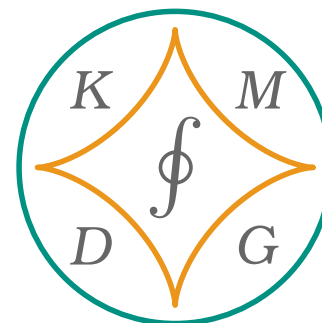


Matematika III: Pracovní listy

Viktor Dubovský, Marcela Jarošová, Jiří Krček, Jitka Krčková, Petra Schreiberová, Petr Volný

Katedra matematiky a deskriptivní geometrie

VŠB - Technická univerzita Ostrava



Předmluva

Studijní materiály jsou určeny pro studenty kombinované i prezenční formy vybraných fakult Vysoké školy báňské - Technické univerzity Ostrava, a to pro předmět Matematika III.

Pracovní listy jsou rozděleny do několika bloků.

Teoretická část (**Listy k přednáškám**) je určena pro přímou výuku v rámci jednotlivých přednášek. Nejedná se o náhradu skript. Proto nedoporučujeme pročitat tento text bez ilustrací, bez vysvětlení významu vět a bez podpůrných příkladů. A také nedoporučujeme považovat tyto materiály za náhradu účasti na výuce.

Blok obsahující řešené příklady (**Řešené příklady**) je zaměřen především na samostudium.

Listy s neřešenými příklady (**Pracovní listy do cvičení**) lze využít v rámci cvičení pro studenty prezenčního studia a pro domácí práci studentů kombinované formy.

Poděkování

Pracovní listy vznikly za finanční podpory projektu FRVŠ 17/2015 „Inovace předmětu Matematika III, FAST, VŠB-TUO“.

Jak pracovat s pracovními listy

Pokud si vytisknete tyto listy, pak si můžete do svého výtisku vpisovat vysvětlující komentáře a příklady, které uslyšíte na přednášce. Nemusíte se tak zdržovat přepisováním definic a vět, ale můžete se lépe soustředit na jejich pochopení.

Příjemně strávený čas s matematikou přeje kolektiv autorů.

Obsah

Listy k přednáškám

1	Kombinatorika	10
1.1	Kombinatorika	11
1.1.1	Úvod	11
1.1.2	Základní pravidla	11
1.2	Výběry bez opakování	12
1.2.1	Permutace bez opakování	12
1.2.2	Variace bez opakování	12
1.2.3	Kombinace bez opakování	12
1.3	Výběry s opakováním	13
1.3.1	Permutace s opakováním	13
1.3.2	Variace s opakováním	13
1.3.3	Kombinace s opakováním	13
1.4	Kombinační čísla	14
1.4.1	Vlastnosti kombinačních čísel	14
2	Pravděpodobnost	15
2.1	Úvod	16
2.1.1	Základní pojmy	16
2.1.2	Typy jevů	16
2.2	Jevové operace a relace	17
2.2.1	Operace s jevy	17
2.2.2	Relace mezi jevy	17
2.2.3	Neslučitelnost jevů	17
2.3	Pravděpodobnost	18
2.3.1	Klasická - Laplaceova - pravděpodobnost	18
2.3.2	Geometrická pravděpodobnost	18
2.3.3	Statistická definice	18
2.3.4	Jevové pole	19
2.3.5	Pravděpodobnost	19
2.3.6	Vlastnosti pravděpodobnosti	19
2.4	Podmíněná pravděpodobnost a nezávislé jevy	20
2.4.1	Podmíněná pravděpodobnost	20
2.4.2	Nezávislé jevy	20
2.5	Úplná pravděpodobnost a Bayesův vzorec	21
2.5.1	Úplná pravděpodobnost	21
2.5.2	Bayesův vzorec	21
2.6	Opakované pokusy	22
2.6.1	Nezávislé	22
2.6.2	Závislé	22
3	Náhodná veličina	23
3.1	Pojem náhodná veličina	24
3.2	Distribuční funkce	25
3.3	Pravděpodobnostní funkce	26
3.4	Funkce hustoty pravděpodobnosti	27
3.5	Číselné charakteristiky	28
3.6	Momentové charakteristiky	29
3.7	Kvantilové a ostatní charakteristiky	30
4	Rozdělení diskrétní náhodné veličiny	31
4.1	Rovnoměrné rozdělení $R(n)$	32
4.2	Alternativní rozdělení $A(p)$	32
4.3	Binomické rozdělení $Bi(n, p)$	33
4.4	Hypergeometrické rozdělení $H(N, M, n)$	34
4.5	Poissonovo rozdělení $Po(\lambda)$	35
5	Rozdělení spojitě náhodné veličiny	36
5.1	Rovnoměrné rozdělení $R(a, b)$	37
5.2	Exponenciální rozdělení $E(\lambda)$	38
5.3	Normální rozdělení $N(\mu, \sigma^2)$	39
5.4	Normované normální rozdělení $N(0, 1)$	40

6	Náhodný vektor	41			
6.1	Náhodný vektor	42			
6.2	Distribuční funkce	42			
6.3	Pravděpodobnostní funkce	42			
6.4	Marginální rozdělení	43			
6.5	Nezávislost složek náhodného vektoru	43			
6.6	Podmíněné rozdělení	44			
6.7	Smíšené momenty	45			
6.8	Podmíněné charakteristiky	45			
6.9	Kovariance	46			
6.10	Korelační koeficient	46			
7	Popisná analýza	47			
7.1	Popisná statistika	48			
7.2	Řazení	49			
7.3	Třídění	49			
7.3.1	Tabulka četností	49			
7.3.2	Četnosti	50			
7.4	Grafická znázornění	50			
7.5	Charakteristiky polohy	51			
7.6	Charakteristiky variability	52			
7.7	Charakteristiky tvaru rozdělení	53			
8	Odhady parametrů	54			
8.1	Bodové odhady	55			
8.2	Intervalové odhady	56			
8.2.1	Intervalový odhad střední hodnoty	57			
8.2.2	Intervalový odhad rozptylu	58			
9	Testování hypotéz	59			
9.1	Statistické hypotézy - úvod	60			
9.1.1	Kroky při testování hypotézy - klasický postup	60			
9.2	Parametrické testy	61			
9.2.1	Test hypotézy o střední hodnotě μ (jednovýběrový t-test)	61			
9.2.2	Test hypotézy o rozptylu σ^2	62			
9.2.3	Dvouvýběrový F-test	63			
9.2.4	Dvouvýběrový t-test	64			
			9.2.5	Studentův test pro párované hodnoty	66
			9.3	Neparametrické testy	67
			9.3.1	Pearsonův test dobré shody (χ^2 -test) pro jeden výběr	67
			9.3.2	Kolmogorovův-Smirnovův test dobré shody pro jeden výběr	68
			9.4	Testy extrémních hodnot	69
			9.4.1	Dixonův test extrémních odchylek	69
			9.4.2	Grubbsův test extrémních odchylek	70
			10	Lineární regrese	71
			10.1	Závislost dvou číselných proměnných	72
			10.2	Regresní analýza	73
			10.3	Metoda nejmenších čtverců	74
			10.4	Verifikace modelu	76
			10.5	Korelační analýza	77
			Řešené příklady	78	
			Řešené příklady – Kombinatorika	79	
			Kombinatorické pravidlo součtu	80	
			Kombinatorické pravidlo součinu	81	
			Permutace bez opakování	82	
			Variace bez opakování	83	
			Kombinace bez opakování	84	
			Variace s opakováním	85	
			Permutace s opakováním	86	
			Kombinace s opakováním	87	
			Vlastnosti kombinačních čísel I	88	
			Vlastnosti kombinačních čísel II	89	
			Úpravy výrazů s kombinačními čísly I	90	
			Úpravy výrazů s kombinačními čísly II	91	
			Rovnice s kombinačními čísly	92	
			Využití více kombinatorických postupů, pravidel a vzorců	93	

Řešené příklady – Pravděpodobnost	94	Normální rozdělení - část 1.	128
Náhodný pokus, náhodný jev	95	Normální rozdělení - část 2.	129
Operace s náhodnými jevy	96	Řešené příklady – Náhodný vektor	130
Relace mezi náhodnými jevy	97	Sdružené rozdělení pravděpodobnosti - část 1.	131
Klasická pravděpodobnost	98	Sdružené rozdělení pravděpodobnosti - část 2.	132
Pravděpodobnost a operace s náhodnými jevy	99	Marginální rozdělení, nezávislost složek	133
Pravděpodobnost a relace mezi náhodnými jevy	100	Marginální rozdělení pravděpodobnosti	134
Geometrická pravděpodobnost	101	Podmíněné rozdělení pravděpodobnosti	135
Podmíněná pravděpodobnost	102	Číselné charakteristiky	136
Podmíněná pravděpodobnost a nezávislé jevy	103	Kovariance a korelační koeficient	137
Úplná pravděpodobnost	104	Řešené příklady – Popisná analýza	138
Úplná pravděpodobnost	105	Třídění, četnosti	139
Bayesova věta	106	Charakteristiky polohy	140
Bayesova věta	107	Charakteristiky variability	141
Opakované pokusy - nezávislé	108	Charakteristiky tvaru	142
Opakované pokusy - nezávislé	109	Řešené příklady – Odhady parametrů	143
Opakované pokusy - závislé	110	Interval spolehlivosti pro střední hodnotu	144
Řešené příklady – Náhodná veličina	111	Interval spolehlivosti pro střední hodnotu	145
Rozdělení diskrétní náhodné veličiny - část 1.	112	Interval spolehlivosti pro střední hodnotu	146
Rozdělení diskrétní náhodné veličiny - část 2.	113	Interval spolehlivosti pro rozptyl, směrodatnou odchylku	147
Číselné charakteristiky	114	Řešené příklady – Testování hypotéz	148
Rozdělení spojité náhodné veličiny	115	Hypotéza o střední hodnotě	149
Číselné charakteristiky	116	Hypotéza o rozptylu	150
Náhodná veličina - část 1.	117	Hypotéza o rovnosti rozptylů	151
Náhodná veličina - část 2.	118	Hypotéza o rovnosti středních hodnot	152
Kvantilové charakteristiky - část 1.	119	Párový t-test	153
Kvantilové charakteristiky - část 2.	120	Testy dobré shody - část 1.	154
Řešené příklady – Rozdělení diskrétní náhodné veličiny	121	Testy dobré shody - část 2.	155
Binomické rozdělení	122	Testy extrémních hodnot	156
Hypergeometrické rozdělení	123	Řešené příklady – Lineární regrese	157
Poissonovo rozdělení	124	Regresní model	158
Řešené příklady – Rozdělení spojité náhodné veličiny	125	Regresní model - verifikace modelu	159
Rovnoměrné rozdělení	126	Regresní model - korelační analýza	160
Exponenciální rozdělení	127		

Pracovní listy do cvičení	161	Náhodná veličina	195
Příklady – Kombinatorika	162	Náhodná veličina	196
Kombinatorické pravidlo součtu	163	Náhodná veličina	197
Kombinatorické pravidlo součinu	164	Příklady – Rozdělení diskrétní náhodné veličiny	198
Kombinatorická pravidla součtu a součinu	165	Binomické rozdělení	199
Permutace bez opakování	166	Hypergeometrické rozdělení	200
Variace bez opakování	167	Poissonovo rozdělení	201
Kombinace bez opakování	168	Binomické, hypergeometrické a Poissonovo rozdělení	202
Variace s opakováním	169	Příklady – Rozdělení spojité náhodné veličiny	203
Permutace s opakováním	170	Rovnoměrné rozdělení	204
Kombinace s opakováním	171	Exponenciální rozdělení	205
Úprava výrazů s kombinačními čísly	172	Normální rozdělení	206
Rovnice s kombinačními čísly	173	Rovnoměrné, exponenciální a normální rozdělení	207
Směs kombinatorických postupů	174	Příklady – Náhodný vektor	208
Směs kombinatorických postupů	175	Marginální rozdělení	209
Směs kombinatorických postupů	176	Podmíněné pravděpodobnosti	210
Příklady – Pravděpodobnost	177	Číselné charakteristiky	211
Náhodný pokus, náhodný jev	178	Nezávislost náhodných veličin	212
Náhodný pokus, náhodný jev	179	Náhodný vektor	213
Klasická pravděpodobnost	180	Příklady – Popisná analýza	214
Geometrická pravděpodobnost	181	Četnosti, histogram	215
Podmíněná pravděpodobnost	182	Číselné charakteristiky	216
Nezávislé jevy	183	Číselné charakteristiky	217
Úplná pravděpodobnost	184	Popisná analýza	218
Bayesův vzorec	185	Popisná analýza	219
Bayesův vzorec	186	Popisná analýza	220
Opakované pokusy	187	Příklady – Odhady parametrů	221
Závislé opakované pokusy	188	Interval spolehlivosti pro střední hodnotu	222
Příklady – Náhodná veličina	189	Interval spolehlivosti pro střední hodnotu	223
Náhodná veličina	190	Interval spolehlivosti pro rozptyl, směrodatnou odchylku	224
Náhodná veličina	191		
Náhodná veličina	192		
Náhodná veličina	193		
Náhodná veličina	194		

Příklady – Testování hypotéz	225	Příklady – Lineární regrese	232
Hypotéza o střední hodnotě	226	Regresní model	233
Hypotéza o rozptylu	227	Regresní a korelační analýza	234
Hypotéza o rovnosti středních hodnot	228	Verifikace modelu	235
Párový t-test	229	Regresní model	236
Testy dobré shody	230		
Testy extrémních hodnot	231	Literatura	237

Matematika III: Listy k přednáškám

Viktor Dubovský, Marcela Jarošová, Jiří Krček, Jitka Krčková, Petra Schreiberová, Petr Volný

Katedra matematiky a deskriptivní geometrie

VŠB - Technická univerzita Ostrava

Kapitola 1

Kombinatorika

11 - Kombinatorika

Řešené příklady: 80,81
Příklady: 163,164

1.1 Kombinatorika

1.1.1 Úvod

Kombinatorika je část matematiky zabývající se vlastnostmi konečných množin a jejich podmnožin. Speciálně zkoumá počty výběrů ze základní množiny, přičemž tyto výběry splňují dané podmínky.

Při takových výběrech rozlišujeme, zda záleží na pořadí, v němž k výběru dojde, či nikoliv. V prvním případě hovoříme o *uspořádaných* a ve druhém o *neuspořádaných* výběrech, skupinách, *n*-ticích apod.

Dalším posuzovaným hlediskem je, zda se prvky mohou ve výběrech opakovat. Takto rozlišujeme výběry *s opakováním* a *bez opakování*.

1.1.2 Základní pravidla

Mnoho z kombinatorických úloh lze vyřešit za pomoci dvou jednoduchých pravidel, a to *pravidla součinu* a *pravidla součtu*.

Věta 1.1.1: Kombinatorické pravidlo součinu Počet všech uspořádaných *k*-tic, jejichž první člen lze vybrat n_1 způsoby, druhý n_2 , atd., až *k*-tý člen lze vybrat n_k způsoby, je roven součinu $n_1 \cdot n_2 \cdot n_3 \cdots n_k$.

Věta 1.1.2: Kombinatorické pravidlo součtu Necht' $A_1, A_2, \dots, A_k$ jsou konečné množiny, které mají $p_1, p_2, \dots, p_k$ prvků, a které jsou po dvou vzájemně disjunktní, tj. nemají společný prvek, pak počet prvků sjednocení $A_1 \cup A_2 \cup \dots \cup A_k$ je roven součtu $p_1 + p_2 + \dots + p_k$.

12 - Výběry bez opakování

Řešené příklady: 82,83,84
Příklady: 166,167,168

1.2 Výběry bez opakování

Ze základní množiny vybíráme skupiny prvků, přičemž prvek se ve výběru může objevit nejvýše jednou. Celou situaci si lze představit tak, že prvek, který dle požadavků ze základní množiny vybereme, do ní nevracíme zpět. Rozlišujeme tři základní typy takových výběrů, dva uspořádané, *permutace*, *variace* a neuspořádané *kombinace*.

1.2.1 Permutace bez opakování

Z n prvkové základní množiny vybíráme skupinu n prvků, přitom dbáme na jejich pořadí a na to, aby se prvky neopakovaly. Takový výběr nazýváme **permutací bez opakování** a značíme $P(n)$. Platí

Věta 1.2.3: Počet uspořádaných n -tic utvořených bez opakování z n prvků je roven

$$P(n) = n!.$$

1.2.2 Variace bez opakování

Opět z n prvků vybíráme, sledujeme pořadí a neopakujeme prvky. Avšak ve výběru nejsou obsaženy všechny prvky základní množiny, ale jen k z nich. Tento výběr je **variací k -té třídy z n prvků bez opakování**, značíme $V_k(n)$ a platí

Věta 1.2.4: Počet uspořádaných k -tic utvořených bez opakování z n prvků je roven

$$V_k(n) = \frac{n!}{(n-k)!}.$$

1.2.3 Kombinace bez opakování

Oproti předchozím výběrům nebereme zřetel na pořadí výběru k prvků, který bez opakování provádíme z n prvků základní množiny. Nazýváme **kombinací k -té třídy z n prvků bez opakování**, $C_k(n)$.

Věta 1.2.5: Počet neuspořádaných k -tic utvořených bez opakování z n prvků je roven

$$C_k(n) = \binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Poznámka: **Faktoriál** přirozeného čísla n , je definován jako součin

$$n! = \prod_{i=1}^n i,$$

tj. $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$. Pro $n = 0$ je dodefinován $0! = 1$.

Počet $C_k(n)$ nazýváme **kombinační číslo** a používáme pro něj symbol $\binom{n}{k}$.

13 - Výběry s opakováním

Řešené příklady: 85,86,87
Příklady: 169,170,171

1.3 Výběry s opakováním

V tomto výběru ze základní množiny se prvek může opakovat. Lze si představit, že základní skupina obsahuje prvky, které nedokážeme rozlišit, nebo tak, že prvek, který dle zadání ze základní množiny vybereme, do ní opět vrátíme. I zde jsou tři základní typy výběrů, dva uspořádané, *permutace*, *variace* a neuspořádané *kombinace*.

1.3.1 Permutace s opakováním

Vybíráme uspořádanou skupinu n prvků. Přitom sama základní množina má $n = n_1 + \dots + n_k$ prvků, kde k značí počet jejich podmnožin se vzájemně nerozlišitelnými prvky. Tento výběr nazýváme **permutací s opakováním** a značíme $P'(n_1, n_2, \dots, n_k)$. Platí

Věta 1.3.6: Počet uspořádaných n -tic utvořených s opakováním z $n = n_1 + n_2 + \dots + n_k$ prvků je roven

$$P'(n_1, n_2, \dots, n_k) = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}.$$

1.3.2 Variace s opakováním

Z n prvků základní množiny vybíráme k , sledujeme pořadí výběru, opakování je povoleno. Oproti permutaci zde nemusí být ve výběru obsaženy všechny prvky základní množiny. Takový výběr nazveme **variací k -té třídy z n prvků s opakováním**, značíme $V'_k(n)$ a platí

Věta 1.3.7: Počet uspořádaných k -tic utvořených s opakováním z n prvků je roven

$$V'_k(n) = n^k.$$

1.3.3 Kombinace s opakováním

Z n různých typů prvků tvoříme k -prvkovou skupinu. Nezáleží na pořadí v ní a prvky (typy) se mohou opakovat. Jde o **kombinaci k -té třídy z n prvků s opakováním**, $C'_k(n)$.

Věta 1.3.8: Počet neuspořádaných k -tic utvořených s opakováním z n prvků je roven

$$C'_k(n) = C_k(n + k - 1) = \binom{n + k - 1}{k} = \frac{(n + k - 1)!}{k!(n - 1)!}.$$

14 - Vlastnosti kombinačních čísel

Řešené příklady: 88-92
Příklady: 172,173**1.4 Kombinační čísla**

V několika předchozích vzorcích se objevila kombinační čísla. Jejich využití se však neomezuje pouze a jen na kombinatoriku, což můžeme ukázat na příkladě binomické věty, s níž jste se již jistě setkali.

$$(a + b)^n = \sum_{i=0}^n \binom{n}{i} a^i b^{n-i}, \quad (a - b)^n = \sum_{i=0}^n (-1)^i \binom{n}{i} a^i b^{n-i}.$$

I vzhledem k jejich rozšíření si zde shrneme základní vlastnosti.

1.4.1 Vlastnosti kombinačních čísel**Definice**

$$\binom{n}{k} = \frac{n!}{(n-k)! k!}$$

Výpočet

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)!}{k!} = \frac{\overbrace{n \cdot (n-1) \cdots (n-k+1)}^{k \text{ členů}}}{\underbrace{k \cdot (k-1) \cdots 1}_{k \text{ členů}}}$$

Okrajové vlastnosti

$$\binom{n}{0} = \binom{n}{n} = 1$$

Symetrie

$$\binom{n}{k} = \binom{n}{n-k}$$

Součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Kapitola 2

Pravděpodobnost

16 - Teorie pravděpodobnosti

Řešené příklady: 95,96,97
Příklady: 178,179

2.1 Úvod

Uvažujme provedení pokusu. Jeho výsledek může být jednoznačně určen podmínkami, za nichž pokus provádíme a při jeho opakování dostáváme vždy stejný výsledek. Takový pokus nazýváme *deterministický*. Ovšem existují také pokusy, jejichž výsledek se i při dodržení postupů a zachování předpokladů může v jednotlivých opakováních lišit. Tyto pokusy nazýváme *stochastickými*. A právě takové pokusy jsou předmětem zkoumání teorie pravděpodobnosti.

2.1.1 Základní pojmy

Náhodný pokus je proces, jehož výsledek nelze předem přesně určit, i když jsou známy všechny podmínky a předpoklady jeho provedení.

Základní prostor je množina všech možných výsledků zkoumaného pokusu, značíme jej Ω .

Elementární jev je každý jednotlivý výsledek náhodného pokusu, tj. každý prvek základního prostoru, značíme $\omega_1, \omega_2, \dots, \omega_n$.

Náhodný jev je každá (i víceprvková) podmnožina základního prostoru Ω .

2.1.2 Typy jevů

Jev jistý je takový jev, který nastává v každém opakování pokusu, takovým jevem je Ω .

Jev nemožný je jev, který nikdy nastat nemůže, značíme jej $\emptyset$.

Opačný jev k jevu A je jev, který nastupuje, když nenastupuje jev A , značíme $\bar{A}$.

17 - Jevové operace a relace

Řešené příklady: 95,96,97
Příklady: 178,179



2.2 Jevové operace a relace

Při zkoumání výsledků náhodného pokusu se lze setkat s případy, kdy náhodné jevy nastávají společně či naopak, nastoupení jednoho vylučuje nastoupení jiných výsledků. V takových situacích pomáhají operace, které s jevy můžeme provádět. Dále pak popis relací (vztahů) mezi jevy.

2.2.1 Operace s jevy

Sjednocení jevů A, B je jev $A \cup B$, který nastane, nastal-li alespoň jeden z jevů A, B , zapisujeme $A \cup B = \{\forall \omega \in \Omega : \omega \in A \vee \omega \in B\}$.

Průnik jevů A, B je jev $A \cap B$, který nastává, nastanou-li oba jevy A, B společně, zápis $A \cap B = \{\forall \omega \in \Omega : \omega \in A \wedge \omega \in B\}$.

Rozdíl jevů A, B je jev $A - B$, při němž nastoupí jev A a zároveň nenastane jev B , $A - B = \{\forall \omega \in \Omega : \omega \in A \wedge \omega \notin B\}$.

2.2.2 Relace mezi jevy

Podjev jevu A je takový jev B , z jehož nastoupení vyplývá nastoupení jevu A , zapisujeme $B \subseteq A$, a platí $B \subseteq A \Leftrightarrow \{\forall \omega \in \Omega : \omega \in B \Rightarrow \omega \in A\}$.

Rovnost jevů A, B , nastává, právě když oba jevy nastupují vždy společně, značíme $A = B$ a $A = B \Leftrightarrow \{\forall \omega \in \Omega : \omega \in B \Leftrightarrow \omega \in A\}$.

2.2.3 Neslučitelnost jevů

Neslučitelné jevy A, B jsou jevy, které nemohou nastat společně, to znamená, že nastoupení A vylučuje B a naopak. Platí pro ně $A \cap B = \emptyset$.

Systém neslučitelných jevů je množina jevů $A_i, i = 1, \dots, n$, pro které platí, že jsou po dvou neslučitelné, tj. že $A_i \cap A_j = \emptyset$, kdykoli je $i \neq j$.

Úplný systém neslučitelných jevů je systém neslučitelných jevů, jejichž sjednocením je celý základní prostor Ω , tj. $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$.

Poznámka: Množinový zápis a množinové operace jsou zvoleny, protože náhodné jevy chápeme jako podmnožiny základního prostoru Ω .

K operaci *sjednocení* $\cup$ přísluší logická spojka „nebo“ $\vee$.

Operaci *průnik* $\cap$ přísluší logická spojka „a zároveň“ $\wedge$.

18 - Pravděpodobnost - definice, typy

Řešené příklady: 98,101
Příklady: 180,181

2.3 Pravděpodobnost

Jak bylo řečeno, výsledek náhodného pokusu nelze předem určit. Můžeme se však pokusit odhadnout, změřit, spočítat, jak velká je možnost, že výsledkem bude ten který náhodný jev. Tento „odhad“ nazýváme pravděpodobností jevu. Nejběžněji používáme tři typy postupů jejího určení, a to *klasickou*, *geometrickou* a *statistickou* definici pravděpodobnosti.

2.3.1 Klasická - Laplaceova - pravděpodobnost

Definice 2.3.9: Necht' $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ je konečný základní prostor tvořený n různými, vzájemně se vylučujícími elementárními jevy, které jsou stejně možné. Pak je **pravděpodobnost nastoupení jevu** A , který tvoří $m \leq n$ z těchto elementárních jevů, rovna podílu

$$P(A) = \frac{m}{n}.$$

2.3.2 Geometrická pravděpodobnost

Definice 2.3.10: Necht' základní prostor Ω tvoří geometrický útvar $\mathcal{O}$ a náhodný jev A útvar $\mathcal{A}$, přičemž platí $\mathcal{A} \subseteq \mathcal{O}$, pak **pravděpodobnost nastoupení jevu** A lze spočítat jako podíl

$$P(A) = \frac{\mu(\mathcal{A})}{\mu(\mathcal{O})},$$

kde $\mu(\mathcal{A})$, resp. $\mu(\mathcal{O})$ značí *měru* útvaru $\mathcal{A}$, resp. $\mathcal{O}$.

2.3.3 Statistická definice

Definice 2.3.11: Necht' A je hromadný jev. Nastal-li tento jev v n pokusech právě f_n krát, definujeme jeho **pravděpodobnost** jako

$$P(A) = \lim_{n \rightarrow \infty} \frac{f_n}{n}.$$

Číslo f_n se nazývá **absolutní četnost** jevu A , a $\frac{f_n}{n}$ **relativní četnost** jevu A při n pokusech.

Poznámka:

Obvykle je $\mathcal{O}$ konečná, uzavřená oblast na přímce, v rovině či v prostoru a $\mathcal{A}$ její uzavřená podmnožina, a jejich měrou rozumíme délku, plochu či objem.

Hromadným jevem rozumíme jev, který lze za daných podmínek libovolně krát zopakovat, či jej lze pozorovat na objektech stejného typu s hromadným výskytem.

19 - Pravděpodobnost - axiomatická definice a vlastnosti

V předchozím bylo uvedeno několik definic pravděpodobnosti, všechny však do značné míry vycházely z „intuitivního“ rozboru situace. Uveďme tedy *axiomatickou* definici pravděpodobnosti dle A.N. Kolmogorova.

2.3.4 Jevové pole

Definice 2.3.12: *Jevové pole* $\mathfrak{D}$ je systém různých podmnožin základního prostoru Ω s těmito vlastnostmi:

- $\Omega \in \mathfrak{D}$, (jev jistý patřící do $\mathfrak{D}$),
- pro jevy $A, B \in \mathfrak{D}$ platí také $\bar{A} \in \mathfrak{D}, A \cup B \in \mathfrak{D}, A \cap B \in \mathfrak{D}, A - B \in \mathfrak{D}$ (uzavřenost na jevové operace).

2.3.5 Pravděpodobnost

Definice 2.3.13: *Pravděpodobností jevu* A z jevového pole $\mathfrak{D}$ je reálné číslo $P(A)$ pro něž platí:

- $P(A) \geq 0$ (axiom nezápornosti),
- $P(\Omega) = 1$ (axiom jednotky),
- $P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$, kde $A_1, A_2, \dots, A_i \in \mathfrak{D}$ tvoří systém navzájem neslučitelných jevů (axiom aditivity).

2.3.6 Vlastnosti pravděpodobnosti

- $P(\emptyset) = 0$
- $P(\bar{A}) = 1 - P(A)$
- Jestliže $A \subseteq B$, pak platí $P(A) \leq P(B)$.
- Jestliže $A \subseteq B$, pak platí $P(B - A) = P(B) - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

20 - Podmíněná pravděpodobnost a nezávislé jevy

Řešené příklady: 102,103
Příklady: 182,183

2.4 Podmíněná pravděpodobnost a nezávislé jevy

Z předchozího již víme, že předem nelze s určitostí zjistit, zda bude jev A výsledkem náhodného pokusu. Lze však spočítat pravděpodobnost $P(A)$ jeho nastoupení. Otázkou však je, jak a zda vůbec se tato pravděpodobnost $P(A)$ změní, pokud víme, že nastal jev B .

2.4.1 Podmíněná pravděpodobnost

Definice 2.4.14: Pravděpodobnost nastoupení jevu A za předpokladu, že nastal jev B , značíme $P(A|B)$ a nazýváme **podmíněná pravděpodobnost** jevu A jevem B . Spočteme ji dle vzorce

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ kde } P(B) \neq 0.$$

- Pro jev B s $P(B) = 0$ podmíněnou pravděpodobnost nedefinujeme.
- Platí-li $A \cap B = \emptyset$, pak $P(A|B) = 0$.
- Vzorec využíváme také k výpočtu $P(A \cap B) = P(A|B) \cdot P(B)$.

2.4.2 Nezávislé jevy

Definice 2.4.15: Jevy pro něž platí $P(A|B) = P(A)$, nazýváme **nezávislé jevy**.

- Platí-li $P(A|B) = P(A)$, platí také $P(B|A) = P(B)$.
- Jevy jsou nezávislé, právě když $P(A \cap B) = P(A) \cdot P(B)$.
- Jsou-li A, B neslučitelné jevy s nenulovými pravděpodobnostmi, pak jsou závislé.

Definice 2.4.16: Jevy $A_1, \dots, A_n$ jsou **vzájemně nezávislé**, jestliže pro každou jejich podmnožinu platí, že pravděpodobnost průniku v něm zastoupených jevů je rovna součinu jejich pravděpodobností.

21 - Úplná pravděpodobnost a Bayesův vzorec

Řešené příklady: 104,105,106,107
Příklady: 184,185,186

2.5 Úplná pravděpodobnost a Bayesův vzorec

Podmíněnou pravděpodobností byla zodpovězena otázka, jak a zda se změní pravděpodobnost jevu A , nastane-li jev B . Nyní odpovíme na otázku, jaká je pravděpodobnost jevu A , známe-li jeho pravděpodobnost podmíněnou nastoupením předpokladů, hypotéz H_i a jejich pravděpodobností. Je přirozené ptát se i „opačnou“ otázkou, tedy jaká je pravděpodobnost, že byla splněna (právě) hypotéza H_i , víme-li, že jev A skutečně nastal.

2.5.1 Úplná pravděpodobnost

Věta 2.5.17: Necht' $H_1, H_2, \dots, H_n$ tvoří úplný systém neslučitelných jevů. Potom pro každý náhodný jev $A \subseteq \Omega$ platí

$$P(A) = \sum_{i=1}^n P(A|H_i) P(H_i).$$

2.5.2 Bayesův vzorec

Věta 2.5.18: Necht' $H_1, H_2, \dots, H_n$ tvoří úplný systém neslučitelných jevů. Potom pro každý náhodný jev $A \subseteq \Omega$ a pro všechna $k = 1, 2, \dots, n$ platí

$$P(H_k|A) = \frac{P(A|H_k) P(H_k)}{\sum_{i=1}^n P(A|H_i) P(H_i)}.$$

22 - Opakované pokusy

Řešené příklady: 108,109,110
Příklady: 187,188

2.6 Opakované pokusy

Pokud budeme náhodný pokus vícekrát opakovat, budou se mezi výsledky pravděpodobnější náhodné jevy vyskytovat častěji. S větším počtem opakování pokusu však bude také narůstat složitost výpočtu četnosti, resp. pravděpodobnosti četnosti, výskytu daného jevu. V případě série *nezávislých* pokusů odpovídá na tuto otázku *Bernoulliho* vzorec, pro pokusy závislé využijeme vzorec *kombinatorický*.

2.6.1 Nezávislé

Sérii náhodných pokusů, v nichž není pravděpodobnost $P(A)$ ovlivněna výsledky předcházejících pokusů, nazveme (*Bernoulliho*) *posloupností nezávislých pokusů*.

Věta 2.6.19: Nechť A je náhodný jev, jehož pravděpodobnost nastoupení v náhodném pokusu je $P(A) = p$. Pravděpodobnost jevu A_k , že jev A nastane v posloupnosti n nezávislých pokusů právě k -krát je

$$P(A_k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

2.6.2 Závislé

Pokud pravděpodobnost nastoupení jevu A v sérii pokusů ovlivňují výsledky pokusů předcházejících, hovoříme o *posloupnosti závislých jevů*.

Věta 2.6.20: Nechť v souboru n prvků má k , $0 \leq k \leq n$, z nich určitou vlastnost a $n - k$ tuto vlastnost nemá. Ze souboru vybíráme postupně j prvků, které nevracíme. Označme A_i jev, odpovídající tomu, že v tomto výběru má právě i prvků sledovanou vlastnost. Pravděpodobnost $P(A_i)$ vypočteme podle vzorce

$$P(A_i) = \frac{\binom{k}{i} \cdot \binom{n-k}{j-i}}{\binom{n}{j}}.$$

Kapitola 3

Náhodná veličina

24 - Náhodná veličina

3.1 Pojem náhodná veličina

Definice 3.1.21: Náhodná veličina X je reálná funkce definovaná na množině všech elementárních jevů, která každému jevu přiřadí reálné číslo.

Podle oboru hodnot, kterých může náhodná veličina (NV) nabývat rozdělujeme náhodné veličiny na:

- **diskrétní (DNV)** - obor hodnot je konečná nebo nekonečná posloupnost,
- **spojité (SNV)** - obor hodnot je otevřený nebo uzavřený interval.

Pravidlo, které každé hodnotě (popř. každému intervalu) přiřazuje pravděpodobnost, že náhodná veličina nabude této hodnoty (popř. hodnoty z tohoto intervalu), nazýváme **rozdělením pravděpodobnosti náhodné veličiny**.

Popis rozdělení pravděpodobnosti provádíme nejčastěji pomocí distribuční a frekvenční funkce a pomocí číselných charakteristik.

25 - Náhodná veličina

Řešené příklady: 112,113,114
Příklady: 190,191

3.2 Distribuční funkce

Definice 3.2.22: Distribuční funkce náhodné veličiny X je reálná funkce $F(x) = P(X < x)$, definovaná na množině reálných čísel.

Vlastnosti distribuční funkce:

- $0 \leq F(x) \leq 1$,
- $P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$ pro $x_1 < x_2$,
- $F(x)$ je neklesající funkce:

$$\forall x_1, x_2 \in \mathbb{R} : x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2),$$

- $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$,
- $F(x)$ je zleva spojitá: $\forall a \in \mathbb{R} : \lim_{x \rightarrow a^+} F(x) = F(a)$.

Distribuční funkce je definována shodně pro diskrétní i spojitou NV.

26 - Náhodná veličina - frekvenční funkce

Řešené příklady: 112,113,114
Příklady: 190,191,192,193

3.3 Pravděpodobnostní funkce

Frekvenční funkce je definována pro diskrétní a spojitou náhodnou veličinu odlišně.

Pro DNV se užívá termín pravděpodobnostní funkce náhodné veličiny a značí se $p(x)$.

Definice 3.3.23: Pravděpodobnostní funkcí náhodné veličiny X nazýváme funkci $p(x) = P(X = x)$.

Vlastnosti pravděpodobnostní funkce:

- $p(x_i) \geq 0$ pro všechny hodnoty z oboru hodnot,
- $\sum_{i=1}^n p(x_i) = 1$,
- $F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$.

Pravděpodobnostní funkci lze zadat: předpisem, tabulkou, grafem.

27 - Náhodná veličina - frekvenční funkce

Řešené příklady: 115
Příklady: 194, 195

3.4 Funkce hustoty pravděpodobnosti

V případě SNV se pro frekvenční funkci užívá názvu hustota pravděpodobnosti náhodné veličiny a značí se $f(x)$.

Definice 3.4.24: Hustota pravděpodobnosti $f(x)$ spojitě náhodné veličiny X definované na intervalu $\langle a, b \rangle$ je nezáporná reálná funkce definovaná vztahem:

$$f(x) = \lim_{h \rightarrow 0} \frac{P(x \leq X < x + h)}{h},$$

kde pro $x \notin \langle a, b \rangle$ je $f(x) = 0$; $x, x + h \in \langle a, b \rangle$.

Vlastnosti funkce hustoty:

- $\forall x \in \mathbb{R} : f(x) \geq 0$,
- $f(x) = F'(x)$ ($F(x)$ je primitivní funkcí $f(x)$),
- $\lim_{x \rightarrow \infty} f(x) = 0, \lim_{x \rightarrow -\infty} f(x) = 0$,
- $\int_{-\infty}^{\infty} f(x) = 1$,
- $P(x_1 \leq X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$.

28 - Náhodná veličina

Řešené příklady: 116,117,118
Příklady: 195,197

3.5 Číselné charakteristiky

Je výhodné shrnout informace o náhodné veličině do několika čísel, které ji dostatečně charakterizují. Tato čísla nazýváme **číselné charakteristiky** a dělíme je:

a) podle způsobu konstrukce na charakteristiky:

- momentové,
- kvantilové,
- ostatní,

b) podle toho, které vlastnosti rozdělení pravděpodobnosti charakterizují:

- charakteristiky polohy - určují jakýsi „střed“ kolem něhož kolísají náhodné veličiny X (střední hodnota, modus, medián, kvantily),
- charakteristiky rozptýlu - ukazují rozptýlenost hodnot náhodné veličiny kolem střední hodnoty (rozptyl, směrodatná odchylka),
- charakteristiky šikmosti a špičatosti - charakterizují průběh rozdělení náhodné veličiny X .

29 - Náhodná veličina - číselné charakteristiky

Řešené příklady: 116, 117, 118
Příklady: 195

3.6 Momentové charakteristiky

1. Počáteční (obecný) moment k -tého stupně μ_k

$$\mu_k = \begin{cases} \sum_{i=-\infty}^{\infty} x_i^k p(x_i) & \text{pro diskrétní náhodnou veličinu} \\ \int_{-\infty}^{\infty} x^k f(x) dx & \text{pro spojitou náhodnou veličinu} \end{cases}$$

2. Centrální moment k -tého stupně ν_k

$$\nu_k = \begin{cases} \sum_{i=-\infty}^{\infty} (x_i - \mu)^k p(x_i) & \text{pro diskrétní náhodnou veličinu} \\ \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx & \text{pro spojitou náhodnou veličinu} \end{cases}$$

kde $\mu = \mu_1$ je počáteční moment 1. stupně náhodné veličiny X .Praktický význam mají čtyři momentové charakteristiky: $\mu_1, \nu_2, \nu_3, \nu_4$.**První počáteční moment μ_1** představuje **střední hodnotu** náhodné veličiny X , ozn. $E(X)$ či μ .**Druhý centrální moment ν_2** představuje **rozpyl** (disperzi, variaci) náhodné veličiny X , ozn. $D(X)$ či σ^2 .

- Výpočet rozptylu náhodné veličiny X se obvykle počítá podle vztahu $D(X) = E(X^2) - [E(X)]^2 = \mu_2 - \mu^2$.
- Druhá odmocnina z rozptylu se nazývá **směrodatná odchylka** a značí se σ : $\sqrt{D(X)} = \sqrt{\sigma^2} = \sigma$.

Třetí centrální moment ν_3 slouží k určení **koefficientu šikmosti** (asymetrie), ozn. $\bar{\nu}_3 = A$.

$$A = \bar{\nu}_3 = \frac{\nu_3}{\sigma^3}$$

Čtvrtý centrální moment ν_4 slouží k výpočtu **koefficientu špičatosti** (excesu), ozn. $\bar{\epsilon}$.

$$\bar{\epsilon} = \bar{\nu}_4 = \frac{\nu_4}{\sigma^4} - 3$$

Výpočet centrálních momentů lze provádět podle výše uvedeného, a nebo s využitím μ_k :

- $\nu_2 = \mu_2 - \mu_1^2$,
- $\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$,
- $\nu_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4$,
- $\vdots$
- $\nu_k = \binom{k}{0}\mu_k\mu_1^0 - \binom{k}{1}\mu_{k-1}\mu_1^1 + \binom{k}{2}\mu_{k-2}\mu_1^2 + \dots + (-1)^k \binom{k}{k}\mu_1^k$.

30 - Náhodná veličina - číselné charakteristiky

Řešené příklady: 119
Příklady: 196,197

3.7 Kvantilové a ostatní charakteristiky

Definice 3.7.25: Necht' $F(x)$ je distribuční funkce spojité náhodné veličiny X . Pak hodnota x_p , pro kterou platí $F(x_p) = p$, kde $p \in \langle 0; 1 \rangle$, se nazývá **p -kvantil**.

Poznámka: p -kvantil dělí plochu pod grafem funkce hustoty v poměru $p : (1 - p)$.

Nejpoužívanější kvantily:

- **kvartily:** $x_{0,25}; x_{0,5}; x_{0,75}$
- rozdělí obor možných hodnot na čtyři části, v nichž se náhodná veličina nachází s pravděpodobností 0,25,
 $x_{0,5} = Me \dots$ **medián** : dělí plochu pod křivkou hustoty pravděpodobnosti na dvě stejné části,
- **decily:** $x_{0,1}; x_{0,2}; \dots; x_{0,9}$
- rozdělí obor možných hodnot na deset částí se stejnou pravděpodobností výskytu,
- **percentily:** $x_{0,01}; x_{0,02}; \dots; x_{0,99}$
- rozdělí obor možných hodnot na sto částí se stejnou pravděpodobností výskytu.

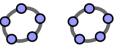
Modus: Mo - je hodnota, v níž nabývá frekvenční funkce maxima:

- u diskrétní náhodné veličiny je to hodnota, v níž pravděpodobnostní funkce $p(x)$ dosahuje maxima,
- u spojité náhodné veličiny je to hodnota, v níž funkce hustoty $f(x)$ nabývá lokálního maxima.

Kapitola 4

Rozdělení diskrétní náhodné veličiny

32 - Rovnoměrné a alternativní rozdělení



4.1 Rovnoměrné rozdělení $R(n)$

- popisuje náhodnou veličinu, která může nabývat n stejně pravděpodobných hodnot.

Definice 4.1.26: Náhodná veličina X má **rovnoměrné rozdělení** $R(n)$, právě když je její pravděpodobnostní funkce dána rovnicí

$$p(x) = \frac{1}{n},$$

kde n je počet možných výsledků.

4.2 Alternativní rozdělení $A(p)$

- popisuje náhodnou veličinu, která může nabývat pouze dvou hodnot, a to 1 v případě, že sledovaný jev nastal (pravděpodobnost p) a 0 v případě, že sledovaný jev nenastal (pravděpodobnost $1 - p$).

Definice 4.2.27: Náhodná veličina X má **alternativní rozdělení** $A(p)$, právě když je její pravděpodobnostní funkce dána rovnicí

$$p(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \end{cases}.$$

Vlastnosti

- $E(X) = p$
- $D(X) = p(1 - p)$

33 - Binomické rozdělení

Řešené příklady: 122
Příklady: 199, 202**4.3 Binomické rozdělení $Bi(n, p)$**

- popisuje četnost výskytu náhodného jevu v n nezávislých pokusech, v nichž má jev stále stejnou pravděpodobnost p .

Definice 4.3.28: Náhodná veličina X má **binomické rozdělení** $Bi(n, p)$, právě když je její pravděpodobnostní funkce dána rovnicí

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n,$$

kde n je počet nezávislých pokusů a p je pravděpodobnost výskytu sledovaného jevu v každém pokusu.

Vlastnosti

- $E(X) = np$
- $D(X) = np(1-p)$

Poznámka: Pro $n = 1$ jde o alternativní rozdělení, tj. $A(p) \sim Bi(1, p)$.

EXCEL:

$$P(X = x) = \text{BINOM.DIST}(x; n; p; 0)$$

$$P(X \leq x) = \text{BINOM.DIST}(x; n; p; 1)$$

34 - Hypergeometrické rozdělení

Řešené příklady: 123
Příklady: 200,202

4.4 Hypergeometrické rozdělení $H(N, M, n)$

- popisuje četnost výskytu sledovaného jevu v n opakovaných závislých pokusech - výběr bez vracení.

Definice 4.4.29: Náhodná veličina X má **hypergeometrické rozdělení** $H(N, M, n)$, právě když je její pravděpodobnostní funkce dána rovnicí

$$p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, \dots, n,$$

kde

- N je celkový počet prvků základního souboru,
- M je počet prvků v základním souboru, které mají sledovanou vlastnost,
- n je počet prvků ve výběru,
- x je počet prvků ve výběru, které mají sledovanou vlastnost.

Vlastnosti

- $E(X) = \frac{nM}{N}$
- $D(X) = \frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$

EXCEL:

$$P(X = x) = \text{HYPGEOM.DIST}(x; n; M; N; 0)$$

$$P(X \leq x) = \text{HYPGEOM.DIST}(x; n; M; N; 1)$$

35 - Poissonovo rozdělení

Řešené příklady: 124
Příklady: 201, 202 4.5 Poissonovo rozdělení $Po(\lambda)$

- popisuje četnost výskytu náhodného (Poissonovského) jevu v daném intervalu (časovém, délkovém, prostorovém).

Definice 4.5.30: Náhodná veličina X má **Poissonovo rozdělení** $Po(\lambda)$, právě když je její pravděpodobnostní funkce dána rovnicí

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, \dots,$$

kde λ je střední hodnota počtu výskytů sledovaného jevu v daném intervalu.

Vlastnosti

- $E(X) = \lambda$
- $D(X) = \lambda$
- $A = \frac{1}{\sqrt{\lambda}}$
- $\bar{e} = \frac{1}{\lambda}$

Poznámka: Binomické rozdělení $Bi(n, p)$ lze pro velká n a p blížící se k nule aproximovat Poissonovým rozdělením s parametrem $\lambda = np$, tj.

$$Bi(n, p) \sim Po(np) \quad \text{pro } n \rightarrow \infty \text{ a } p \rightarrow 0.$$

Dále pro velké λ se hodnoty Poissonova rozdělení blíží k hodnotám rozdělení normálního.

EXCEL:

$$P(X = x) = \text{POISSON.DIST}(x; \lambda; 0)$$

$$P(X \leq x) = \text{POISSON.DIST}(x; \lambda; 1)$$

Poissonovský jev:

- výskyt jevu v daném časovém (prostorovém) intervalu není ovlivněn tím, co se stalo jindy (jinde),
- v každém bodě intervalu je pravděpodobnost výskytu jevu na jeho okolí stejná,
- v jednom časovém okamžiku (bodě v prostoru) nemohou nastat dva jevy současně.

Kapitola 5

Rozdělení spojité náhodné veličiny

37 - Rovnoměrné rozdělení

Řešené příklady: 126
Příklady: 204, 2075.1 Rovnoměrné rozdělení $R(a, b)$

- popisuje náhodnou veličinu, jejíž realizace vyplňují interval $\langle a, b \rangle$ se stejnou možností výskytu v každém bodě.

Definice 5.1.31: Náhodná veličina X má **rovnoměrné rozdělení** $R(a, b)$, právě když je její hustota pravděpodobnosti dána funkcí

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in \langle a, b \rangle \\ 0 & x \notin \langle a, b \rangle \end{cases}.$$

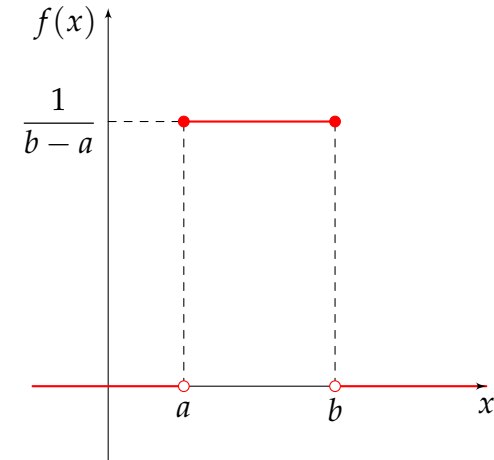
Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, a) \\ \frac{x-a}{b-a} & x \in \langle a, b \rangle \\ 1 & x \in (b, +\infty) \end{cases}$$

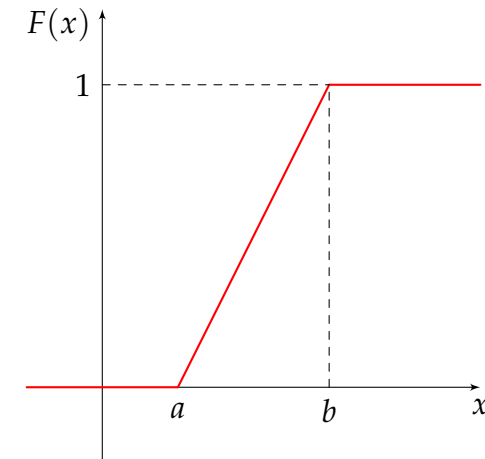
Vlastnosti

- $E(X) = \frac{a+b}{2}$
- $D(X) = \frac{(b-a)^2}{12}$

Graf hustoty pravděpodobnosti:



Graf distribuční funkce:



38 - Exponenciální rozdělení

Řešené příklady: 127
Příklady: 205, 2075.2 Exponenciální rozdělení $E(\lambda)$

- popisuje dobu čekání na (Poissonovský) náhodný jev, resp. délku intervalu (časového nebo délkového) mezi dvěma takovými jevy.

Definice 5.2.32: Náhodná veličina X má **exponenciální rozdělení** $E(\lambda)$, právě když je její hustota pravděpodobnosti dána funkcí

$$f(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ \lambda e^{-\lambda x} & x \in \langle 0, +\infty \rangle \end{cases},$$

kde λ je převrácená hodnota střední hodnoty doby čekání na sledovaný jev.

Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ 1 - e^{-\lambda x} & x \in \langle 0, +\infty \rangle \end{cases}$$

Vlastnosti

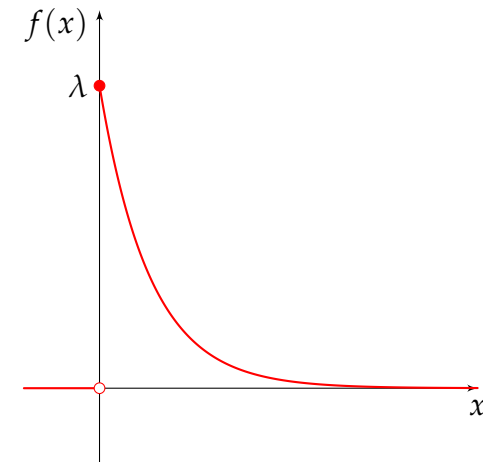
- $E(X) = \frac{1}{\lambda}$
- $D(X) = \frac{1}{\lambda^2}$

EXCEL:

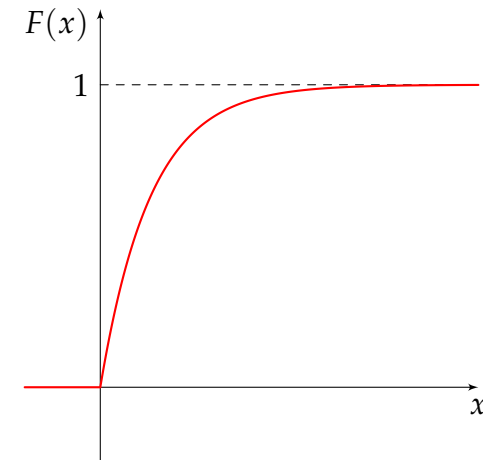
$$f(x) = \text{EXPON.DIST}(x; \lambda; 0)$$

$$F(x) = \text{EXPON.DIST}(x; \lambda; 1)$$

Graf hustoty pravděpodobnosti:



Graf distribuční funkce:



39 - Normální rozdělení

Řešené příklady: 128
Příklady: 206, 2075.3 Normální rozdělení $N(\mu, \sigma^2)$

- jedno z nejdůležitějších rozdělení spojité náhodné veličiny,
- dobře aproximuje mnoho jiných rozdělení spojité i diskrétní náhodné veličiny.

Definice 5.3.33: Náhodná veličina X má **normální rozdělení** $N(\mu, \sigma^2)$, právě když je její hustota pravděpodobnosti dána funkcí

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in (-\infty, +\infty),$$

kde μ je střední hodnota a σ^2 je rozptyl náhodné veličiny X .

Distribuční funkce

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt, \quad x \in (-\infty, +\infty)$$

Vlastnosti

- $E(X) = \mu$
- $D(X) = \sigma^2$
- $A = 0$
- $\bar{e} = 0$

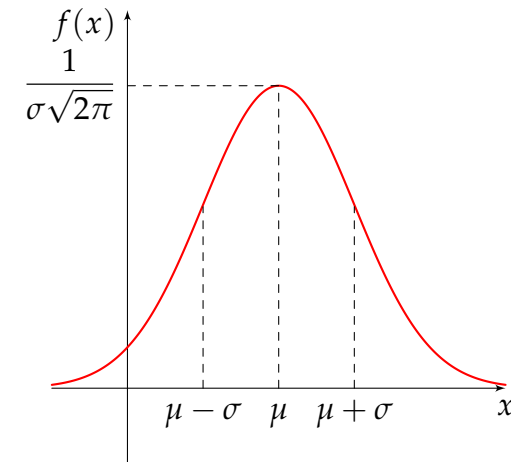
EXCEL:

$$f(x) = \text{NORM.DIST}(x; \mu; \sigma; 0)$$

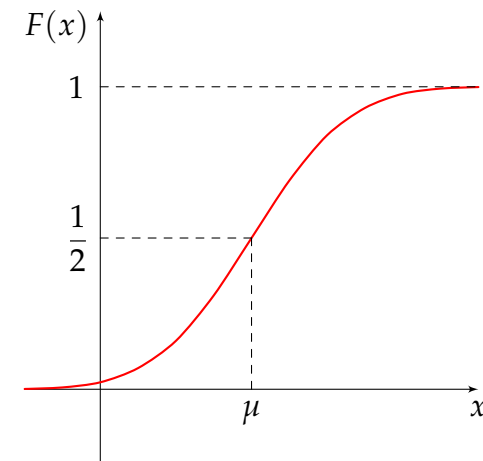
$$F(x) = \text{NORM.DIST}(x; \mu; \sigma; 1)$$

$$x_p = \text{NORM.INV}(p; \mu; \sigma)$$

Graf hustoty pravděpodobnosti (Gaussova křivka):



Graf distribuční funkce:



40 - Normované normální rozdělení

5.4 Normované normální rozdělení $N(0, 1)$

- speciální případ normálního rozdělení s parametry $\mu = 0, \sigma^2 = 1$,
- hodnoty distribuční funkce tohoto rozdělení lze nalézt v tabulkách.

Hustota pravděpodobnosti

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in (-\infty, +\infty)$$

Distribuční funkce

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \quad x \in (-\infty, +\infty)$$

Věta 5.4.34: Má-li spojitá náhodná veličina X normální rozdělení $N(\mu, \sigma^2)$, pak veličina

$$T = \frac{X - \mu}{\sigma}$$

má normované normální rozdělení $N(0, 1)$.

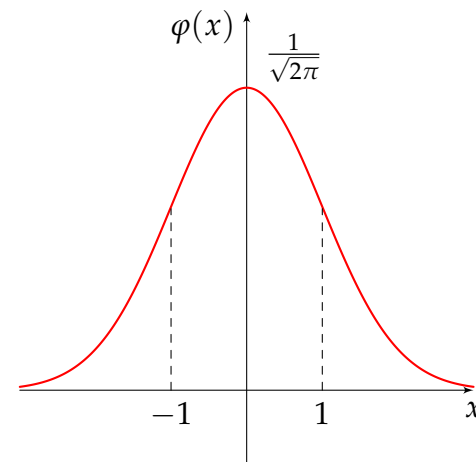
Poznámka: Hodnoty distribuční funkce $\Phi(x)$ jsou tabelovány pouze pro nezáporná x . K určení hodnot pro záporná x využijeme vztah

$$\Phi(x) = 1 - \Phi(-x).$$

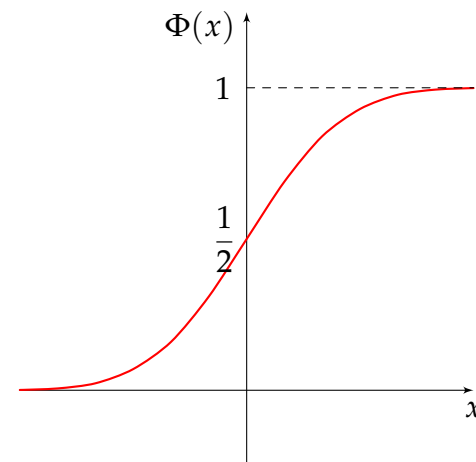
Poznámka: Binomické rozdělení $Bi(n, p)$ lze pro velká n a hodnotu p blíží se k 0,5 (obvykle se uvádí pro $p \in \langle 0, 3; 0, 7 \rangle$) aproximovat normálním rozdělením s parametry $\mu = np$ a $\sigma^2 = np(1 - p)$, tj.

$$Bi(n, p) \sim N(np, np(1 - p)) \quad \text{pro } n \rightarrow \infty \text{ a } p \rightarrow 0,5.$$

Graf hustoty pravděpodobnosti (Gaussova křivka):



Graf distribuční funkce:



Kapitola 6

Náhodný vektor

42 - Náhodný vektor - distribuční a pravděpodobnostní funkce

Řešené příklady: 131
Příklady: 209, 213

6.1 Náhodný vektor

Definice 6.1.35: Uspořádaná n -tice náhodných veličin $X_1, X_2, \dots, X_n$ se nazývá **n -rozměrný náhodný vektor** a značí se $\mathbf{X} = (X_1, X_2, \dots, X_n)$.

Poznámka: Náhodné veličiny $X_1, X_2, \dots, X_n$ tvoří složky náhodného vektoru $\mathbf{X}$.

Poznámka: Dále budeme pracovat s dvourozměrným náhodným vektorem $\mathbf{X} = (X, Y)$. Náhodným vektorem může být například délka a váha vysoustružené součástky.

6.2 Distribuční funkce

Definice 6.2.36: Reálnou funkci $F : \mathbb{R}^2 \rightarrow \langle 0, 1 \rangle$ definovanou předpisem

$$F(x, y) = P(X < x, Y < y)$$

nazýváme **simultánní (sdružená) distribuční funkce** dvourozměrného náhodného vektoru $\mathbf{X} = (X, Y)$.

6.3 Pravděpodobnostní funkce

Definice 6.3.37: **Simultánní (sdružená) pravděpodobnostní funkce** diskrétního náhodného vektoru $\mathbf{X}$ je funkce

$$p(x, y) = P(X = x, Y = y).$$

Poznámka: Náhodný vektor má diskrétní rozdělení, pokud existuje nejvýše spočetně mnoho hodnot náhodného vektoru tak, že

$$\sum_i \sum_j p(x_i, y_j) = 1,$$

sdruženou distribuční funkci poté definujeme:

$$F(x, y) = \sum_{x_i < x} \sum_{y_j < y} p(x_i, y_j).$$

43 - Marginální rozdělení a nezávislost složek

Řešené příklady: 133,134
Příklady: 209,210,212,213

6.4 Marginální rozdělení

V souvislosti s rozdělením náhodného vektoru $\mathbf{X} = (X, Y)$ nazýváme rozdělení jeho složek marginálními rozděleními náhodného vektoru $\mathbf{X}$.

Definice 6.4.38: **Marginální distribuční funkci** definujeme:

$$F_X(x) = P(X < x) = \lim_{y \rightarrow \infty} F(x, y),$$
$$F_Y(y) = P(Y < y) = \lim_{x \rightarrow \infty} F(x, y).$$

Definice 6.4.39: **Marginální pravděpodobnostní funkci** definujeme:

$$p_X(x) = P(X = x) = \sum_{y_j} P(X = x, Y = y_j),$$
$$p_Y(y) = P(Y = y) = \sum_{x_i} P(X = x_i, Y = y).$$

6.5 Nezávislost složek náhodného vektoru

Definice 6.5.40: Necht' $\mathbf{X} = (X, Y)$ je náhodný vektor. Řekneme, že náhodné veličiny X, Y jsou **nezávislé**, pokud $\forall (x, y) \in \mathbb{R}^2$ platí:

$$F(x, y) = F_X(x) \cdot F_Y(y).$$

Vlastnost:

Diskrétní náhodné veličiny X, Y jsou nezávislé právě tehdy, když

$$p(x, y) = p_X(x) \cdot p_Y(y), \quad \forall x, y \in \mathbb{R}.$$

44 - Podmíněné rozdělení

Řešené příklady: 135
Příklady: 210,213

6.6 Podmíněné rozdělení

Rozdělení náhodné veličiny X za předpokladu, že náhodná veličina Y nabyla hodnoty y , popisuje **podmíněné rozdělení** X vzhledem k Y .

Definice 6.6.41: Podmíněné pravděpodobnostní funkce jsou definovány:

$$P(x|y) = \frac{p(x,y)}{p_Y(y)}, \quad p_Y(y) \neq 0,$$

$$P(y|x) = \frac{p(x,y)}{p_X(x)}, \quad p_X(x) \neq 0.$$

Definice 6.6.42: Podmíněné distribuční funkce definujeme:

$$F(x|y) = \frac{\sum_{x < x_i} p(x_i, y)}{p_Y(y)}, \quad p_Y(y) \neq 0,$$

$$F(y|x) = \frac{\sum_{y < y_j} p(x, y_j)}{p_X(x)}, \quad p_X(x) \neq 0.$$

45 - Číselné charakteristiky

Řešené příklady: 136
Příklady: 211

6.7 Smíšené momenty

Kromě momentů náhodných veličin X, Y (složek náhodného vektoru) jsou ještě definovány tzv. smíšené momenty.

Smíšený obecný moment řádu k, n :

$$E(X^k Y^n) = \sum_i \sum_j x_i^k y_j^n p(x_i, y_j).$$

Smíšený centrální moment řádu k, n :

$$E[(X - E(X))^k (Y - E(Y))^n] = \sum_i \sum_j (x_i - E(X))^k (y_j - E(Y))^n p(x_i, y_j).$$

6.8 Podmíněné charakteristiky

Popisují vlastnosti podmíněných rozdělání.

Podmíněná střední hodnota:

$$E(X|y) = \sum_i x_i p(x_i|y).$$

Podmíněný rozptyl:

$$D(X|y) = \sum_i [x_i - E(X|y)]^2 p(x_i|y).$$

46 - Kovariance a koeficient korelace

Řešené příklady: 137
Příklady: 211, 212

6.9 Kovariance

Kovariance $\text{cov}(X, Y)$ je nejjednodušším ukazatelem souvislosti dvou náhodných veličin X, Y . Je definována jako smíšený centrální moment řádu 1, 1,

$$\begin{aligned}\text{cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] = E(X, Y) - E(X)E(Y) \\ &= \sum_i \sum_j x_i y_j p(x_i, y_j) - E(X)E(Y).\end{aligned}$$

Kladná (záporná) hodnota znamená, že se zvětšením hodnoty X se pravděpodobně zvýší (sníží) i hodnota Y . Jsou-li X, Y nezávislé, pak $\text{cov}(X, Y) = 0$.

Kovarianční matice

$$\begin{pmatrix} D(X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & D(Y) \end{pmatrix}$$

6.10 Korelační koeficient

Koeficient korelace $\rho(X, Y)$ určuje míru **lineární** závislosti náhodných veličin X, Y ,

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}}.$$

Vlastnosti:

- Pokud $\rho(X, Y) = 0$, pak jsou veličiny nekorelované (nemusí být nezávislé).
- Pokud $\rho(X, Y) > 0$, říkáme, že X, Y jsou pozitivně korelované (s rostoucím X roste Y).
- Pokud $\rho(X, Y) < 0$, říkáme, že X, Y jsou negativně korelované (s rostoucím X klesá Y).
- Hodnoty $\rho(X, Y)$ blízké ± 1 znamenají silnou lineární závislost.
- Hodnoty $\rho(X, Y)$ blízké 0 znamenají velmi slabou lineární závislost.

Kapitola 7

Popisná analýza

48 - Popisná statistika - základní pojmy

7.1 Popisná statistika

Metody popisné statistiky mají za cíl sumarizovat pozorované hodnoty tak, abychom mohli lépe a jednodušeji pracovat s informací v nich uloženou. Výstupy popisného zpracování dat často slouží jako podklad pro stanovení hypotéz, které jsou následně ověřeny dalšími experimenty.

Statistická jednotka - předmět sledování (např. člověk, zvíře, materiál, událost, ...)

Statistický znak - vlastnost jednotky, kterou jsme schopni číselně nebo slovně popsat (např. pohlaví, věk, počet obyvatel, teplota, doba čekání, ...)

Základní soubor - populace - všechny jednotky, které existují v rámci nějakého logického celku; bývá zadán buď výčtem prvků, nebo vymezením některých společných vlastností (např. všichni obyvatelé ČR, všechny telefonní hovory v síti, ...)

Výběrový soubor - výběr - zkoumaná část populace, vybrané jednotky, nejčastěji volíme náhodný výběr, počet prvků ve výběru označujeme n (např. náhodně oslovení lidé na ulici, hovory monitorované v daném období, ...)

Popisná statistika bývá prvním krokem k odhalení informací skrytých ve velkém množství proměnných a jejich variant.

Statistické proměnné (znaky) dělíme na:

- kvantitativní (číselné) - diskrétní, spojitě,
- kvalitativní (slovní) - nominální, ordinální.

49 - Řazení, třídění

Řešené příklady: 139
Příklady: 215

7.2 Řazení

- kvantitativní proměnná - podle velikosti
- kvalitativní proměnná - podle významu či abecedně

7.3 Třídění

Jedná se o zpřehlednění dat do tabulek, např. uspořádání do **tabulky četností**. Nejčastěji postupujeme tak, že data uspořádáme podle velikosti a stanovíme intervaly, které odpovídají jednotlivým třídám. Mluvíme pak o intervalovém třídění. Jinak se jedná o jednoduché (prosté) třídění - přímo z dat.

7.3.1 Tabulka četností

Postup:

1. zjistíme, v jakém rozmezí se hodnoty proměnné pohybují - nejmenší (min) a největší (max) hodnotu
2. rozhodneme, zda provedeme jednoduché či intervalové třídění - závisí na typu proměnné a počtu obměn
3. určíme počet tříd k (případně **rozpětí tříd**: $(\max - \min) / k$)
4. vypočteme počet pozorování v jednotlivých třídách

ad 3.

jednoduché třídění - počet tříd (řádků) v tabulce se rovná počtu obměn diskrétní proměnné.

intervalové třídění - neexistuje obecný předpis, každý případ je jedinečný. Používá se např. Sturgesovo pravidlo $k \approx 1 + 3,3 \log n$, odmocninové pravidlo $k \approx \sqrt{n}$ či subjektivní pohled.

Třídy musí zahrnovat všechny hodnoty, nejčastěji se volí stejně široké a nesmí se překrývat.

50 - Četnosti, grafické znázornění

Řešené příklady: 139
Příklady: 215, 218, 220

7.3.2 Četnosti

absolutní četnost - f_i ; počet prvků souboru spadající do i -té třídy

relativní četnost - poměr četnosti třídy k celkovému počtu dat

$$\varphi_i = \frac{f_i}{n}$$

kumulativní absolutní četnost - počet prvků souboru v i -té třídě a třídách předcházejících

$$F_i = \sum_{l=1}^i f_l$$

kumulativní relativní četnost - poměr kumulativní četnosti třídy k celkovému počtu dat

$$\phi_i = \frac{F_i}{n}$$

7.4 Grafická znázornění

Pro větší názornost používáme místo tabulek znázornění pomocí grafů. Vizualizace nám dává informaci nejen o charakteru dat, případně procesu, při kterém vznikla, ale i o problematických záznamech a chybných hodnotách. Bez adekvátní vizualizace můžeme v datech nechat pozorování, která negativně ovlivní další hodnocení a znemožní správnou interpretaci výsledků

bodový graf - zobrazuje každou měřenou hodnotu jako bod plochy

histogram - nejpoužívanější grafický nástroj pro vizualizaci poměrových a intervalových dat, (sloupcový) graf, kdy na vodorovnou osu znázorníme třídy a na svislou osu jejich četnosti; jednotlivé hodnoty četností jsou zobrazeny jako výšky sloupců

kruhový graf - (výsečový) je znázornění pomocí výsečí kruhu, kde každé třídě odpovídá jedna výseč; velikosti obsahů výsečí odpovídají četnostem třídy

číslicový graf - lodyha s listy (stem and leaf plot), přehledný zápis číselné hodnoty výběru

krabicový graf - graf ve tvaru obdélníku s tzv. fousky, jednotlivé prvky krabicového grafu nejčastěji odpovídají významným kvantilům vypočteným na základě pozorovaných dat, znázorňuje významné a extrémní hodnoty v souboru

51 - Číselné charakteristiky - část 1.

Řešené příklady: 140
Příklady: 216,218,219,220

Na základě pozorovaných hodnot chceme vypočítat hodnoty, které slouží k "uložení" informace obsažené v datech, neboť použití všech pozorovaných hodnot je nepraktické a často vlastně i nemožné.

7.5 Charakteristiky polohy

Statistický soubor je popsán jediným číslem, které v jistém smyslu vyjadřuje typickou hodnotu popisující celý soubor.

aritmetický průměr - je citlivý na extrémní (odlehle) hodnoty

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$$

Další využívané průměry - vážený aritmetický průměr (výpočet aritmetického průměru hodnot uspořádaných do tabulky četností), useknutý průměr, geometrický průměr (pro analýzu vývoje ukazatele v čase), harmonický průměr (v indexní teorii) a kvadratický průměr.

modus - $\hat{x}$; hodnota, která má největší absolutní četnost (z dat uspořádaných v tabulce je modus možno odhadnout jako střed třídy s nejvyšší absolutní četností); modů může být více, nebo i žádný

kvantily - x_p ; hodnota kvantilu říká, že 100p procent hodnot souboru nabývá hodnoty stejné nebo menší, než je hodnota kvantilu x_p ; nejčastější kvantily jsou medián, kvartily, decily a percentily

medián - $x_{0,5}$; $\tilde{x}$; hodnota, která rozděluje seřazený soubor na dvě části o stejném počtu prvků

- lichý počet hodnot souboru - $\tilde{x}$ je prostřední prvek seřazeného souboru
- sudý počet hodnot souboru - $\tilde{x}$ je průměr dvou prostředních prvků seřazeného souboru
- data uspořádaná do tabulky - $\tilde{x}$ je střed první třídy s kumulativní relativní četností vyšší než 50 %

dolní kvartil - $x_{0,25}$; čtvrtina hodnot je menší nebo rovna této hodnotě

horní kvartil - $x_{0,75}$; tři čtvrtiny hodnot jsou menší nebo rovny této hodnotě

Výběrový průměr

EXCEL: =PRŮMĚR(číslo1;číslo2;...)

Modus

EXCEL: =MODE(číslo1;číslo2;...)

Kvantily

EXCEL: =PERCENTIL(oblast;p)

EXCEL: =MEDIAN(číslo1;číslo2;...)

EXCEL: =QUARTIL.EXC(matice;1)

EXCEL: =QUARTIL.EXC(matice;3)

Využití analytického nástroje Popisná statistika v Excelu.

52 - Číselné charakteristiky - část 2.

Řešené příklady: 141
Příklady: 216-220

7.6 Charakteristiky variability

Umožňují popis variability (rozptýlenosti) výběrového souboru. Vyjadřují proměnlivost hodnot, zda jsou si hodně podobné, či zda se od sebe liší. Některé míry umožňují srovnávat více souborů.

variační rozpětí - stejně jako průměr je citlivé na extrémní (odlehle) hodnoty

$$R = x_{\max} - x_{\min}$$

interkvartilové rozpětí - rozdíl mezi horním a dolním kvantilem, není citlivý na extrémní hodnoty

$$IQR = x_{0,75} - x_{0,25}$$

výběrový rozptyl - vystihuje rozptýlení jednotlivých hodnot kolem aritmetického průměru

$$s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

výběrová směrodatná odchylka - nejužívanější míra variability; kladná odmocnina výběrového rozptylu; je uvedena ve stejných jednotkách jako aritmetický průměr

$$s = \sqrt{s^2}$$

Malá hodnota značí, že prvky souboru jsou si většinou navzájem podobné, a naopak velká směrodatná odchylka signalizuje velké vzájemné odlišnosti. Nevýhodou rozptylu i směrodatné odchylky je, že neumožňují porovnávat variabilitu proměnných vyjádřených v různých jednotkách.

variační koeficient - udává relativní variabilitu vztahenou k průměru, uvádí se v procentech, pomáhá odhalit odlehle hodnoty

$$V_x = \frac{s}{\bar{x}}$$

Použití zejména při srovnání variability dvou různorodých proměnných, které jsou vyjádřeny v různých měrných jednotkách. Je-li $V_x > 50 \%$, může to ukazovat na nesourodost souboru a není např. vhodné používat aritmetický průměr jako charakteristiku polohy.

Výběrový rozptyl

EXCEL: =VAR.S(číslo1;číslo2;...)

Výběrová směrodatná odchylka

EXCEL: =SMODCH.VÝBĚR.S(číslo1;číslo2;...)

Využití analytického nástroje Popisná statistika v Excelu.

Odlehlá pozorování

- hodnoty ležící mimo interval

$$\langle x_{0,25} - 1,5 \cdot IQR; x_{0,75} + 1,5 \cdot IQR \rangle$$

nazýváme odlehlými

53 - Číselné charakteristiky - část 3.

Řešené příklady: 142
Příklady: 217,218,219,220

7.7 Charakteristiky tvaru rozdělení

Popisují tvar rozdělení, rozložení hodnot v souboru, jaké hodnoty převažují, atd.

šikmost - vyjadřuje, jsou-li hodnoty kolem průměru rozloženy symetricky, nebo zdali převažují spíše hodnoty podprůměrné či nadprůměrné

$$A = \frac{n}{(n-1)(n-2)s^3} \sum_{j=1}^n (x_j - \bar{x})^3$$

- $A > 0$ - kladné zešikmení (převládají nízké hodnoty)
- $A = 0$ - symetrické (hodnoty rozloženy rovnoměrně)
- $A < 0$ - záporné zešikmení (převládají vysoké hodnoty)

špičatost - vyjadřuje, jaký průběh má rozdělení hodnot kolem zvoleného středu, čím více je rozdělení špičatější, tím více jsou hodnoty soustředěny kolem daného středu

$$\tilde{e} = \frac{n(n+1)}{(n-1)(n-2)(n-3)s^4} \sum_{j=1}^n (x_j - \bar{x})^4 - 3 \frac{(n-1)^2}{(n-2)(n-3)}$$

- $\tilde{e} > 0$ - špičaté rozdělení (hodnoty koncentrovány kolem středu)
- $\tilde{e} = 0$ - normální rozdělení (hodnoty rozloženy normálně)
- $\tilde{e} < 0$ - ploché rozdělení (hodnoty nejsou koncentrovány kolem středu)

Zaokrouhlování:

směrodatná odchylka - na 2 platné cifry směrem nahoru

průměr, rozptyl, kvantily - stejný řád jako směrodatná odchylka

min, max - stejný řád, ve kterém jsou data

var. koeficient, šikmost, špičatost - max. na 3 des. místa

Šikmost

EXCEL: =SKEW(číslo1;číslo2;...)

Špičatost - exces

EXCEL: =KURT(číslo1;číslo2;...)

Využití analytického nástroje Popisná statistika v Excelu.

Kapitola 8

Odhady parametrů

55 - Induktivní statistika, bodové odhady

Dva typy odhadů:

- bodové odhady - odhadem parametru základního souboru je jedna hodnota,
- intervalové odhady - odhadem parametru základního souboru je interval hodnot.

8.1 Bodové odhady

Definice 8.1.43: **Bodový odhad** (estimátor) parametru β je statistika B , která aproximuje parametr β s předepsanou přesností.

Definice 8.1.44: **Nevychýlený odhad** parametru β je taková statistika β_n , jejíž očekávaná hodnota

$$E(\beta_n) = \beta,$$

tzn. je to každá statistika, která statisticky (stochasticky) konverguje k parametru β . V opačném případě se veličina β_n nazývá **odhadem vychýleným**, a to vpravo nebo vlevo, podle toho, zda $E(\beta_n) - \beta > 0$, resp. $E(\beta_n) - \beta < 0$.

Definice 8.1.45: **Konzistentní (nesporný) odhad** parametru β je taková statistika β_n , že pro dosti velká n je

$$P(\beta_n - \beta \leq \varepsilon) > 1 - \eta,$$

kde $\varepsilon > 0$, $\eta > 0$ jsou jakákoliv (libovolně malá) předem zvolená čísla.

K získávání bodových odhadů se používá **metoda momentů** a **metoda maximální věrohodnosti**.

Bodový odhad střední hodnoty: $\hat{\mu} = \bar{x}$.

Bodový odhad rozptylu: $\hat{\sigma}^2 = s^2$.

56 - Intervalové odhady

Řešené příklady: 144,145,146,147
Příklady: 222

8.2 Intervalové odhady

Definice 8.2.46: **Intervalový odhad** parametru β základního souboru je interval $\langle B_1, B_2 \rangle$, v němž leží skutečná hodnota parametru s pravděpodobností $1 - \alpha$, tj.

$$P(B_1 \leq \beta \leq B_2) = 1 - \alpha.$$

Interval $\langle B_1, B_2 \rangle$ se nazývá **interval spolehlivosti (konfidenční interval)** pro parametr β na **hladině významnosti** α (nebo se **stupněm spolehlivosti** $1 - \alpha$). Hodnoty B_1, B_2 jsou kritické hodnoty pro parametr β . Intervaly $(-\infty, B_1)$ a (B_2, ∞) se nazývají **kritické intervaly**.

Definice 8.2.47: **Kritické hodnoty** rozdělení na hladině významnosti α jsou kvantily, kde index α vyjadřuje pravděpodobnost, že náhodná veličina (u symetrických rozdělení její absolutní hodnota), překročí tuto hodnotu.

Definice 8.2.48: **Hladina významnosti** α je pravděpodobnost, že skutečná hodnota parametru neleží uvnitř intervalu spolehlivosti.

Definice 8.2.49: **Stupeň spolehlivosti** $1 - \alpha$ je pravděpodobnost, že skutečná hodnota parametru leží uvnitř intervalu spolehlivosti.

Intervaly spolehlivosti (IS) konstruujeme jako jednostranné (důležitá je pouze jedna mez)

$$\text{levostranný IS : } P(\beta \in \langle B_1, \infty \rangle) = 1 - \alpha$$

$$\text{pravostranný IS : } P(\beta \in (-\infty, B_2]) = 1 - \alpha$$

nebo oboustranné

$$\beta \in \langle B_1, B_2 \rangle.$$

57 - Intervalový odhad střední hodnoty

Řešené příklady: 146
Příklady: 222 

8.2.1 Intervalový odhad střední hodnoty

Intervalový odhad střední hodnoty - známe σ^2

Mějme sledovanou veličinu $X \rightarrow N(\mu; \sigma^2)$, pak při známém rozptylu má náhodná veličina

$$T(X) = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$$

normované normální rozdělení pravděpodobnosti $N(0, 1)$.

Hledáme interval spolehlivosti na hladině významnosti α :

$$P(z_{\frac{\alpha}{2}} \leq T(X) \leq z_{1-\frac{\alpha}{2}}) = 1 - \alpha,$$

kde $z_{\frac{\alpha}{2}}$ je $\frac{\alpha}{2}$ -tý kvantil normovaného normálního rozdělení,

$$P\left(\frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \bar{x} - \mu \leq \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

Využijeme vlastnosti symetrie normovaného normálního rozdělení kolem nuly (platí $z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$) a dostáváme tedy:

$$P\left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

Intervalový odhad střední hodnoty - neznáme σ^2

Mějme sledovanou veličinu $X \rightarrow N(\mu; \sigma^2)$, pak při neznámém rozptylu má náhodná veličina

$$T(X) = \frac{\bar{x} - \mu}{s} \sqrt{n}$$

t-rozdělení pravděpodobnosti s $(n - 1)$ stupni volnosti. t-rozdělení je také symetrické kolem nuly, tzn. postup odvození intervalu spolehlivosti je obdobný jako při odvození se známým rozptylem.

EXCEL:

kvantil normovaného normálního rozdělení $z_{1-\frac{\alpha}{2}}$

$$= \text{NORM.S.INV}\left(1 - \frac{\alpha}{2}\right)$$

58 - Intervalový odhad střední hodnoty a rozptylu

Řešené příklady: 144,145,147
Příklady: 222,223,224

Interval spolehlivosti na hladině významnosti α lze vyjádřit:

$$P(t_{\frac{\alpha}{2}}(n-1) \leq T(X) \leq t_{1-\frac{\alpha}{2}}(n-1)) = 1 - \alpha,$$

kde $t_{\frac{\alpha}{2}}(n-1)$ je $\frac{\alpha}{2}$ -tý kvantil t-rozdělení s $n-1$ stupni volnosti. Po úpravách dostáváme:

$$P\left(\bar{x} - \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1)\right) = 1 - \alpha.$$

8.2.2 Intervalový odhad rozptylu

Mějme sledovanou veličinu $X \rightarrow N(\mu; \sigma^2)$, pak při neznámé střední hodnotě (a to zpravidla platí) má náhodná veličina

$$\chi^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2}$$

chí-kvadrát rozdělení s $(n-1)$ stupni volnosti.

Oboustranný intervalový odhad NV χ^2 můžeme zapsat pravděpodobnostní rovnicí:

$$P\left(\chi_{\frac{\alpha}{2}}^2(n-1) \leq \chi^2 \leq \chi_{1-\frac{\alpha}{2}}^2(n-1)\right) = 1 - \alpha,$$

kde $\chi_{\frac{\alpha}{2}}^2(n-1)$ je $\frac{\alpha}{2}$ -tý kvantil chí-kvadrát rozdělení s $n-1$ stupni volnosti. Dostáváme:

$$P\left(\frac{(n-1) \cdot s^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \leq \sigma^2 \leq \frac{(n-1) \cdot s^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}\right) = 1 - \alpha.$$

Z intervalového odhadu rozptylu lze odvodit intervalový odhad pro směrodatnou odchylku:

$$P\left(\sqrt{\frac{(n-1) \cdot s^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}} \leq \sigma \leq \sqrt{\frac{(n-1) \cdot s^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}}\right) = 1 - \alpha.$$

EXCEL:

kvantil t-rozdělení s $n-1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n-1\right)$$

kvantil chí-kvadrát rozdělení s $n-1$ stupni volnosti $\chi_{1-\frac{\alpha}{2}}^2(n-1)$

$$=CHISQ.INV\left(1 - \frac{\alpha}{2}; n-1\right)$$

Kapitola 9

Testování hypotéz

60 - Testování hypotéz

9.1 Statistické hypotézy - úvod

Definice 9.1.50: **Statistická hypotéza** je určitý předpoklad o rozdělení (jednoho či více) základního souboru.

Hypotézy se mohou týkat pouze neznámých číselných parametrů rozložení NV, pak jde o **testy parametrické**, ostatní typy jsou **testy neparametrické**.

Nulová hypotéza H_0 - hypotéza, jejíž platnost ověřujeme.

Alternativní hypotéza H_1 - opačná k H_0 . Může být buď oboustranná nebo jednostranná.

Statistické testy jsou postupy, jimiž ověřujeme platnost nulové hypotézy. Na základě nich pak hypotézu buď přijmeme, nebo zamítneme.

Testovací kritérium (statistika) t_h je NV závislá na náhodném výběru mající vztah k nulové hypotéze.

Stejně jako u konstrukce intervalů spolehlivosti, nelze zjistit, zda hypotéza platí na 100 %. Nejčastěji se volí testy s jistotou 99 %, 95 % nebo 90 %, obecně $1 - \alpha$, kde α je **hladina významnosti**, tedy pravděpodobnost chyby testu, kdy zamítneme H_0 , přestože platí (chyba 1. druhu).

Kritické hodnoty testovacího kritéria a_α, b_α oddělují interval prakticky možných hodnot $\langle a_\alpha, b_\alpha \rangle$ od kritických intervalů, v nichž se hodnoty NV X vyskytují s pravděpodobností α .

Porovnání hodnoty testovacího kritéria s jeho kritickými hodnotami slouží k rozhodnutí o výsledku testu. Hypotézu přijímáme, leží-li pozorovaná hodnota testovacího kritéria v intervalu prakticky možných hodnot, tzn. rozdíl mezi pozorovanou a teoretickou hodnotou testovacího kritéria je vysvětlitelný na dané hladině významnosti náhodností výběru.

9.1.1 Kroky při testování hypotézy - klasický postup

- Formulace výzkumné otázky ve formě nulové a alternativní hypotézy
- Zvolení přijatelné úrovně chyby rozhodování (volba hladiny významnosti α)
- Volba (konkretizace) testu a výpočet testovacího kritéria
- Sestrojení kritického intervalu
- Rozhodnutí a závěr testu (nezamítnutí, nebo zamítnutí H_0)

Druhý způsob vyhodnocení testu je na základě p-hodnoty.

p-hodnota představuje nejmenší hladinu významnosti testu, při níž na daných datech ještě zamítneme nulovou hypotézu.

Platí tedy, že čím nižší p-hodnota testu je, tím menší nám tento test indikuje pravděpodobnost, že platí nulová hypotéza. Vyjde-li nám při vyhodnocení statistického testu p-hodnota "blízká nule", znamená to, že naše nulová hypotéza má velmi malou oporu v pozorovaných datech a můžeme ji zamítnout.

Výpočet p-hodnoty pro pozorovanou hodnotu testovacího kritéria t_h :

$$H_1 : \dots \neq \dots \Rightarrow p = 2 \cdot \min(F(t_h), 1 - F(t_h))$$

$$H_1 : \dots < \dots \Rightarrow p = F(t_h)$$

$$H_1 : \dots > \dots \Rightarrow p = 1 - F(t_h)$$

61 - Parametrické testy - test o střední hodnotě

Řešené příklady: 149
Příklady: 226

9.2 Parametrické testy

9.2.1 Test hypotézy o střední hodnotě μ (jednovýběrový t-test)

Předpoklady:

Je dán výběr ze ZS s rozdělením $N(\mu, \sigma^2)$ o rozsahu n s výběrovou střední hodnotou $\bar{x}$ a výběrovým rozptylem s^2 .

Nulová hypotéza:

$$H_0 : \mu = \mu_0$$

Alternativní hypotéza:

$H_1 : \mu \neq \mu_0$... oboustranná hypotéza

$H_1 : \mu > \mu_0$... pravostranná hypotéza

$H_1 : \mu < \mu_0$... levostranná hypotéza

Testovací kritérium:

$$T = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}$$

Testovací statistika má Studentovo rozdělení pravděpodobnosti s $n - 1$ stupni volnosti.

Kritická hodnota:

$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n-1)$... oboustranný test

$t_{\text{krit}} = t_{1-\alpha}(n-1)$... jednostranný test

Závěr:

$$|T| < t_{\text{krit}} \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil t -rozdělení s $n - 1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

62 - Parametrické testy - test o rozptylu

Řešené příklady: 150
Příklady: 2279.2.2 Test hypotézy o rozptylu σ^2

Nulová hypotéza:

$$H_0 : \sigma^2 = \sigma_0^2$$

Alternativní hypotéza:

$$H_1 : \sigma^2 \neq \sigma_0^2 \dots \text{oboustranná hypotéza}$$

$$H_1 : \sigma^2 > \sigma_0^2 \dots \text{pravostranná hypotéza}$$

$$H_1 : \sigma^2 < \sigma_0^2 \dots \text{levostranná hypotéza}$$

Testovací kritérium:

$$\chi^2 = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

Testovací statistika má chí-kvadrát rozdělení s $n - 1$ stupni volnosti.

Kritická hodnota:

$$\chi_{\frac{\alpha}{2}}^2(n-1); \chi_{1-\frac{\alpha}{2}}^2(n-1) \dots \text{oboustranný test}$$

$$\chi_{\alpha}^2(n-1); \chi_{1-\alpha}^2(n-1) \dots \text{jednostranný test}$$

Závěr:

oboustranný test:

$$\chi^2 \in \left\langle \chi_{\frac{\alpha}{2}}^2(n-1); \chi_{1-\frac{\alpha}{2}}^2(n-1) \right\rangle \Rightarrow H_0 \text{ nezamítáme}$$

jednostranný test:

$$\chi^2 > \chi_{\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 < \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

$$\chi^2 < \chi_{1-\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 > \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil chí-kvadrát rozdělení s $n - 1$ stupni volnosti $\chi_{1-\frac{\alpha}{2}}^2(n-1)$

$$=\text{CHISQ.INV}\left(1 - \frac{\alpha}{2}; n - 1\right)$$

63 - Parametrické testy - test shody dvou rozptylů

Řešené příklady: 151
Příklady: 228**9.2.3 Dvouvýběrový F-test**

Mějme dva výběrové soubory s rozsahem n_1, n_2 a výběrovým rozptylem s_1^2, s_2^2 . Ověřujeme předpoklad, že variabilita souborů je srovnatelná.

Nulová hypotéza:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

Alternativní hypotéza:

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \dots \text{oboustranná hypotéza}$$

$$H_1 : \sigma_1^2 > \sigma_2^2 \dots \text{pravostranná hypotéza}$$

Testovací kritérium:

$$F = \frac{s_1^2}{s_2^2}$$

Indexy 1,2 volíme, aby $F > 1$. Testovací statistika má F rozdělení s $n_1 - 1, n_2 - 1$ stupni volnosti.

Kritická hodnota:

$$F_{\text{krit}} = F_{1-\frac{\alpha}{2}}(n_1 - 1; n_2 - 1) \dots \text{oboustranný test}$$

$$F_{\text{krit}} = F_{1-\alpha}(n_1 - 1; n_2 - 1) \dots \text{jednostranný test}$$

Závěr:

$$F < F_{\text{krit}} \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil F-rozdělení s $(n_1 - 1; n_2 - 1)$ stupni volnosti

$$=F.INV\left(1 - \frac{\alpha}{2}; n_1 - 1; n_2 - 1\right)$$

64 - Parametrické testy - test shody dvou středních hodnot

Řešené příklady: 152
Příklady: 228

9.2.4 Dvouvýběrový t-test

Předpoklady:

Jsou dány dva výběry o rozsazích n_1, n_2 s výběrovými středními hodnotami $\bar{x}_1, \bar{x}_2$ a výběrovými rozptyly s_1^2, s_2^2 , vybrané ze dvou ZS s rozděleními $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$.

Nulová hypotéza:

$$H_0 : \mu_1 = \mu_2$$

Alternativní hypotéza:

$H_1 : \mu_1 \neq \mu_2$... oboustranná hypotéza

$H_1 : \mu_1 > \mu_2$... pravostranná hypotéza

$H_1 : \mu_1 < \mu_2$... levostranná hypotéza

Pro výpočet testovací a kritické hodnoty musíme první ověřit, zda lze předpokládat platnost hypotézy $\sigma_1^2 = \sigma_2^2$. Prověříme F -testem.

a) Dvouvýběrový t-test s rovností rozptylů (platí $\sigma_1^2 = \sigma_2^2$)

Testovací kritérium:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

Testovací statistika má t rozdělení s $n_1 + n_2 - 2$ stupni volnosti.

Kritická hodnota:

$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$... oboustranný test

$t_{\text{krit}} = t_{1-\alpha}(n_1 + n_2 - 2)$... jednostranný test

Závěr:

$$|T| < t_{\text{krit}} \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil t -rozdělení s $n_1 + n_2 - 2$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n_1 + n_2 - 2\right)$$

65 - Parametrické testy - test shody dvou středních hodnot

Řešené příklady: 152
Příklady: 228**b) Dvouvýběrový t-test s nerovností rozptylů** (za podmínky $\sigma_1^2 \neq \sigma_2^2$)**Testovací kritérium:**

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_2 - 1)s_1^2 + (n_1 - 1)s_2^2}} \sqrt{(n_1 - 1)(n_2 - 1)},$$

které je složeno ze dvou Studentových rozdělení pravděpodobnosti.

Kritická hodnota:

$$t_{\text{krit}} = \frac{(n_2 - 1)s_1^2 t_{1-\frac{\alpha}{2}}(n_1 - 1) + (n_1 - 1)s_2^2 t_{1-\frac{\alpha}{2}}(n_2 - 1)}{(n_2 - 1)s_1^2 + (n_1 - 1)s_2^2} \dots \text{oboustranný test}$$

$$t_{\text{krit}} = \frac{(n_2 - 1)s_1^2 t_{1-\alpha}(n_1 - 1) + (n_1 - 1)s_2^2 t_{1-\alpha}(n_2 - 1)}{(n_2 - 1)s_1^2 + (n_1 - 1)s_2^2} \dots \text{jednostranný test}$$

Závěr:

$$|T| < t_{\text{krit}} \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:kvantil t -rozdělení s $n - 1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n - 1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

66 - Parametrické testy - párový test

Řešené příklady: 153
Příklady: 229**9.2.5 Studentův test pro párované hodnoty****Předpoklady:**

Ze dvou ZS s rozděleními $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$ byly vybrány dva výběry se stejnými rozsahy n . Každému prvku prvního výběru x_{1i} odpovídá právě jeden prvek druhého výběru x_{2i} ($d = x_{2i} - x_{1i}$).

Nulová hypotéza:

$$H_0 : \mu_1 = \mu_2 \ (\bar{d} = 0)$$

Alternativní hypotéza:

$$H_1 : \mu_1 \neq \mu_2 \ (\bar{d} \neq 0) \dots \text{oboustranná hypotéza}$$

$$H_1 : \mu_1 > \mu_2 \ (\bar{d} < 0) \dots \text{pravostranná hypotéza}$$

$$H_1 : \mu_1 < \mu_2 \ (\bar{d} > 0) \dots \text{levostranná hypotéza}$$

Testovací kritérium:

$$t = \frac{|\bar{x}_d|}{s_d} \cdot \sqrt{n},$$

kde $\bar{x}_d$, s_d jsou výběrové charakteristiky souboru tvořených z rozdílů párových hodnot, testovací statistika má Studentovo rozdělení pravděpodobnosti $t(n-1)$.

Kritická hodnota:

$$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n-1) \dots \text{oboustranný test}$$

$$t_{\text{krit}} = t_{1-\alpha}(n-1) \dots \text{jednostranný test}$$

Závěr:

$$|T| < t_{\text{krit}} \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil t -rozdělení s $n-1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

67 - Neparametrické testy - testy dobré shody

Řešené příklady: 154,155
Příklady: 230

9.3 Neparametrické testy

9.3.1 Pearsonův test dobré shody (χ^2 -test) pro jeden výběr

Předpoklady:

Výsledky pozorování jsou rozděleny do k skupin a v každé skupině je zjištěna skupinová četnost f_{ei} (experimentální četnosti). Uvažujme určité rozdělení, které budeme považovat za model pro náš výběr. Pro každou třídu určíme teoretické (očekávané) četnosti f_{oi} pro $i = 1, \dots, k$.

Nulová hypotéza:

ZS má očekávané rozdělení pravděpodobnosti, tj. četnosti f_{ei} a f_{oi} pro $i = 1, \dots, n$ se liší pouze náhodně.

Testovací kritérium:

$$\chi^2 = \sum_{i=1}^k \frac{(f_{ei} - f_{oi})^2}{f_{oi}}$$

Testovací kritérium podléhá Pearsonovu rozdělení pravděpodobnosti $\chi^2(k - s - 1)$, s značí počet parametrů očekávaného rozdělení pravděpodobnosti.

Závěr:

$$\chi^2 < \chi_{1-\alpha}^2(k - s - 1) \Rightarrow H_0 \text{ nezamítáme}$$

Podmínky použití testu:

- všechny f_{oi} mají být > 1
- nejvýše 20 % f_{oi} může být < 5
- počet tříd k nevolíme větší než 20

Poznámka: Pokud podmínky nejsou splněny, musíme sloučit sousední třídy.

EXCEL:

kvantil chí-kvadrát rozdělení s $k - s - 1$ stupni volnosti $\chi_{1-\alpha}^2(k - s - 1)$
=CHISQ.INV($1 - \alpha; k - s - 1$)

68 - Neparametrické testy - testy dobré shody

Řešené příklady: 154,155
Příklady: 230

9.3.2 Kolmogorovův-Smirnovův test dobré shody pro jeden výběr

Předpoklady:

Výsledky pozorování jsou rozděleny do k skupin a v každé skupině je zjištěna skupinová četnost f_{ei} (experimentální četnosti). Uvažujme určité rozdělení, které budeme považovat za model pro náš výběr. Pro každou třídu určíme teoretické (očekávané) četnosti f_{oi} a kumulativní četnosti F_{ei} , F_{oi} .

Nulová hypotéza:

ZS má očekávané rozdělení pravděpodobnosti, tj. četnosti f_{ei} , f_{oi} pro $i = 1, \dots, k$ se liší pouze náhodně.

Testovací kritérium:

$$D_1 = \frac{1}{n} \max |F_{ei} - F_{oi}|$$

Testovací kritérium má speciální rozdělení pravděpodobnosti, jehož hodnoty jsou pro $n < 40$ tabelovány a pro $n \geq 40$ určovány dle vzorců.

Kritická hodnota:

$$\alpha = 0,05 \Rightarrow D_{1,0,05}(n) = \frac{1,36}{\sqrt{n}}$$

$$\alpha = 0,01 \Rightarrow D_{1,0,01}(n) = \frac{1,63}{\sqrt{n}}$$

Závěr:

$$D_1 < D_{1,\alpha}(n) \Rightarrow H_0 \text{ nezamítáme}$$

69 - Neparametrické testy - testy extrémních hodnot

Řešené příklady: 156
Příklady: 231

9.4 Testy extrémních hodnot

9.4.1 Dixonův test extrémních odchylek

Předpoklady:

Ve výběrovém souboru o rozsahu n je $x_1 = \min(x_i)$, resp. $x_n = \max(x_i)$.

Nulová hypotéza:

H_0 : Hodnota x_1 , resp. x_n (největší hodnota) se neliší významně od ostatních hodnot souboru.

Testovací kritérium:

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1}$$

$$Q_n = \frac{x_n - x_{n-1}}{x_n - x_1}$$

podle toho, testujeme-li minimální nebo maximální hodnotu ve výběru.

Kritické hodnoty $Q_{1;\alpha}$, resp. $Q_{n;\alpha}$ jsou tabelovány.

Závěr:

$$Q_1 < Q_{1;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

$$Q_n < Q_{n;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

70 - Neparametrické testy - testy extrémních hodnot

Řešené příklady: 156
Příklady: 231

9.4.2 Grubbsův test extrémních odchylek

Předpoklady:

Ve výběrovém souboru o rozsahu n je $x_1 = \min(x_i)$, resp. $x_n = \max(x_i)$ (např. hodnoty jsou seřazeny podle velikosti od x_1 do x_n). $\bar{x}$ je střední hodnota výběru, s je výběrová směrodatná odchylka.

Nulová hypotéza:

H_0 : Hodnota x_1 (nejmenší hodnota), resp. x_n (největší hodnota) se neliší významně od ostatních hodnot souboru.

Testovací kritérium:

$$T_1 = \frac{\bar{x} - x_1}{s}$$

$$T_n = \frac{x_n - \bar{x}}{s}$$

podle toho, testujeme-li minimální nebo maximální hodnotu ve výběru.

Kritické hodnoty $T_{1;\alpha}$, resp. $T_{n;\alpha}$ jsou tabelovány.

Závěr:

$$T_1 < T_{1;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

$$T_n < T_{n;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

Poznámka: Vede-li test k závěru, že extrémní hodnotu je třeba ze souboru vyloučit, je třeba sestavit znovu všechny výběrové charakteristiky (ze souboru bez extrémní hodnoty) pro případné další výpočty.

Kapitola 10

Lineární regrese

72 - Lineární regrese

10.1 Závislost dvou číselných proměnných

Často je potřeba porovnat vztahy mezi proměnnými, změny proměnné, atd. Sledujeme tedy 2 veličiny (jedna je závislá a jedna nezávislá) a zajímá nás vztah mezi veličinami.

Grafická analýza - veličiny vyneseme do bodového grafu (osa y ...závislá proměnná, osa x ...nezávislá proměnná), graf napomáhá odhalení závislosti i naznačuje sílu případné závislosti

závislá proměnná (vysvětlovaná) - její chování se snažíme vysvětlit

nezávislá proměnná (vysvětlující) - její chování vysvětluje chování závislé proměnné

Závislost může být

- pevná (funkční) - lineární (hodnoty leží na přímce), nelineární (hodnoty leží na křivce jiné než přímka)
- volná (stochastická) - hodnoty neleží přímo na křivce, ale je patrný jejich průběh kolem pomyslné křivky, čím jsou body blíže pomyslné křivce, tím je závislost těsnější
- nezávislost - pomyslná křivka je rovnoběžná s osou x nebo pomyslnou křivku procházející body nelze vůbec nalézt

EXCEL: Vložení -> Bodový graf

73 - Regresní analýza

Řešené příklady: 158
Příklady: 233

10.2 Regresní analýza

Regresní analýza se věnuje vystižení charakteru závislosti a určení její konkrétní formy.

- Jednoduchá regresní analýza - popisuje závislost dvou číselných proměnných (jedna je vysvětlující a jedna je vysvětlovaná)
- Vícenásobná regresní analýza - popisuje závislost více číselných proměnných (více je vysvětlujících a jen jedna je vysvětlovaná)

Regresní model - zjednodušené zobrazení reality, závislost popisuje pomocí rovnice (v grafu křivka)

- regresní přímka: $y = ax + b$
- regresní parabola: $y = ax^2 + bx + c$
- regresní hyperbola: $y = \frac{a}{x} + b$
- logaritmická funkce: $y = a \ln x + b$
- exponenciální funkce: $y = ae^{bx}$
- polynom stupně k : $y = a_k x^k + a_{k-1} x^{k-1} + \dots + a_1 x + a_0$

$a, b, c, a_k, \dots, a_0$ nazýváme **regresní koeficienty**

74 - Lineární regresní model

Řešené příklady: 158
Příklady: 233 

10.3 Metoda nejmenších čtverců

Sledujeme dvě veličiny x , y . Předpokládáme, že mezi nimi existuje vztah $y = f(x)$. Chceme najít funkci, která vystihuje tento vztah; v případě lineárního regresního modelu se jedná o regresní přímku

$$y = ax + b.$$

Pro nalezení rovnice regresní přímky (prochází nejbližše všem bodům a je jediná) slouží **metoda nejmenších čtverců**, která vybere ze všech možných přímek takovou, pro níž je součet druhých mocnin odchylek bodů od přímky minimální.

rezidua - chyby nalezeného řešení

$$e_i = y_i - \tilde{y}_i$$

y_i ... známé přibližné hodnoty (získané např. měřením)

$\tilde{y}_i$... odhad regresní funkce

Kriteriální funkce:

$$\varphi = \sum_{i=1}^n e_i^2$$

Po dosazení dostaneme:

$$\varphi = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 = \sum_{i=1}^n (y_i - (ax_i + b))^2 = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

Hledáme tedy minimum kriteriální funkce. Koeficienty a , b určíme ze soustavy normálních rovnic

$$\frac{\partial \varphi}{\partial a} = 0,$$

$$\frac{\partial \varphi}{\partial b} = 0.$$

Maticový zápis modelu:

$$Y = X \cdot A,$$

kde

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_3 \end{pmatrix}, X = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_3 & 1 \end{pmatrix}, A = \begin{pmatrix} a \\ b \end{pmatrix}$$

75 - Lineární regresní model - MNČ

Řešené příklady: 158
Příklady: 233,235,236

Po určení parciálních derivací dostáváme soustavu

$$2 \sum_{i=1}^n (y_i - ax_i - b)(-1) = 0,$$

$$2 \sum_{i=1}^n (y_i - ax_i - b)(-x_i) = 0.$$

Upravíme:

$$\sum_{i=1}^n (y_i - ax_i - b) = 0,$$

$$\sum_{i=1}^n (x_i y_i - ax_i^2 - bx_i) = 0.$$

Soustava normálních rovnic:

$$a \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i,$$

$$a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i.$$

Bodové odhady parametrů a, b dostaneme řešením soustavy pomocí Cramerova pravidla:

$$\begin{aligned} D &= \left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2 \\ D_a &= \sum_{i=1}^n y_i \sum_{i=1}^n x_i - n \sum_{i=1}^n x_i y_i \\ D_b &= \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} a &= \frac{D_a}{D} \\ b &= \frac{D_b}{D} \end{aligned}$$

Koeficienty regresní přímky

a ... je směrnici přímky, udává její sklon

b ... je průsečík přímky s osou y

Směrodatné chyby odhadů regresních parametrů

$$s(a_i) = \sqrt{s_e^2 \cdot h_{ii}},$$

kde $i = 1, \dots, p$,

p značí počet koeficientů v modelu,

$h_{11}, \dots, h_{pp}$ jsou prvky na hlavní diagonále matice $(X^T X)^{-1}$ pro daný regresní model.

76 - Verifikace a kvalita modelu

Řešené příklady: 159
Příklady: 235

10.4 Verifikace modelu

Po nalezení konkrétního odhadu regresní funkce na základě výběru si musíme položit otázku, zda je nalezený odhad kvalitní, zda byl zvolen vhodný typ regresní funkce, atd. Je tedy potřeba provést verifikaci modelu a hodnocení kvality modelu.

Koeficient determinace R^2 nabývá hodnot $\langle 0; 1 \rangle$ a vystihuje, jak těsně datové body přiléhají ke křivce. Čím je vyšší hodnota, tím je model vhodnější. Navíc určuje, jaké procento změn vysvětlované proměnné je vysvětleno modelem. V případě nelineární regrese označujeme index determinace I^2 .

Pro naměřené y_i a odhadnuté $\tilde{y}_i$ hodnoty vysvětlované proměnné počítáme:

$$S_T^2 = \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{celkový součet čtverců,}$$

$$S_R^2 = \sum_{i=1}^n (\tilde{y}_i - \bar{y})^2 \quad \text{regresní součet čtverců,}$$

$$SSE = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 \quad \text{reziduální součet čtverců.}$$

Pak

$$R^2 = \frac{S_R^2}{S_T^2} = 1 - \frac{SSE}{S_T^2}.$$

Pokud není významný rozdíl mezi koeficienty determinace u různých modelů, vždy je vhodné zvolit model jednodušší. Hodnota R^2 roste s počtem koeficientů p , proto je nutné modely s více koeficienty porovnávat pomocí upraveného koeficientu determinace

$$R_{ADJ}^2 = 1 - (1 - R^2) \frac{n-1}{n-p}.$$

Reziduální rozptyl - nevychýlený odhad rozptylu σ^2

$$\hat{\sigma}^2 = s_e^2 = \frac{SSE}{n-p}$$

Celkový F-test - testujeme, zda hodnota vysvětlované proměnné závisí na lineární kombinaci vysvětlujících proměnných

H_0 : zvolený model není statisticky významný ($a_1 = \dots = a_p = 0$)

H_1 : $\neg H_0$

Testovací kritérium: $F = \frac{S_R^2}{p-1} : \frac{SSE}{n-p}$

Testovací statistika má F-rozdělení pravděpodobnosti s $p-1$ a $n-p$ stupni volnosti. Kritickou hodnotou je kvantil $F_{1-\alpha}(p-1, n-p)$.

Pokud bychom H_0 nezamítli, znamenalo by to, že model je chybně specifikován.

Dílič t-testy - testujeme oprávněnost setrvání koeficientů a_i v regresním modelu, každý koeficient se testuje zvlášť ($i = 1, \dots, p$)

H_0 : koeficient není statisticky významný ($a_i = 0$)

H_1 : koeficient je statisticky významný ($a_i \neq 0$)

Testovací kritérium: $t = \frac{a_i}{s(a_i)}$

Testovací statistika má t-rozdělení pravděpodobnosti s $n-p$ stupni volnosti. Kritickou hodnotou je kvantil $t_{1-\frac{\alpha}{2}}(n-p)$.

Pokud nelze H_0 zamítnout, je třeba uvážit setrvání příslušného koeficientu v modelu. Pokud vyjde u regresní přímky koeficient a statisticky nevýznamný, znamená to, že proměnné nejsou závislé. Závěrem je třeba poznamenat, že vyřazení (či nové zařazení) proměnné do modelu znamená spustit celý proces tvorby modelu od začátku, tzn. nový odhad regresních parametrů.

77 - Korelační analýza

Řešené příklady: 160
Příklady: 234,236

10.5 Korelační analýza

Korelace určuje míru (intenzitu) závislosti statistických proměnných.

Korelační koeficient nabývá hodnot $\langle -1, 1 \rangle$ a měří těsnost lineární závislosti dvou proměnných. Jde o ukazatel oboustranné závislosti. Nejdůležitější a nejčastěji používanou mírou síly vztahu je **Pearsonův korelační koeficient** $\rho_{x,y}$

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \cdot \sigma_y} = \frac{\sum_i [(x_i - \bar{x})(y_i - \bar{y})]}{\sum_i (x_i - \bar{x}) \cdot \sum_i (y_i - \bar{y})},$$

kde σ_x, σ_y jsou směrodatné odchylky a $\text{cov}(x,y)$ je kovariance proměnných x, y . Předpokládáme normalitu obou proměnných.

Lineární vztah náhodných veličin x a y kvantifikujeme na základě výběrového souboru -**výběrový Pearsonův korelační koeficient** r

$$r = \frac{\sum_i x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y}.$$

Čím více se hodnota korelačního koeficientu blíží k hodnotě jedna, tím je závislost silnější, obě hodnoty společně rostou.

Čím blíže je hodnota korelačního koeficientu k hodnotě -1 , tím je závislost silnější, rostou-li hodnoty jedné proměnné, hodnoty druhé klesají.

Je-li hodnota blízká nule, nejsou proměnné závislé.

O síle závislosti vypovídá nejen korelační koeficient, ale i rozsah souboru n .

Slovní interpretace:

$ r = 0$	korelační nezávislost
$0 < r < 0.3$	nízký stupeň korelační závislosti
$0.3 \leq r < 0.5$	mírný stupeň korelační závislosti
$0.5 \leq r < 0.7$	střední stupeň korelační závislosti
$0.7 \leq r < 0.9$	vysoký stupeň korelační závislosti
$0.9 \leq r < 1$	velmi vysoký stupeň korelační závislosti
$ r = 1$	funkční závislost

Test hypotézy o nulové korelaci dvou náhodných veličin:

$$H_0 : r = 0$$

$$H_1 : r \neq 0$$

Testovací kritérium:

$T = r \sqrt{\frac{n-2}{1-r^2}}$...testovací statistika má Studentovo rozdělení pravděpodobnosti s $n - 2$ stupni volnosti.

EXCEL:

=CORREL(matice1;matice1)

Matematika III: Řešené příklady

Viktor Dubovský, Marcela Jarošová, Jiří Krček, Jitka Krčková, Petra Schreiberová, Petr Volný

Katedra matematiky a deskriptivní geometrie

VŠB - Technická univerzita Ostrava

Řešené příklady – Kombinatorika

80 - Kombinatorické pravidlo součtu

Zadání Pracovní četa je složena z 15 pracovníků, z toho 10 umí zacházet s lopatou, 7 s krumpáčem a 9 s rýčem, lopatu a krumpáč ovládají 3 muži, lopatu a rýč 6 a krumpáč s rýčem 4 muži, všechny tři nástroje zvládá jediný. Je ve skupině muž, který nepracuje s žádným z nástrojů?

Řešení

Teorie: 11 Příklady: 163,165

Před samotným řešením si zavedeme značení. Necht' $\mathbb{V}$ značí vedoucího, $\mathbb{L}$ značí ty, kteří umí s lopatou, $\mathbb{LK}$ ty, kteří pracují s lopatou i krumpáčem, $\mathbb{LKR}$ ty, kteří umí se vším. Obdobně $\mathbb{K}$ pro „krumpáče“ a tak podobně.

Sečteme-li počty členů $\mathbb{L}, \mathbb{K}, \mathbb{R}$ dostáváme $\mathbb{L} + \mathbb{K} + \mathbb{R} = 10 + 7 + 9 = 26$. To zjevně nemůže být správný výsledek, protože četa je tvořena 15 muži. Důvodem chyby je fakt, že muži z $\mathbb{LK}$ jsou započteni dvakrát, jak ve skupině „lopat“ $\mathbb{L}$, tak ve skupině „krumpáčů“ $\mathbb{K}$, a stejně tak muži z $\mathbb{LR}$ a $\mathbb{KR}$.

K tomu abychom toto jejich „zdvojení“ napravili, stačí počty členů $\mathbb{LK}, \mathbb{LR}$ a $\mathbb{KR}$ odečíst, a to následovně

$$\mathbb{L} + \mathbb{K} + \mathbb{R} - \mathbb{LK} - \mathbb{LR} - \mathbb{KR} = 10 + 7 + 9 - 3 - 6 - 4 = 13.$$

Může se zdát, že v pracovní četě je 13 mužů, kteří umí s alespoň jedním z nástrojů. Ani toto však není správný výsledek. Při „odčítání“ $\mathbb{LK}$ jsme samozřejmě odečetli i $\mathbb{LKR}$, totéž platí i pro $\mathbb{LR}$ a $\mathbb{KR}$.

Takže zatímco v $\mathbb{L} + \mathbb{K} + \mathbb{R}$ bylo $\mathbb{LKR}$ započteno třikrát, v $\mathbb{L} + \mathbb{K} + \mathbb{R} - \mathbb{LK} - \mathbb{LR} - \mathbb{KR}$ není $\mathbb{LKR}$ započteno vůbec. Řešení je jednoduché, připočíst jej,

$$\mathbb{L} + \mathbb{K} + \mathbb{R} - \mathbb{LK} - \mathbb{LR} - \mathbb{KR} + \mathbb{LKR} = 10 + 7 + 9 - 3 - 6 - 4 + 1 = 14.$$

Správným výsledkem je, že v četě je 14 mužů, kteří umí zacházet s alespoň jedním z nástrojů. Protože má četa celkem 15 mužů, je jasné, že je v ní $15 - 14 = 1$ člen, který neumí ani s lopatou, ani krumpáčem, ani s rýčem.

Při řešení jsme používali jednoduchého *kombinatorického pravidla součtu*, které nám říká, že počet členů skupiny, tvořené skupinami, které nemají společné členy, je roven součtu členů jednotlivých skupin.

Podmínka neexistence společných členů je zde zásadní a její vynechání může vést k nesprávným výsledkům, viz. $\mathbb{L} + \mathbb{K} + \mathbb{R} = 10 + 7 + 9 = 26 > 15$. Při jejím splnění však může být aplikace pravidla až triviální, viz. „pracují s něčím“ + „pracují s ničím“ jsou „všichni“, tj. $14 + 1 = 15$.

Poznámky

Kombinatorické pravidlo součtu

Necht' $A_1, A_2, \dots, A_k$ jsou konečné množiny, které mají $p_1, p_2, \dots, p_k$ prvků, a které jsou po dvou vzájemně disjunktní, tj. nemají společný prvek, pak počet prvků sjednocení $A_1 \cup A_2 \cup \dots \cup A_k$ je roven součtu

$$p_1 + p_2 + \dots + p_k.$$

81 - Kombinatorické pravidlo součinu

Zadání Kolik možností musíme vyzkoušet, pamatujeme-li si ze čtyřmístného (číselného) PINu, že začíná číslicí 2 nebo 3, neobsahuje číslici 7 a poslední číslice je lichá.

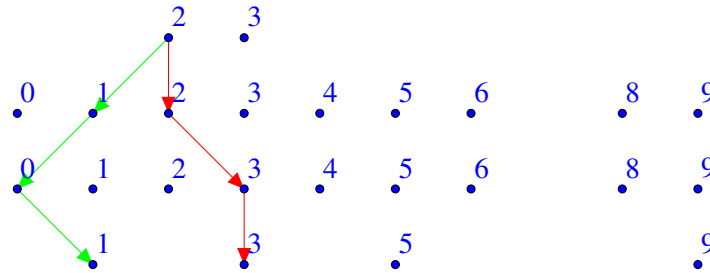
Řešení

Teorie: 11 Příklady: 164

Protože víme, že náš PIN neobsahuje číslici 7 zbývá nám již jen devět možností, jak jednotlivé pozice obsadit, a to číslice 0, 1, 2, 3, 4, 5, 6, 8, 9.

Dále si pamatujeme, že první číslicí je 2 nebo 3, čímž se výběr na této pozici zužuje pouze na dvě možnosti.

Konečně na poslední pozici může figurovat pouze některá z číslic 1, 3, 5, 9, neboť víme, že jde o liché číslo, které však neobsahuje číslici 7. Máme tedy čtyři možnosti pro poslední místo PINu.



V náčrtku jsou naznačeny dva z PINů vyhovujících podmínkám, a to 2101 a 2233. Důležitější však je, že si zde lze znovu uvědomit a lépe představit dvě možnosti pro první pozici, devět pro druhou a třetí pozici a čtyři možné volby pro pozici poslední.

Celkově tedy k prověření zůstává

$$2 \cdot 9 \cdot 9 \cdot 4 = 8 \cdot 9^2 = 648$$

možných složení PINu.

Postup, který jsme při řešení použili, ilustroval *kombinatorické pravidlo součinu*. Tedy fakt, že „celkový počet možností je roven součinu počtu možností na jednotlivých místech.“

Úplným závěrem poznamejme, že počet všech možných čtyřmístných (číselných) PIN kódů určíme dle návodu v příkladě 85, jde totiž o variace s opakováním. Zde máme 10 možností, číslice 0, 1, ..., 9, které můžeme umístit na čtyři místa. Proto lze sestavit $10 \cdot 10 \cdot 10 \cdot 10 = 10^4 = 10000$ různých PINů.

Poznámky

Kombinatorické pravidlo součinu

Počet všech uspořádaných k -tic, jejichž první člen lze vybrat z n_1 možných prvků, druhý člen z n_2 prvků, atd. až k -tý člen z n_k prvků, je rovný součinu

$$n_1 \cdot n_2 \cdot \dots \cdot n_k.$$

82 - Permutace bez opakování

Zadání Výběrového řízení se účastní 3 firmy. Určete počet všech možných výsledků soutěže.

Řešení

Teorie: 12 Příklady: 166

Máme určit počet trojic ze 3 odlišných prvků, tzn. neprovádíme výběr, ale jen různá uspořádání prvků $\Rightarrow$ permutace bez opakování

$$P(3) = 3! = 3 \cdot 2 \cdot 1 = 6.$$

Může dojít k šesti možným výsledkům výběrového řízení.

Detailněji:

Pro přehlednost si označme jednotlivé firmy písmeny A, B, C a pokusme se možné výsledky „modelovat“.

Uvažujme nejprve případ, že vítězem bude firma A . Na druhém místě tedy bude jedna z firem B, C . Bude-li druhou firma B , pak pro firmu C nezbývá jiná možnost než skončit na místě třetím. Výsledek celé soutěže si zapišme (například) ve tvaru $A - B - C$. Stejně tak se ovšem při vítězství firmy A , mohla na druhém místě objevit firma C a třetí by tedy skončila firma B . Tj. výsledek by byl ve tvaru $A - C - B$.

Pokračujme předpokladem vítězství firmy B . Na druhém místě tak bude jedna z firem A, C . Možné výsledky tedy jsou $B - A - C$ a $B - C - A$. Obdobně lze „rozebrat“ situaci při výhře firmy C . Ta vede k výsledkům $C - A - B$ a $C - B - A$.

Závěrem lze konstatovat, že existuje šest možných výsledků výběrového řízení, jehož se zúčastní tři firmy, a to

$$A - B - C, A - C - B, B - A - C, B - C - A, C - A - B, C - B - A.$$

Uvědomme si však, že obdobný postup bude jen těžko aplikovatelný při vyšším počtu účastníků soutěže. To však neznamená, že v něm nelze nalézt návod pro řešení.

Nejprve jsme vybrali vítěze, a to ze tří možností A, B, C . Na druhém místě se pak mohla umístit jedna z dvojice dosud nevybraných firem. Po výběru druhého místa, už zbyla jediná možnost pro „volbu“ místa třetího. Tři možnosti pro vítěze, dvě pro druhé místo a jedna pro třetí dohromady dávají $3 \cdot 2 \cdot 1 = 6$ možných výsledků.

Poznámky

Permutace bez opakování

Permutace z n prvků je uspořádaná n -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje právě jednou. Značí se $P(n)$ a platí, že

$$P(n) = n!.$$

Faktoriál

$$n! = \prod_{i=1}^n i,$$

$$0! = 1.$$

83 - Variace bez opakování

Zadání V sedmičlenné správní radě se volí předseda a místopředseda. Určete počet všech možností, jak tyto posty rozdělit.

Řešení

Teorie: 12 Příklady: 167

Zajímá nás počet možných dvojic ze 7 kandidátů (vybíráme dvojice ze 7 prvků). Jeden kandidát nemůže současně obsadit obě pozice (bez opakování) a záleží na pořadí ve dvojici (je rozdíl být předseda, nebo místopředseda) $\Rightarrow$ variace bez opakování

$$V_2(7) = \frac{7!}{(7-2)!} = 7 \cdot 6 = 42.$$

Existuje 42 možných výsledků voleb.

Detailněji:

Máme k dispozici 7 kandidátů na předsedu, a po jeho zvolení zbývá jen 6 možností, jak volit místopředsedu. Dalším pořadím kandidátů se není třeba zabývat, neboť oba posty jsou již obsazeny.

Celkový počet možností lze vyjádřit jako součin možností obsazení jednotlivých postů, tj. máme $7 \cdot 6 = 42$ možných výsledků.

I když oproti *permutacím* (viz. 82) neuvažujeme pořadí (uspořádání) všech účastníků (prvků), můžeme i zde vhodně využít faktoriálu.

Uvědomíme-li si, že $7!$ je počet možných pořadí všech zúčastněných a $5!$ je počet možných pořadí „pětice, která nás nezajímá“, pak počet obsazení dvou míst získáme jako podíl těchto čísel. Tedy jako

$$\frac{7!}{5!} = \frac{7 \cdot 6 \cdot 5!}{5!} = 7 \cdot 6.$$

Odsud již lze dovodit, jak zjistit počet možných obsazení k pozic v n členném týmu. Stačí spočítat

$$\frac{n!}{(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k)!}{(n-k)!} = n \cdot (n-1) \cdots (n-k+1),$$

kde $n!$ reprezentuje počet všech možných pořadí a $(n-k)!$ počet pořadí na místech, která pro nás nejsou důležitá.

Poznámky

Variace bez opakování

k -členná variace z n prvků je uspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše jednou. Značí se $V_k(n)$ a platí, že

$$V_k(n) = \frac{n!}{(n-k)!}.$$

84 - Kombinace bez opakování

Zadání Z pětičlenné skupiny pracovníků je potřeba vytvořit dvojici a trojici. Kolika způsoby to lze učinit?

Řešení

Teorie: 12 Příklady: 168

Vybíráme dvojici z pěti prvků, nezáleží nám na pořadí ve vybrané dvojici a vybraný pracovník už nemůže být do dvojice vybrán podruhé (prvky se nemohou opakovat) $\Rightarrow$ kombinace bez opakování

$$C_2(5) = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = 10.$$

Lze postupovat i tak, že budeme vybírat z pěti prvků trojice (vlastnost symetrie kombinačních čísel).

$$C_3(5) = \binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5!}{3!2!} = 10.$$

Detailněji:

Nejprve vybereme do dvojice Aloise s Bedřichem, trojici pak tvoří Cyril, Diviš a Emil. Mohlo by se tedy zdát, že můžeme postupovat stejně jako při volbě předsedy a místopředsedy z příkladu na straně 83, ovšem není tomu tak. Zde totiž nezáleží na pořadí výběru, protože dvojice Alois-Bedřich je stejná jako dvojice Bedřich-Alois. Proto musíme naše úvahy upravit.

Začít můžeme stejně a sice výpočtem všech možných uspořádání pětice, což je $5!$. Protože pro nás pořadí v trojici pro nás není důležité, vydělíme počtem těchto pořadí, tj. $3!$. Pořadí v dvojici taky není rozhodující, proto opět dělíme, tenkrát $2!$. Tak dostáváme

$$\frac{5!}{3! \cdot 2!} = \frac{5 \cdot 4 \cdot \cancel{3!}}{\cancel{3!} \cdot 2 \cdot 1} = \frac{5 \cdot 4}{2} = 10.$$

Z pětice lze tedy dvojici vybrat deseti způsoby.

Poznámky

Kombinace bez opakování

k -členná kombinace z n prvků je neuspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše jednou.

Značí se $C_k(n)$ a platí, že

$$C_k(n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Základní vlastnosti kombinačních čísel

$$\binom{n}{0} = 1,$$

$$\binom{n}{n} = 1,$$

$$\binom{n}{1} = n.$$

Symetrie

$$\binom{n}{k} = \binom{n}{n-k}.$$

85 - Variace s opakováním

Zadání Kolika způsoby lze teoreticky zaměstnanci naplánovat směny (ranní, odpolední, noční) na pětidenní pracovní týden.

Řešení

Teorie: 13 Příklady: 169

Vybíráme pěti z 3 prvků (směn), prvky se musí opakovat a záleží na pořadí směn ve výběru $\Rightarrow$ variace s opakováním

$$V'_5(3) = 3^5.$$

Detailněji:

V pondělí lze volit ze tří možností - ranní, odpolední nebo noční. Stejně tak však lze vybírat směnu úterní, tj. ze všech tří možností. Obdobně pro středu, čtvrtek i pátek. Celkový počet možností sestavení pracovního týdne je

$$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5 = 243.$$

Pokud bychom volili pouze mezi ranní odpolední směnou, byl by počet všech možností roven $2^5 = 32$. Pokud bychom ranní, odpolední a noční plánovali na 22 pracovních dní, pak je 3^{22} způsobů jak plán sestavit.

Obecně lze tedy říci, že máme-li k -krát volit z n prvků, přičemž nám při této volbě záleží na pořadí volby, ale opakování je dovoleno, pak máme

$$n^k$$

možností, jak volbu provést.

Poznámky

Variace s opakováním

k -členná variace s opakováním z n prvků je uspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše k -krát. Pro jejich počet platí

$$V'_k(n) = n^k.$$

86 - Permutace s opakováním

Zadání Kolika způsoby lze rozdělit pět kopáčů do tří výkopů?

Řešení

Teorie: 13 Příklady: 170

Tvoříme různá uspořádání pěti kopáčů do tří výkopů.

Můžeme si pomoci nákresem jednoho takového uspořádání: . Tento náčrtek znázorňuje dva kopáče v prvním, jednoho v druhém a dva ve třetím výkopu.

Tvoříme tedy různá uspořádání 7 prvků, ve kterém se vyskytují 2 znaky, kopáč  a dělení -. Jedná se o permutace s opakováním ze dvou prvků, kde první prvek se opakuje pětkrát a druhý dvakrát.

$$P'(5, 2) = \frac{(5+2)!}{5! \cdot 2!} = \frac{7!}{5! \cdot 2!} = \frac{7 \cdot 6 \cdot \cancel{5!}}{\cancel{5!} \cdot 2!} = \frac{7 \cdot 6}{2} = 21.$$

Poznámka:

Počet uspořádání sedmi prvků je 7!. Protože nerozlišujeme mezi jednotlivými kopáči , musíme výsledek dělit všemi možnostmi jejich pořadí, tj. 5!. Stejně tak nerozlišujeme symboly -, proto dělíme ještě 2!.

Uvědomme si, že výsledných 21 možností zahrnuje také rozdělení, při nichž v některém z výkopů nebude žádný z kopáčů. Například  - - tak znázorňuje situaci, kdy všech pět mužů pracuje v prvním výkopu.

Uvažujme nyní, že skupina kopáčů je složena ze dvou pracovníků firmy A, , , a dvou z firmy B, ,  a jednoho  z firmy C. Kolik možných pořadí, v nichž se dostaví na pracoviště existuje, pokud nebudeme rozlišovat jednotlivé muže, ale jen jejich firemní příslušnost? Lehce modifikovanou úvahou dospějeme k výsledku

$$\frac{(2+2+1)!}{2! \cdot 2! \cdot 1!} = \frac{5!}{4} = \frac{5 \cdot \cancel{4} \cdot 3 \cdot 2 \cdot 1}{\cancel{4}} = 30.$$

Obecně, máme-li uspořádat n prvků tvořených n_1 nerozlišitelnými prvky „druhu 1“, n_2 nerozlišitelnými prvky „druhu 2“, až n_k nerozlišitelnými prvky „druhu k “, pak lze počet možností určit následujícím vzorcem

$$\frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}.$$

Poznámky

Permutace s opakováním

Základní množina má n prvků, které jsou k různých typů. Přitom $n_1, \dots, n_k$ značí počet prvků daného typu a platí $n = n_1 + n_2 + \dots + n_k$. Permutace s opakováním je uspořádaná n -tice, z těchto prvků sestavená. Značí se $P'(n_1, n_2, \dots, n_k)$ a jejich počet je

$$P'(n_1, n_2, \dots, n_k) = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}.$$

87 - Kombinace s opakováním

Zadání Kolika způsoby lze 8 dělníků vybavit „montérkami“, máme-li k dispozici 20 ks modrých a 10 ks červených sad.

Řešení

Teorie: 13 Příklady: 171

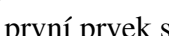
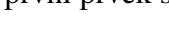
Při rozdělování a vydávání „montérek“ nezáleží pořadí, ve kterém se dělníci dostaví, ani na pořadí v němž jsou přiděleny jednotlivé barvy.

Vybíráme 8 prvků ze dvou možností (modrá, červená), nezáleží na pořadí a musí se opakovat $\Rightarrow$ kombinace s opakováním.

$$C'_8(2) = \binom{2+8-1}{8} = \frac{(2+8-1)!}{8! \cdot 1!} = \frac{9 \cdot \cancel{8!}}{\cancel{8!}} = 9.$$

Poznámka:

S určením počtu možností nám také může pomoci příklad 86.

I zde si lze pomoci obrázkem:  - , který představuje jedno z možných „dělení“. Nyní stačí pouze jednotlivé skupiny „obarvit“, například takto  - . Máme tedy opět dvě skupiny prvků, dělníků  a dělení -. Jedná se o permutace s opakováním ze dvou prvků, kde první prvek se opakuje osmkrát a druhý jedenkrát.

$$P'(8, 1) = \frac{(8+1)!}{8! \cdot 1!} = \frac{9 \cdot \cancel{8!}}{\cancel{8!}} = 9.$$

Poznamenejme ještě, že počet kusů jednotlivých sad nám zajišťuje, že můžeme sestavit opravdu všechny možnosti. Bylo by-li na skladě jen 5 ks červených sad, byla by uvedena varianta, , samozřejmě vyloučena.

Nyní se již můžeme pokusit postup zobecnit, tj. obléci k mužů do n různých barev. Opět si lze představit dělení do n skupin pomocí $(n-1)$ „rozdělovníků“. Celkem určujeme pořadí $k + (n-1)$ prvků, tj. $(k + (n-1))!$. Protože však nerozlišujeme jednotlivé „muže“ ani „rozdělovníky“, dělíme počtem všech možných jejich pořadí, tj. $k!$ a $(n-1)!$. Dostáváme tedy vztah

$$P'(k, n-1) = \frac{(k+n-1)!}{k! (n-1)!} = \frac{(n+k-1)!}{k! (n+k-1-k)!} = \binom{n+k-1}{k} = C'_k(n).$$

Poznámky

Kombinace s opakováním

k -členná kombinace s opakováním z n prvků je neuspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše k -krát.

Značí se $C'_k(n)$ a jejich počet je

$$\begin{aligned} C'_k(n) &= \binom{n+k-1}{k} \\ &= \frac{(n+k-1)!}{k! (n-1)!}. \end{aligned}$$

88 - Vlastnosti kombinačních čísel I

Zadání Kolika způsoby lze ze šesti mužů vybrat čtveřici k provedení úklidu staveniště?

Řešení

Teorie: 14 Příklady: 172

Při výběru nám nejde o pořadí, ani mezi muži nerozlišujeme, nelze také do čtveřice vybrat jednoho muže dvakrát či dokonce ještě vícekrát. Jedná se tedy o kombinace bez opakování a můžeme si pomoci příkladem 84.

Hledaný počet možností tak určíme pomocí kombinačního čísla $\binom{6}{4}$, tj.

$$\binom{6}{4} = \frac{6!}{4!(6-4)!} = \frac{6!}{4!2!} = \frac{6 \cdot 5 \cdot \cancel{4!}}{\cancel{4!}2!} = \frac{30}{2} = 15.$$

Lze tedy sestavit 15 čtveřic určených k úklidu. Zbývající dva muže můžeme využít k jiné práci nebo je poslat domů.

A kolika způsoby lze vybrat dvojici šťastlivců, kterým padla? Jedná se samozřejmě opět o kombinace bez opakování a jejich počet je $\binom{6}{2}$ a spočteme jej následovně

$$\binom{6}{2} = \frac{6!}{2!(6-2)!} = \frac{6!}{2!4!} = \frac{6 \cdot 5 \cdot \cancel{4!}}{2!\cancel{4!}} = \frac{30}{2} = 15.$$

Že je výsledek stejný by nás nemělo nikterak překvapit. Vybrat ze šesti mužů dva, kteří odejdou, jasně určuje čtveřici, která zůstane.

Takto jsme si pomocí jednoduché úvahy odvodili důležitou vlastnost kombinačních čísel, kterou nazýváme symetrií, tj. vztah

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-(n-k))!(n-k)!} = \binom{n}{n-k}.$$

Poznámky

Vlastnosti kombinačních čísel

symetrie

$$\binom{n}{k} = \binom{n}{n-k}.$$

89 - Vlastnosti kombinačních čísel II

Zadání Kolika způsoby lze z 8 studentů vytvořit trojici nebo čtveřici?

Řešení

Teorie: 14 Příklady: 172

Protože jde o kombinace bez opakování využijeme kombinačních čísel. Počet možných trojic je $\binom{8}{3}$ a čtveřic $\binom{8}{4}$.

A celkový počet možností proto je $\binom{8}{3} + \binom{8}{4}$, tj.

$$\begin{aligned}\binom{8}{3} + \binom{8}{4} &= \frac{8!}{3! (8-3)!} + \frac{8!}{4! (8-4)!} = \frac{8!}{3! 5!} + \frac{8!}{4! 4!} = \frac{8!}{3! 5 \cdot 4!} + \frac{8!}{4 \cdot 3! 4!} \\ &= \frac{8!}{3! 4!} \left(\frac{1}{5} + \frac{1}{4} \right) = \frac{8!}{3! 4!} \cdot \frac{4+5}{5 \cdot 4} = \frac{9 \cdot 8!}{5 \cdot 4 \cdot 3! 4!} = \frac{9!}{5! \cdot 4!} = \frac{9!}{4! (9-4)!} = \binom{9}{4}.\end{aligned}$$

Zbývá již jen vyčíslit, že $\binom{9}{4} = 126$. Ke stejnému výsledku dospějeme také po určení toho, že $\binom{8}{3} = 56$ a $\binom{8}{4} = 70$.

Předchozí výpočet tak slouží spíše jako návod k odvození další důležité vlastnosti kombinačních čísel, a sice

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

Poznámky

Vlastnosti kombinačních čísel

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

90 - Úpravy výrazů s kombinačními čísly I

Zadání Vyjádřete $\binom{19}{4} + \binom{19}{16} + \binom{20}{5}$ jediným kombinačním číslem.

Řešení

Teorie: 14 Příklady: 172

Využijeme vlastnosti symetrie kombinačních čísel a samozřejmě také vzorec pro jejich součet. Takto dostaneme

$$\binom{19}{4} + \binom{19}{16} + \binom{20}{5} = \binom{19}{4} + \binom{19}{19-16} + \binom{20}{5} = \underbrace{\binom{19}{4} + \binom{19}{3}}_{\binom{20}{4}} + \binom{20}{5} = \binom{20}{4} + \binom{20}{5} = \binom{21}{5}.$$

Provedení naznačených úprav není nikterak složité a vynaložená energie se může vrátit v podobě časové úspory při výpočtu číselné hodnoty součtu. Je totiž značný rozdíl mezi výpočtem

$$\begin{aligned} \binom{19}{4} + \binom{19}{16} + \binom{20}{5} &= \frac{19 \cdot \cancel{18} \cdot 17 \cdot \cancel{16}^4}{\cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot 1} + \frac{19 \cdot \cancel{18}^3 \cdot 17}{\cancel{3} \cdot \cancel{2} \cdot 1} + \frac{\cancel{20} \cdot 19 \cdot \cancel{18}^3 \cdot 17 \cdot 16}{\cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot 1} \\ &= 19 \cdot 3 \cdot 17 \cdot 4 + 19 \cdot 3 \cdot 17 + 19 \cdot 3 \cdot 17 \cdot 16 = 19 \cdot 3 \cdot 17 \cdot (4 + 1 + 16) = 21 \cdot 19 \cdot 3 \cdot 17, \end{aligned}$$

a výpočtem

$$\binom{21}{5} = \frac{21 \cdot \cancel{20} \cdot 19 \cdot \cancel{18}^3 \cdot 17}{\cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot 1} = 21 \cdot 19 \cdot 3 \cdot 17.$$

Výsledek, 20349, je ovšem samozřejmě v obou případech stejný.

Poznámky

Kombinační čísla

základní

$$\binom{n}{0} = 1,$$

$$\binom{n}{n} = 1,$$

$$\binom{n}{1} = n.$$

symetrie

$$\binom{n}{k} = \binom{n}{n-k}.$$

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

91 - Úpravy výrazů s kombinačními čísly II

Zadání Vyjádřete $\binom{2}{2} + \binom{3}{2} + \binom{4}{2} + \binom{5}{2}$ jediným kombinačním číslem.

Řešení**Teorie: 14 Příklady: 172**

Nejprve využijeme faktu, že $\binom{n}{n} = 1$, pro $\forall n \in \mathbb{N}$ a tedy, že $\binom{2}{2} = \binom{3}{3}$. Dále budeme používat vzorec pro součet kombinačních čísel. Takto dostaneme

$$\binom{2}{2} + \binom{3}{2} + \binom{4}{2} + \binom{5}{2} = \underbrace{\binom{3}{3} + \binom{3}{2}}_{\binom{4}{3}} + \binom{4}{2} + \binom{5}{2} = \underbrace{\binom{4}{3} + \binom{4}{2}}_{\binom{5}{3}} + \binom{5}{2} = \binom{5}{3} + \binom{5}{2} = \binom{6}{3}.$$

Pokusme se obdobně vyjádřit součet $\binom{2}{2} + \binom{3}{2} + \dots + \binom{n}{2}$ pro $n \in \mathbb{N}$.

$$\binom{2}{2} + \binom{3}{2} + \dots + \binom{n}{2} = \underbrace{\binom{3}{3} + \binom{3}{2}}_{\binom{4}{3}} + \binom{4}{2} + \dots + \binom{n}{2} = \dots = \binom{n}{3} + \binom{n}{2} = \binom{n+1}{3}.$$

Na závěr zopakujme zadání pro součet $\binom{k}{k} + \binom{k+1}{k} + \dots + \binom{n}{k}$ pro $k, n \in \mathbb{N}$ a $k \leq n$.

$$\binom{k}{k} + \binom{k+1}{k} + \dots + \binom{n}{k} = \underbrace{\binom{k+1}{k+1} + \binom{k+1}{k}}_{\binom{k+2}{k+1}} + \binom{k+2}{k} + \dots + \binom{n}{k} = \dots = \binom{n}{k+1} + \binom{n}{k} = \binom{n+1}{k+1}.$$

Můžeme tedy psát

$$\sum_{i=k}^n \binom{i}{k} = \binom{n+1}{k+1}.$$

Poznámky*Kombinační čísla**základní*

$$\binom{n}{0} = 1,$$

$$\binom{n}{n} = 1,$$

$$\binom{n}{1} = n.$$

symetrie

$$\binom{n}{k} = \binom{n}{n-k}.$$

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

92 - Rovnice s kombinačními čísly

Zadání Vyřešte rovnici $\binom{x}{x-1} + \binom{x+3}{x+1} = 12$.

Řešení

Teorie: 14 Příklady: 173

Nejprve využijeme vlastností kombinačních čísel k úpravám a zjednodušení rovnice

$$\begin{aligned} \binom{x}{x-1} + \binom{x+3}{x+1} &= 12 \\ \binom{x}{x-(x-1)} + \binom{x+3}{(x+3)-(x+1)} &= 12 && \text{symetrie} \\ \binom{x}{1} + \binom{x+3}{2} &= 12 && \text{symetrie} \\ x + \frac{(x+3)!}{2!(x+1)!} &= 12 && \text{základní} \\ x + \frac{(x+3)(x+2)\cancel{(x+1)!}}{2\cancel{(x+1)!}} &= 12 \\ x + \frac{(x+3)(x+2)}{2} &= 12 \\ 2x + (x+3)(x+2) &= 24 \\ 2x + x^2 + 5x + 6 &= 24 \\ x^2 + 7x - 18 &= 0. \end{aligned}$$

Po úpravách jsme dospěli ke kvadratické rovnici, která má dvě řešení, a to $x_1 = -9$ a $x_2 = 2$. Přičemž řešení x_1 zavrhneme, neboť $-9 \notin \mathbb{N}$ a pro takové hodnoty nemá kombinační číslo smysl.

Provedení zkoušky

$$\binom{2}{2-1} + \binom{2+3}{2+1} = \binom{2}{1} + \binom{5}{3} = 2 + \frac{5 \cdot 4 \cdot \cancel{3!}}{\cancel{3!} 2!} = 2 + 10 = 12,$$

potvrzuje správnost řešení $x_2 = 2$.

Poznámky

Kombinační čísla

základní

$$\begin{aligned} \binom{n}{0} &= 1, \\ \binom{n}{n} &= 1, \\ \binom{n}{1} &= n. \end{aligned}$$

symetrie

$$\binom{n}{k} = \binom{n}{n-k}.$$

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

93 - Využití více kombinatorických postupů, pravidel a vzorců

Poznámky

Zadání Na lístku je 29 čárek. Kolika způsoby lze rozdělit útratu mezi 5 kamarádů, kteří si pamatují jen, že každý z nich měl alespoň 5 piv?

Řešení

Teorie: 11,12,13,14 Příklady: 163-176

Žádný z přátel nevypil méně než pět piv, což dohromady činí 25 kousků. Otázkou tedy je, kdo vypil zbývající čtyři. Nabízí se několik možností

- i) jeden z přátel vypil o 4 víc než ostatní,
- ii) jeden měl o 3 víc, další 1 navíc, ostatní zůstali na pěti
- iii) dva vypili o 2 víc než ostatní,
- iv) jeden vypil o 2 víc, dva vypili o jedno víc, dva měli pět,
- v) čtyři si dali jedno navíc a jeden už ne.

Těchto pět možností odpovídá rozkladům čísla 4 na součet přirozených čísel. Ovšem neodpovídají na otázku kolika způsoby lze útratu rozdělit, spíše nás navádí k tomu, jak ji dělit. K určení počtu rozebereme jednotlivé případy.

V případě **i)** z pětky přátel vybíráme jednoho, a to lze $\binom{5}{1} = 5$ způsoby.

V případě **ii)** vybíráme jednoho s 8 pivy, což lze 5 způsoby, a poté jednoho s 6, kterého však vybíráme již jen ze čtveřice, a tedy 4 způsoby. Celkem tedy $5 \cdot 4 = 20$ způsobů.

Případ **iii)** odpovídá hledání dvojice se 7 vypitými, takže $\binom{5}{2} = 10$ možností.

Pokud bychom i ve variantě **iv)** pouze vybrali trojici z pěti $\binom{5}{3} = 10$, nebude výsledek správný, neboť ve vybrané trojici se liší počty vypitých piv. Protože je tedy v trojici důležité uspořádání, určíme si jeho počet. Jde o permutace s opakováním, a to prvků 2, 1, 1 a jejich počet je $\frac{(1+2)!}{1!2!} = 3$. Celkem tedy $3 \cdot 10 = 30$ možností pro variantu **iv)**.

Stejně tak jsme pro **iv)** mohli 5 způsoby vybrat kamaráda se dvěma navíc a zbylých 4 volit dva s jedním navrch, tj. $5 \cdot \binom{4}{2} = 30$.

Dalším způsobem je volba dvojice z pěti a poté jednoho ze tří, tj. $\binom{5}{2} \cdot 3 = 30$.

Nakonec vybíráme čtveřici, která si dala jedno navíc, což je stejné jako volba jednoho, který si nedal, je totiž $\binom{5}{4} = \binom{5}{1} = 5$.

Zbývá už jen sečíst počty možností v jednotlivých variantách, tj. $5 + 20 + 10 + 30 + 5 = 70$. Přátelé tedy mají na výběr ze 70 možností, jak se o útratu podělit.

Řešené příklady – Pravděpodobnost

95 - Náhodný pokus, náhodný jev

Zadání Náhodným pokusem je vytažení 1 karty z „pokerového“ balíčku 52 karet. Popište základní prostor elementárních jevů Ω , zajímá-li nás pouze „barva“ karty, tj. zda jde o $\spadesuit$, $\heartsuit$, $\clubsuit$ nebo $\diamondsuit$. Vypište všechny elementární jevy ω_i a všechny náhodné jevy A_i .

Řešení

Teorie: 16,17 Příklady: 178,179

Začneme výpisem všech elementárních jevů ω_i , tj. výpisem všech možných výsledků náhodného pokusu. Vybíráme-li si z balíčku jednu kartu, může být srdcová. Označme tedy ω_1 elementární jev vytažení srdce, $\omega_1 = \{\heartsuit\}$. A obdobně pro tažení dalších barev, $\omega_2 = \{\spadesuit\}$, $\omega_3 = \{\diamondsuit\}$ a $\omega_4 = \{\clubsuit\}$. Základní prostor elementárních jevů Ω tedy je $\Omega = \{\heartsuit, \spadesuit, \diamondsuit, \clubsuit\}$.

Náhodným jevem je každá podmnožina základního prostoru elementárních jevů Ω . Jednou z takových podmnožin je například množina $\{\heartsuit, \diamondsuit\}$, která odpovídá tomu, že byla vytažena červená karta, tedy srdce nebo káry.

Protože základní prostor Ω je čtyřprvkový, existují $\binom{4}{1} = 4$ jeho jednoprvkové, $\binom{4}{2} = 6$ dvouprvkové, $\binom{4}{3} = 4$ tříprvkové a $\binom{4}{4} = 1$ čtyřprvková podmnožina. Spolu s prázdnou podmnožinou, tak máme $\binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 2^4 = 16$ podmnožin, resp. možných náhodných jevů, které souvisí s náhodným pokusem „tažení karty.“ Jsou to

- $A_0 = \emptyset$ - nevytažení žádné karty, nemožný jev,
- $A_1 = \omega_1 = \{\heartsuit\}$, $A_2 = \omega_2 = \{\spadesuit\}$, $A_3 = \omega_3 = \{\diamondsuit\}$, $A_4 = \omega_4 = \{\clubsuit\}$ - elementární jevy,
- $A_5 = \{\heartsuit, \spadesuit\}$, $A_6 = \{\heartsuit, \diamondsuit\}$, $A_7 = \{\heartsuit, \clubsuit\}$, $A_8 = \{\spadesuit, \diamondsuit\}$, $A_9 = \{\spadesuit, \clubsuit\}$, $A_{10} = \{\diamondsuit, \clubsuit\}$,
- $A_{11} = \{\heartsuit, \spadesuit, \diamondsuit\}$, $A_{12} = \{\heartsuit, \spadesuit, \clubsuit\}$, $A_{13} = \{\heartsuit, \diamondsuit, \clubsuit\}$, $A_{14} = \{\spadesuit, \diamondsuit, \clubsuit\}$,
- $A_{15} = \Omega = \{\heartsuit, \spadesuit, \diamondsuit, \clubsuit\}$ - jev jistý.

Poznamejme ještě, že např. jev $A_5 = \{\heartsuit, \spadesuit\}$ značí tažení srdcí nebo pik. Obdobně $A_{12} = \{\heartsuit, \spadesuit, \clubsuit\}$ je tažení srdce, pik nebo křížů, což jde popsat také jako netažení kár.

Poznámky

Náhodný pokus

pokus, jehož výsledek nelze předem jednoznačně určit, přestože podmínky, za nichž pokus probíhá, jsou pevně dány,

Základní prostor

základní prostor jevů Ω je množina všech možných výsledků náhodného pokusu,

Elementární jev

každý prvek $\omega \in \Omega$, tj. každý možný výsledek náhodného pokusu,

Náhodný jev

každá podmnožina $A \subset \Omega$

Jistý jev

nastane vždy, $I = \Omega$,

Nemožný jev

nenastane nikdy $\emptyset$

96 - Operace s náhodnými jevy

Zadání Uvažujme opět náhodný pokus vytažení 1 karty z „pokerového“ balíčku 52 karet. Zapište následující náhodné jevy

A - vytažení srdcové karty,

D - vytažení karty menší než 5,

G - sjednocení jevů C a D ,

B - vytažení červené karty,

E - průnik jevů A a C ,

H - rozdíl jevů A a D ,

C - vytažení dámy,

F - sjednocení jevů A a C ,

CH - jev opačný k jevu B .

Řešení

Teorie: 16,17 Příklady: 178,179

Začneme jednoduchým výpisem jevů A, B, C a D

$$A = \{2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 5^\heartsuit, 6^\heartsuit, 7^\heartsuit, 8^\heartsuit, 9^\heartsuit, 10^\heartsuit, J^\heartsuit, Q^\heartsuit, K^\heartsuit, A^\heartsuit\},$$

$$B = \{2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 5^\heartsuit, 6^\heartsuit, 7^\heartsuit, 8^\heartsuit, 9^\heartsuit, 10^\heartsuit, J^\heartsuit, Q^\heartsuit, K^\heartsuit, A^\heartsuit, 2^\diamondsuit, 3^\diamondsuit, 4^\diamondsuit, 5^\diamondsuit, 6^\diamondsuit, 7^\diamondsuit, 8^\diamondsuit, 9^\diamondsuit, 10^\diamondsuit, J^\diamondsuit, Q^\diamondsuit, K^\diamondsuit, A^\diamondsuit\},$$

$$C = \{Q^\heartsuit, Q^\spadesuit, Q^\diamondsuit, Q^\clubsuit\},$$

$$D = \{2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 2^\spadesuit, 3^\spadesuit, 4^\spadesuit, 2^\diamondsuit, 3^\diamondsuit, 4^\diamondsuit, 2^\clubsuit, 3^\clubsuit, 4^\clubsuit\}.$$

Průnikem dvou jevů je jev, který nastává nastanou-li oba jevy společně. Pro jevy A a B jde tedy o tažení srdcové dámy,

$$E = A \cap C = \{Q^\heartsuit\}.$$

Sjednocení dvou náhodných jevů nastává, nastane-li alespoň jeden z nich. Pro jevy A a C to znamená, že je tažena srdcová karta nebo dáma,

$$F = A \cup C = \{2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 5^\heartsuit, 6^\heartsuit, 7^\heartsuit, 8^\heartsuit, 9^\heartsuit, 10^\heartsuit, J^\heartsuit, Q^\heartsuit, K^\heartsuit, A^\heartsuit, Q^\spadesuit, Q^\diamondsuit, Q^\clubsuit\}.$$

$$G = C \cup D = \{2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 2^\spadesuit, 3^\spadesuit, 4^\spadesuit, 2^\diamondsuit, 3^\diamondsuit, 4^\diamondsuit, 2^\clubsuit, 3^\clubsuit, 4^\clubsuit, Q^\heartsuit, Q^\spadesuit, Q^\diamondsuit, Q^\clubsuit\}.$$

Rozdílem dvou náhodných jevů, rozumíme jev, při němž první jev nastává, zatímco druhý nikoliv. Pro jevy A a D jde o tažení „srdcí“ větších než 4, tj.

$$H = A - D = \{5^\heartsuit, 6^\heartsuit, 7^\heartsuit, 8^\heartsuit, 9^\heartsuit, 10^\heartsuit, J^\heartsuit, Q^\heartsuit, K^\heartsuit, A^\heartsuit\}.$$

Jev opačný k jevu B nastane právě tehdy, když jev B nenastane. Zde jde o vytažení černé karty, jev značíme $\bar{B}$,

$$CH = \bar{B} = \{2^\spadesuit, 3^\spadesuit, 4^\spadesuit, 5^\spadesuit, 6^\spadesuit, 7^\spadesuit, 8^\spadesuit, 9^\spadesuit, 10^\spadesuit, J^\spadesuit, Q^\spadesuit, K^\spadesuit, A^\spadesuit, 2^\clubsuit, 3^\clubsuit, 4^\clubsuit, 5^\clubsuit, 6^\clubsuit, 7^\clubsuit, 8^\clubsuit, 9^\clubsuit, 10^\clubsuit, J^\clubsuit, Q^\clubsuit, K^\clubsuit, A^\clubsuit\}.$$

Poznámky

Sjednocení (součet) jevů A, B

$$A \cup B =$$

$$\{\omega \in \Omega : (\omega \in A) \vee (\omega \in B)\}$$

Průnik (součin) jevů A, B

$$A \cap B =$$

$$\{\omega \in \Omega : (\omega \in A) \wedge (\omega \in B)\}$$

Rozdíl jevů A, B

$$A - B =$$

$$\{\omega \in \Omega : (\omega \in A) \wedge (\omega \notin B)\}$$

Jev opačný k A

$$\bar{A} = \Omega - A =$$

$$\{\omega \in \Omega : \omega \notin A\}$$

97 - Relace mezi náhodnými jevy

Zadání Uvažujme náhodné jevy z příkladu 96. Vyberte z nich

- a) všechny podjevy jevu B ,
 b) všechny jevy, které si jsou rovny,
 c) všechny jevy neslučitelné s E ,
 d) úplný systém neslučitelných jevů.

Řešení

Teorie: 16,17 Příklady: 178,179

a) Podjevem jevu B jsou jevy, z jejichž nastoupení plyne také nastoupení jevu B . Jevem B bylo tažení „červené“ karty. Vytažení srdcové karty, jev A , samozřejmě znamená také vytažení karty červené. Proto je jev A podjevem jevu B , což značíme $A \subset B$. Podobně pro jevy E a H . Můžeme tedy zapsat

$$A \subset B, E \subset B \text{ a } H \subset B \text{ a navíc platí také } H \subset A \subset B \text{ a } E \subset A \subset B,$$

b) Dva jevy, X a Y , jsou si rovny, pokud z nastoupení jevu X plyne nastoupení Y a zároveň z nastoupení Y plyne jev X . Taková dvojice mezi jevy z daného příkladu není.

c) Jevy nazýváme neslučitelnými nemohou-li nastat současně. Jevem E je vytáhnutí $Q^\heartsuit$. Protože jde o červenou kartu, vylučuje její vytažení nastoupení jevu CH , tj. tažení černé. Obdobně protože $Q^\heartsuit$ je zjevně vyšší než 5, je nemožné, aby jev E nastal současně s jevem D . Jevy neslučitelné s E jsou D a CH .

d) Jevy tvoří systém neslučitelných jevů, je-li každá dvojice z nich vytvořená neslučitelná. Platí-li navíc, že sjednocením všech jevů tohoto systému je základní prostor Ω , pak systém nazýváme úplným systémem neslučitelných jevů. Takovým systémem je dvojice jevů B a CH . Nelze totiž vytáhnout zároveň černou i červenou, tj. $B \cap CH = \emptyset$, jevy jsou neslučitelné. A navíc $B \cup CH = \Omega$.

$$\{B, CH\} = \left\{ \left\{ 2^\heartsuit, 3^\heartsuit, 4^\heartsuit, 5^\heartsuit, 6^\heartsuit, 7^\heartsuit, 8^\heartsuit, 9^\heartsuit, 10^\heartsuit, J^\heartsuit, Q^\heartsuit, K^\heartsuit, A^\heartsuit, \right. \right. \\ \left. \left. 2^\diamondsuit, 3^\diamondsuit, 4^\diamondsuit, 5^\diamondsuit, 6^\diamondsuit, 7^\diamondsuit, 8^\diamondsuit, 9^\diamondsuit, 10^\diamondsuit, J^\diamondsuit, Q^\diamondsuit, K^\diamondsuit, A^\diamondsuit \right\}, \right. \\ \left. \left\{ 2^\spadesuit, 3^\spadesuit, 4^\spadesuit, 5^\spadesuit, 6^\spadesuit, 7^\spadesuit, 8^\spadesuit, 9^\spadesuit, 10^\spadesuit, J^\spadesuit, Q^\spadesuit, K^\spadesuit, A^\spadesuit, \right. \right. \\ \left. \left. 2^\clubsuit, 3^\clubsuit, 4^\clubsuit, 5^\clubsuit, 6^\clubsuit, 7^\clubsuit, 8^\clubsuit, 9^\clubsuit, 10^\clubsuit, J^\clubsuit, Q^\clubsuit, K^\clubsuit, A^\clubsuit \right\} \right\}.$$

Poznámky

Jev A je podjevem B

$$A \subset B \Leftrightarrow \{\forall \omega \in \Omega : (\omega \in A) \Rightarrow (\omega \in B)\}$$

Jevy A, B se rovnají

$$A = B \Leftrightarrow \{\forall \omega \in \Omega : (\omega \in A) \Leftrightarrow (\omega \in B)\}$$

Jevy A, B jsou neslučitelné

$$A \cap B = \emptyset$$

Systém neslučitelných jevů $A_1, A_2, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

Úplný systém neslučitelných jevů $A_1, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

$$A_1 \cup A_2 \cup \dots \cup A_n = \Omega$$

98 - Klasická pravděpodobnost

Zadání Je opravdu *poker* cennější než *full house*?

Řešení

Teorie: 18 Příklady: 180

Pro zpřehlednění popisu uvažujme klasickou pokerovou hru, při níž je hráči rozdáno pět karet a určíme pravděpodobnosti, že v těchto pěti kartách je *poker* resp. *full house*.

Nejprve určíme kolik různých petic může hráč obdržet. Protože jde o výběr bez možnosti opakování, u kterého nezáleží na pořadí, lze počet takových petic určit jako kombinace bez opakování (viz. př. 84). To znamená

$$n = \binom{52}{5} = \frac{52!}{(52-5)!5!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2\,598\,960.$$

Počtu možných rozdání *pokeru* zjistíme pomocí jednoduché úvahy a kombinatorického pravidla součinu (viz. př. 10). *Poker* totiž mohou být např. čtyři čtverky, $4♥, 4♠, 4♦, 4♣$ nebo (nejlépe) čtyři esa $A♥, A♠, A♦, A♣$, odsud vidíme, že je 13 různých *pokerů*. Krom čtyř karet *pokeru* má hráč v ruce pátou libovolnou kartu z balíčku, v němž zbývá 48 karet. Takto je počet *pokerů*

$$n_{poker} = 13 \cdot 48 = 624.$$

V případě *full house*, který je složen z trojice a dvojice, postupuje obdobně. Nejprve si uvědomme, že je 13 možností pro trojici. Pro dvojici jich zbývá 12. Dále je však třeba brát zřetel na to, že trojice je tvořena pouze třemi trojkami, např. $3♥, 3♠, 3♦$, nebo $3♥, 3♠, 3♣$. Stejně tak je dvojice složena pouze ze dvou karet stejné hodnoty, $2♥, 2♣$, nebo $2♥, 2♠$. Počet trojic, resp. dvojic, ze čtyř prvků určíme kombinačním číslem. Zapišeme-li vše kombinačními čísly, dostáváme

$$n_{fullhouse} = \binom{13}{1} \cdot \binom{4}{3} \cdot \binom{12}{1} \cdot \binom{4}{2} = 13 \cdot 4 \cdot 12 \cdot 6 = 3\,744.$$

Nyní lze spočítat pravděpodobnost rozdání *pokeru*, resp. *full house*. Samozřejmě přitom předpokládáme, že žádné z možných rozdání není nikterak zvýhodněno a že tedy mají stejnou šanci na nastoupení. Pak pravděpodobnost nastoupení požadovaného jevu spočteme jako podíl příznivých výsledků k počtu všech možných výsledků. Takto dostáváme

$$p_{poker} = \frac{n_{poker}}{n} = \frac{624}{2\,598\,960} \doteq 0,000240, \quad p_{fullhouse} = \frac{n_{fullhouse}}{n} = \frac{3\,744}{2\,598\,960} \doteq 0,001441.$$

Přestože se může zdát, že by *full house* měl mít vyšší hodnotu, protože oproti *pokeru* využívá všech pěti karet v hráčově ruce, není tomu tak.

Poznámky

Klasická pravděpodobnost

Nechť Ω je základní prostor tvořený n různými, vzájemně se vylučujícími elementárními jevy, které jsou stejně možné. Pak je pravděpodobnost nastoupení jevu A , který tvoří $m \leq n$ z těchto elementárních jevů, rovna podílu

$$P(A) = \frac{m}{n}.$$

Zjednodušeně lze zapsat

$$P(A) = \frac{\text{počet příznivých jevů}}{\text{počet všech jevů}}.$$

Pro každý jev $A \subseteq \Omega$ platí

$$0 \leq P(A) \leq 1.$$

99 - Pravděpodobnost a operace s náhodnými jevy

Zadání Opět uvažujme náhodný pokus tažení 1 karty z „pokerového“ balíčku 52 karet a s ním spojené náhodné jevy z příkladu ze strany 96. Určeme jejich pravděpodobnost.

A - vytažení srdcové karty,

D - vytažení karty menší než 5,

G - sjednocení jevů C a D ,

B - vytažení červené karty,

E - průnik jevů A a C ,

H - rozdíl jevů A a D ,

C - vytažení dámy,

F - sjednocení jevů A a C ,

CH - jev opačný k jevu B .

Řešení

Teorie: 16,17 Příklady: 178,179

V řešení příkladu 96 jsou všechny jevy vypsány, takže můžeme jednoduše určit počty příznivých jevů. Počet všech možných jevů je 52, což je počet všech karet.

Jev A - vytažení srdcové karty - může nastat 13 způsoby. Je zde tedy 13 příznivých jevů a pravděpodobnost jevu A určíme jako podíl $\frac{13}{52} = \frac{1}{4}$. Stejně tak, jako poměr počtu příznivých jevů ku všem jevům, spočteme i pravděpodobnosti dalších jevů.

$$P(A) = \frac{13}{52} = \frac{1}{4},$$

$$P(D) = \frac{12}{52} = \frac{3}{13},$$

$$P(G) = P(C \cup D) = \frac{16}{52} = \frac{4}{13},$$

$$P(B) = \frac{26}{52} = \frac{1}{2},$$

$$P(E) = P(A \cap C) = \frac{1}{52},$$

$$P(H) = P(A - D) = \frac{10}{52} = \frac{5}{26},$$

$$P(C) = \frac{4}{52} = \frac{1}{13},$$

$$P(F) = P(A \cup C) = \frac{16}{52} = \frac{4}{13},$$

$$P(\bar{B}) = \frac{26}{52} = \frac{1}{2}.$$

Všimněme si pozorně jevů F a G a jejich pravděpodobností. V obou případech jde o sjednocení dvou jevů. Ovšem zatímco pro jev G je $P(G) = \frac{4}{13} = \frac{1}{3} + \frac{3}{13} = P(C) + P(D)$. V případě jevu F podobná rovnost neplatí, $P(A) + P(C) = \frac{17}{52} > \frac{16}{52} = P(F)$. To je dáno tím, že průnik $A \cap C = \{\text{Q♥}\}$ je neprázdný a tato karta je tedy v příznivých jevech započtena dvakrát. Proto je třeba dbát při výpočtu pozornosti a nenechat se zmást tím, že pro sjednocení jevů používáme někdy také výraz „součet jevů“. Správný vzorec je

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Dále je zjevně $B \cap \bar{B} = \emptyset$ jev nemožný a jeho pravděpodobnost nulová. Naopak $B \cup \bar{B} = \Omega$ je jev jistý a $P(\Omega) = 1$. Odsud a z předchozího odvodíme, že $P(B) + P(\bar{B}) = 1$ neboli

$$P(\bar{B}) = 1 - P(B).$$

Poznámky

Vlastnosti pravděpodobnosti

jev nemožný

$$P(\emptyset) = 0,$$

jev jistý

$$P(\Omega) = 1,$$

sjednocení jevů

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

jev opačný

$$P(\bar{B}) = 1 - P(B).$$

100 - Pravděpodobnost a relace mezi náhodnými jevy

Zadání Uvažujme náhodné jevy z příkladu 96. Vyberte z nich všechny podjevy jevu B a určete jejich pravděpodobnosti.

Řešení

Teorie: 16,17 Příklady: 178,179

Podjevy jevu B byly nalezeny a popsány už v příkladu 97. Jde o jevy A, E, H , přičemž konkrétně platí $E \subset A \subset B$ a $H \subset A \subset B$. Pravděpodobnosti jednotlivých jevů jsou určeny již v příkladě 99. V tomto příkladě se zaměříme spíše na popis jejich vztahu. Platí totiž

$$\frac{1}{52} \leq \frac{13}{52} \leq \frac{1}{2} \text{ a tedy } P(E) \leq P(A) \leq P(B),$$
$$\frac{10}{52} \leq \frac{13}{52} \leq \frac{1}{2} \text{ a tedy } P(H) \leq P(A) \leq P(B).$$

Odsud si odvodíme, že pravděpodobnost jevu A , který je podjevem jevu B , je menší než pravděpodobnost jevu B . Přesněji

$$A \subseteq B \Rightarrow P(A) \leq P(B).$$

Dále si uvědomíme, že jev nemožný je podjevem všech uvažovaných jevů. A obdobně všechny uvažované jevy jsou podjevem jevu jistého. Můžeme si z $\emptyset \subseteq A \subseteq \Omega$ dále odvodit, že

$$0 = P(\emptyset) \leq P(A) \leq P(\Omega) = 1.$$

Poznámky

Vlastnosti pravděpodobnosti

pravděpodobnost podjevu

$$A \subseteq B \Rightarrow P(A) \leq P(B)$$

a

$$0 = P(\emptyset) \leq P(A) \leq P(\Omega) = 1.$$

101 - Geometrická pravděpodobnost

Zadání Dvoumetrová tyč byla rozlomena na tři díly. Určete pravděpodobnost, že každý z dílů bude delší než půl metru.

Řešení

Označme si délky jednotlivých dílů x, y, z a zapišme známé a požadované vztahy. Zjevně platí $x + y + z = 2$, odsud vyjádříme $z = 2 - x - y$. Pro délky dílů x, y i z platí, že jsou větší než 0 a menší než 2. Dále odvozujeme $0 < z = 2 - x - y$ a tedy, že $x + y < 2$.

Nyní si podobně shrňme požadavky. Všechny díly mají být delší než 0,5m, tj. $0,5 < x$, $0,5 < y$ a $0,5 < z = 2 - x - y$. Z posledního vztahu plyne $x + y < 1,5$.

S dalším postupem pomůže jednoduchý náčrtek, do kterého si všechny podmínky přehledně zaznamenejme.

První sada podmínek ohraničuje světlý trojúhelník, který reprezentuje všechny možné výsledky lámání tyče. Například červený bod má souřadnice $x = 1,4$ a $y = 1,6$, což samozřejmě nemůže odpovídat délkám dílů dvoumetrové tyče. Naproti modrý bod, o souřadnicích $x = 1,3$, $y = 0,6$, z nichž dopočteme $z = 0,1$, jeden z možných výsledků reprezentuje. Nejde však o výsledek příznivý.

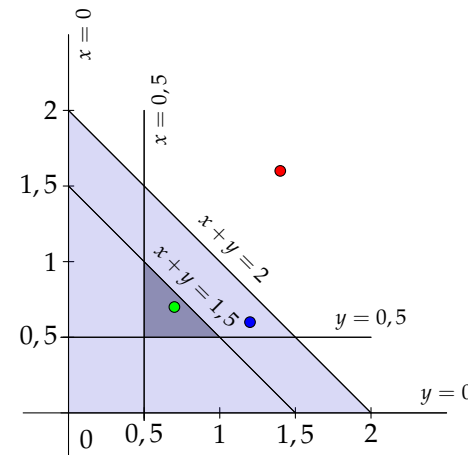
Oblast, která odpovídá požadovaným výsledkům, je zvýrazněna tmavě. Zelený bod v ní obsažený odpovídá lomu $x = 0,7$, $y = 0,7$ a $z = 0,6$, a tedy splňuje podmínky zadání.

Pravděpodobnost, že náhodný lom tyče na tři díly, splní podmínku minimálně půlmetrové délky každého z nich, určíme jako podíl velikostí ploch obou trojúhelníků. Přičemž plocha S světlého, většího z nich, reprezentuje všechny možné jevy. Plocha S_p tmavého trojúhelníku reprezentuje všechny jevy příznivé.

Není těžké spočítat, že $S = 2$ a $S_p = \frac{1}{8}$ a odtud určit pravděpodobnost žádaného jevu,

$$p = \frac{S_p}{S} = \frac{\frac{1}{8}}{2} = \frac{1}{16}.$$

Teorie: 18 Příklady: 181



Poznámky

Geometrická pravděpodobnost
Pokud lze základní prostor Ω modelovat jako geometrický útvar $\mathcal{O}$ a náhodný jev A jako útvar $\mathcal{A}$, přičemž platí $\mathcal{A} \subseteq \mathcal{O}$, pak pravděpodobnost nastoupení jevu A lze spočítat jako podíl

$$P(A) = \frac{\mu(\mathcal{A})}{\mu(\mathcal{O})},$$

kde $\mu(\mathcal{A})$ resp. $\mu(\mathcal{O})$ značí míru útvaru $\mathcal{A}$, resp. $\mathcal{O}$.

Obvykle je $\mathcal{O}$ konečná, uzavřená oblast na přímce, v rovině či v prostoru a $\mathcal{A}$ její uzavřená podmnožina, a jejich mírou rozumíme délku, plochu či objem.

102 - Podmíněná pravděpodobnost

Zadání Sledováním své produkce firma zjistila, že ze 100 000 výrobků jich 87 921 pracovalo bez poruchy dva roky a 68 383 nemělo poruchu tři roky. Jaká je pravděpodobnost, že výrobek, který neměl v prvních dvou letech poruchu, bude bez závad i po třech letech?

Řešení **Teorie: 20 Příklady: 182,183**

Nejprve popíšeme a označíme náhodné jevy. Bezporuchový provoz v prvních dvou letech označme A a jevem B je provoz bez poruch po dobu tří let. Protože nemáme informace o všech výrobcích, které firma vyprodukovala, ale pouze o dostatečně velkém vzorku, využijeme statistické pravděpodobnosti.

Dle definice statistické pravděpodobnosti totiž můžeme poměr $\frac{87\,921}{100\,000} = 87,921$ považovat za pravděpodobnost bezporuchového provozu po dobu dvou let. I v této definici je pravděpodobnost zavedena jako poměr příznivých jevů ku všem možným, přičemž všemi jevy zde myslíme pouze jevy z daného vzorku. Obdobně je $\frac{68\,383}{100\,000} = 68,383$ pravděpodobností tříleté práce bez závad.

Stroj, který pracoval bez závad již dva roky, je jeden z 87 921 takových exemplářů. Tento počet nyní představuje počet všech možných jevů a 68 383 představuje počet jevů příznivých. Pravděpodobnost nastoupení jevu B za předpokladu nastoupení jevu A určíme jako poměr těchto „nových“ příznivých a všech možných. Tj.

$$P(B|A) = \frac{68,383}{87,921} = \frac{7}{9} = 0,\bar{7}.$$

S využitím vzorce pro výpočet podmíněné pravděpodobnosti lze výsledek určit následovně. Nejprve určíme, kdy oba jevy nastávají společně, tj. pravděpodobnost jevu $A \cap B$. Ve vyšetřovaném případě je průnikem jevů A, B jev B , neboť stroj, který neměl poruchu za tři roky provozu, ji nemohl mít ani v prvních dvou letech. Je B je podjevem jevu A . Dostáváme:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{68,383}{87,921} = \frac{7}{9} = 0,\bar{7}.$$

Poznámky

Podmíněná pravděpodobnost

Pravděpodobnost nastoupení jevu B za předpokladu, že nastoupil jev A .

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Nezávislé jevy

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$$

$$P(A \cap B) = P(A) P(B),$$

přičemž je $P(A) > 0$ a $P(B) > 0$.

Závislé jevy

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

103 - Podmíněná pravděpodobnost a nezávislé jevy

Zadání Určete pravděpodobnosti jevů $A - H$ z příkladu 96, pokud víte, že nastal jev B .

Řešení **Teorie: 20 Příklady: 182, 183**

Využijeme vzorce pro podmíněnou pravděpodobnost, přičemž podmínkou je nastoupení jevu B - vytažení červené karty. Takto podmíněná pravděpodobnost pro jev X je

$$P(X|B) = \frac{P(X \cap B)}{P(B)}.$$

Pravděpodobnost jevu B známe, je rovna $P(B) = \frac{1}{2}$. Je zjevné, že pravděpodobnost $P(B|B) = 1$, nastoupení jevu B , víme-li, že jev B nastoupil, je jev jistý. Určíme pravděpodobnosti $P(X \cap B)$, pro $X \neq B$. S pomocí zápisu v příkladě 96 dostáváme

$$\begin{aligned} P(A \cap B) &= \frac{1}{4}, & P(D \cap B) &= \frac{3}{26}, & P(F \cap B) &= \frac{7}{26}, & P(H \cap B) &= \frac{5}{26}, \\ P(C \cap B) &= \frac{1}{26}, & P(E \cap B) &= \frac{1}{52}, & P(G \cap B) &= \frac{2}{13}, & P(\bar{B} \cap B) &= 0. \end{aligned}$$

Dosazením do vzorce zjistíme podmíněné pravděpodobnosti

$$\begin{aligned} P(A|B) &= \frac{1}{2}, & P(D|B) &= \frac{3}{13} = P(D), & P(F|B) &= \frac{7}{13}, & P(H|B) &= \frac{5}{13}, \\ P(C|B) &= \frac{1}{13} = P(C), & P(E|B) &= \frac{1}{26}, & P(G|B) &= \frac{4}{13} = P(G), & P(\bar{B}|B) &= 0. \end{aligned}$$

Pravděpodobnosti jevů C, D a G se splněním podmínky nastoupení jevu B nezměnila. Je přirozené nazvat takové jevy nezávislé. Spočítáme-li např. $P(B|C) = \frac{\frac{1}{26}}{\frac{1}{13}} = \frac{1}{2} = P(B)$. Ani nastoupení jevu C neovlivňuje pravděpodobnost jevu B . Pro nezávislé jevy tedy platí

$$P(C|B) = \frac{P(C \cap B)}{P(B)} = P(C) \quad \text{a} \quad P(B|C) = \frac{P(B \cap C)}{P(C)} = P(B).$$

Odtud je patrné, že pro nezávislé jevy platí

$$P(C \cap B) = P(B)P(C).$$

Poznámky

Podmíněná pravděpodobnost

Pravděpodobnost nastoupení jevů B za předpokladu, že nastoupil jev A .

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Nezávislé jevy

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$$

$$P(A \cap B) = P(A)P(B),$$

přičemž je $P(A) > 0$ a $P(B) > 0$.

Závislé jevy

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

104 - Úplná pravděpodobnost

Zadání V krabičce je 53 sirek, z nichž tři jsou vypálené. Jaká je pravděpodobnost, že v pořadí druhá vytažená zápalka bude vypálená?

Řešení

Teorie: 21 Příklady: 184

Pravděpodobnost, že druhá sirka je vypálená, samozřejmě závisí na stavu zápalky vybrané jako první. Označme si tedy D_1 jev výběru dobré sirky v prvním tahu. V tomto případě zůstane v krabičce 52 sirek, v nichž jsou stále tři vypálené. V případě tažení vypálené, jev označme V_1 , zůstane po prvním tahu v krabičce také 52 zápalek, ale ovšem jen dvě vypálené.

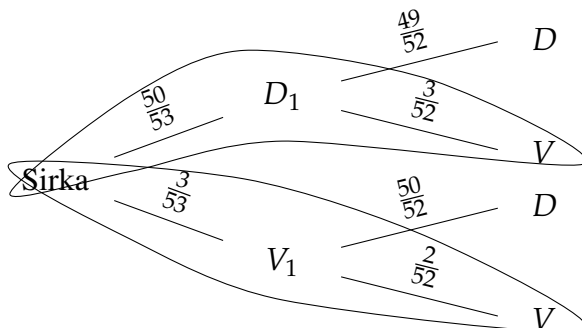
Jevy D_1 a V_1 jsou vzájemně neslučitelné, tj. $D_1 \cap V_1 = \emptyset$. Protože je navíc $D_1 \cup V_1 = \Omega$, tvoří jevy D_1, V_1 úplný systém neslučitelných jevů. Toho využijeme k určení pravděpodobnosti jevu V - tažení vypálené sirky „ve druhém kole.“ Jev V může nastat buď po nastoupení jevu D_1 , a to s pravděpodobností $P(V|D_1)$ nebo po jevu V_1 s pravděpodobností $P(V|V_1)$. Protože pravděpodobnost neslučitelných jevů je rovna součtu jejich pravděpodobností, můžeme spočítat $P(V)$ jako

$$P(V) = P(V|D_1) P(D_1) + P(V|V_1) P(V_1).$$

Pravděpodobnost jevu D_1 je $P(D_1) = \frac{50}{53}$ a jevu V_1 je $P(V_1) = \frac{3}{53}$. Po vytažení dobré sirky jsou mezi 52 zápalkami stále tři špatné, a tedy $P(V|D_1) = \frac{3}{52}$. Obdobně je $P(V|V_1) = \frac{2}{52}$. Odtud dostáváme

$$P(V) = P(V|D_1) P(D_1) + P(V|V_1) P(V_1) = \frac{3}{52} \cdot \frac{50}{53} + \frac{2}{52} \cdot \frac{3}{53} = \frac{3}{53} \doteq 0,0566.$$

Při řešení úplné pravděpodobnosti je výhodné využít rozhodovacího stromu. V našem případě jsou dvě různé cesty, jak vytáhnout v druhém tahu vypálenou sirku.



Poznámky

Úplná pravděpodobnost

neslučitelné jevy

$$A_1 \cap A_2 = \emptyset$$

úplný systém neslučitelných jevů

$$A_1 \cap A_2 = \emptyset$$

$$A_1 \cup A_2 = \Omega$$

úplná pravděpodobnost

$$P(B) = P(B|A_1) P(A_1) + P(B|A_2) P(A_2)$$

105 - Úplná pravděpodobnost

Zadání V krabici je 53 sirek, z nichž tři jsou vypálené. Jaká je pravděpodobnost, že v pořadí třetí vytažená zápalka bude vypálená?

Řešení **Teorie: 21 Příklady: 184**

Jedná se drobnou modifikaci příkladu 104. Opět tedy budeme počítat úplnou pravděpodobnost jevu V , kterým je tentokrát tažení vypálené zápalky ve třetím tahu.

Rozeberme si jevy, které třetímu tahu mohly předcházet. Jev A_1 vytažení dvou dobrých sirek, jev A_2 vytažení dobré a špatné, a konečně jev A_3 vytažení dvou vypálených. Tyto jevy tvoří úplný systém neslučitelných jevů.

Pravděpodobnost jevu A_1 spočteme jako $\frac{50}{53} \cdot \frac{49}{52}$, kde $\frac{50}{53}$ odpovídá tahu dobré v první kole a $\frac{49}{52}$ je výběr dobré v kole druhém, v němž již krabice obsahuje pouze 52 sirek, z nichž je 49 v pořádku. Takto máme $P(A_1) = \frac{1225}{1378}$. Obdobně je $P(A_3) = \frac{3}{53} \cdot \frac{2}{52} = \frac{3}{1378}$. Pro určení pravděpodobnosti jevu A_2 využijeme toho, že je jevem opačným k jevu $A_1 \cup A_3$, a proto je $P(A_2) = 1 - (P(A_1) + P(A_3)) = \frac{150}{1378}$.

Nyní si určíme podmíněné pravděpodobnosti $P(V|A_1)$, $P(V|A_2)$ a $P(V|A_3)$. V případě jevu A_1 zůstává v krabici 51 sirek, z nichž jsou tři vypálené. V případě A_2 jsou mezi 51 sirkami v krabici dvě vypálené. A v případě A_3 zůstala v krabici jediná vypálená sirka. Hledané pravděpodobnosti tedy jsou $P(V|A_1) = \frac{3}{51}$, $P(V|A_2) = \frac{2}{51}$ a $P(V|A_3) = \frac{1}{51}$.

Dosazením těchto hodnot do vzorce dostáváme (úplnou) pravděpodobnost $P(V)$

$$\begin{aligned} P(V) &= P(V|A_1) P(A_1) + P(V|A_2) P(A_2) + P(V|A_3) P(A_3) \\ &= \frac{3}{51} \cdot \frac{1225}{1378} + \frac{2}{51} \cdot \frac{150}{1378} + \frac{1}{51} \cdot \frac{3}{1378} = \frac{3}{53} \doteq 0,0566. \end{aligned}$$

Poznámky

Úplná pravděpodobnost

úplný systém neslučitelných jevů $A_1, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

$$\bigcup_{i=1}^n A_i = \Omega$$

úplná pravděpodobnost

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

106 - Bayesova věta

Zadání Podle odhadů tvoří nevyžádaná pošta (SPAM) až dvě třetiny celkového objemu emailové komunikace. Spam filtr je schopen zachytit 90 % nevyžádaných zpráv. Test však jako spam označí i 12 % zpráv, které spamem nejsou. Určete pravděpodobnost, že

- filtr zprávu označí jako spam,
- zpráva jako spam označená, je opravdu spam.

Řešení **Teorie: 21 Příklady: 185,186**

V případě **a)** se jedná o úplnou pravděpodobnost. Označme T označení zprávy za spam a S skutečnost, že zpráva spamem opravdu je. Jev opačný k S , tj. jev $\bar{S}$, označuje zprávu, která spamem není. Z popisu situace uvedené v zadání známe podmíněné pravděpodobnosti $P(T|S) = 0,9$, $P(T|\bar{S}) = 0,12$. Dále dlouhodobé pozorování vede ke statistické definici pravděpodobností $P(S) = \frac{2}{3}$ a $P(\bar{S}) = \frac{1}{3}$. Úplná pravděpodobnost označení zprávy za spam tedy je

$$P(T) = P(T|S)P(S) + P(T|\bar{S})P(\bar{S}) = 0,9 \cdot \frac{2}{3} + 0,12 \cdot \frac{1}{3} = \frac{1,8}{3} + \frac{0,12}{3} = \frac{1,92}{3} = 0,64.$$

Spam filtr tedy jako spam označí 64 % ze všech emailů, které do schránky přijdou.

V případě **b)** máme určit $P(S|T)$. K tomu využijeme vzorců pro výpočet podmíněné pravděpodobnosti (viz. př. 102.) Ze vztahu $P(T|S) = \frac{P(S \cap T)}{P(S)}$ si vyjádříme $P(S \cap T)$ a vyjádření dosadíme do $P(S|T) = \frac{P(S \cap T)}{P(T)}$. Tak obdržíme

$$P(S|T) = \frac{P(T|S)P(S)}{P(T)} = \frac{P(T|S)P(S)}{P(T|S)P(S) + P(T|\bar{S})P(\bar{S})}.$$

V tomto jsme také využili výpočet $P(T)$ z části **a)**. Tento vztah se nazývá Bayesův vzorec. Zbývá do ní pouze dosadit číselné hodnoty, i při tom využíváme již známých výsledků.

$$P(S|T) = \frac{P(T|S)P(S)}{P(T)} = \frac{\frac{1,8}{3}}{\frac{1,92}{3}} = \frac{1,8}{1,92} = \frac{15}{16} = 0,9375.$$

Je-li zpráva označena jako spam, pak s pravděpodobností 93,75 % spamem opravdu je.

Poznámky

Bayesův vzorec

neslučitelné jevy

$$A_1 \cap A_2 = \emptyset$$

úplný systém neslučitelných jevů

$$A_1 \cap A_2 = \emptyset$$

$$A_1 \cup A_2 = \Omega$$

úplná pravděpodobnost

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2)$$

Bayesův vzorec

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

107 - Bayesova věta

Zadání Materiál na stavbu dodávají tři hlavní prodejci D_1 , D_2 a D_3 s podíly 40 %, 30 % a 25 %. Zbývajících 5 % je nakupováno průběžně dle aktuální potřeby u různých dodavatelů D_r . Spolehlivost dodavatelů D_1 , D_2 a D_3 je 95 %, 97 % a 95 %. Při nákupu u jiných dodavatelů dochází k omylu v 10 % objednávek. Určete pravděpodobnost, že

- a) dodávka je chybná,
- b) že chyba vznikla u dodavatele D_3 .

Řešení

Teorie: 21 Příklady: 185, 186

Postup bude úpravou postupu z příkladu 106. Nejprve si zavedeme označení jednotlivých jevů a jejich pravděpodobností. Podíl dodavatelů „na trhu“ představuje pravděpodobnost, že nákup byl proveden právě u nich, tj. $P(D_1) = 0,4$, $P(D_2) = 0,3$, $P(D_3) = 0,25$ a $P(D_r) = 0,05$. Spolehlivost dodavatele značí pravděpodobnost bezvadné dodávky. Protože je úkolem spočítat pravděpodobnost chybné dodávky, uvažujme jev CH_1 , který značí chybu vzniklou u dodavatele D_1 , jeho pravděpodobnost je $P(CH|D_1) = 0,05$. (Jde o jev opačný k bezchybnému dodání, tj. spolehlivosti dodavatele.) Obdobně $P(CH|D_2) = 0,03$, $P(CH|D_3) = 0,05$ a konečně $P(CH|D_r) = 0,1$.

Nyní již můžeme počítat. Nejprve v případě a) úplnou pravděpodobnost chybné dodávky $P(CH)$.

$$\begin{aligned} P(CH) &= P(CH|D_1) P(D_1) + P(CH|D_2) P(D_2) + P(CH|D_3) P(D_3) + P(CH|D_r) P(D_r) = \\ &= 0,05 \cdot 0,4 + 0,03 \cdot 0,3 + 0,05 \cdot 0,25 + 0,1 \cdot 0,05 = 0,0465. \end{aligned}$$

Pravděpodobnost chybné dodávky je 0,0465.

Případ b) popisuje situaci, kdy byla objednávka vyřízena chybně a máme určit $P(D_3|CH)$. Opět lze využít vzorců pro výpočet podmíněné pravděpodobnosti (viz. př. 102) a z nich si vyjádřit a odvodit Bayesův vzorec. Dosazením do něj získáme hledanou pravděpodobnost

$$P(D_3|CH) = \frac{P(CH|D_3) P(D_3)}{P(CH)} = \frac{0,05 \cdot 0,25}{0,0465} \doteq 0,269.$$

Dodejme ještě hodnoty $P(D_1|CH) \doteq 0,430$, $P(D_2|CH) \doteq 0,194$ a $P(D_r|CH) \doteq 0,107$, a že součet $P(D_1|CH) + P(D_2|CH) + P(D_3|CH) + P(D_r|CH) = 1$.

Poznámky

Bayesův vzorec

úplný systém neslučitelných jevů $A_1, \dots, A_n$

$$\begin{aligned} A_i \cap A_j &= \emptyset, \text{ pro } i \neq j \\ \bigcup_{i=1}^n A_i &= \Omega \end{aligned}$$

úplná pravděpodobnost

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

Bayesův vzorec

$$P(A_k|B) = \frac{P(B|A_k) P(A_k)}{\sum_{i=1}^n P(B|A_i) P(A_i)}$$

108 - Opakované pokusy - nezávislé

Zadání Na stavbě pracuje 7 mužů z firmy A a 4 muži firmy B . Každé ráno je jeden z mužů vybrán k testu na přítomnost alkoholu. Muži jsou vybíráni náhodně a mohou být testováni opakovaně. Jaká je pravděpodobnost, že během pracovního týdne jsou testováni právě dva muži firmy B .

Řešení **Teorie: 22 Příklady: 187**

Výběr pracovníka firmy A je náhodným jevem s pravděpodobností $P(A) = \frac{7}{7+4} = \frac{7}{11}$. Pravděpodobnost výběru muže firmy B je $P(B) = \frac{4}{11}$. Uvažme nejprve, že dva muži B jsou vybráni již v pondělí a úterý. To znamená, že ve zbývajících dny jsou zvoleni muži firmy A . Protože muži mohou být voleni opakovaně, neovlivňuje výsledek provedené volby výsledky následující, můžeme pravděpodobnost uvažovaného výběru - $BBAAA$ - spočítat jako součin

$$\frac{4}{11} \cdot \frac{4}{11} \cdot \frac{7}{11} \cdot \frac{7}{11} \cdot \frac{7}{11} = \left(\frac{4}{11}\right)^2 \cdot \left(\frac{7}{11}\right)^3.$$

Samozřejmě mohou být muži B vybráni na konci týdne $AAABB$ nebo ve středu a v pátek $AABAB$. Počet všech takových možností zjistíme jako počet permutací s opakováním, přesně to je $\frac{5!}{3!2!} = \binom{5}{2}$. Protože každá z těchto voleb může nastat se stejnou pravděpodobností, můžeme hledanou pravděpodobnost výběru právě 2 mužů, označme ji $P(V)$, vyjádřit jako

$$P(V) = \binom{5}{2} \cdot \left(\frac{4}{11}\right)^2 \cdot \left(\frac{7}{11}\right)^3 = \frac{10 \cdot 4^2 \cdot 7^3}{11^5} \doteq 0,34.$$

Pokud bychom chtěli výsledek zobecnit a zjistit pravděpodobnost výběru k mužů firmy B , pak získáme tvar

$$\binom{5}{k} \cdot \left(\frac{4}{11}\right)^k \cdot \left(\frac{7}{11}\right)^{5-k}.$$

Další zobecnění pro n nezávislých pokusů, v nichž jev A s pravděpodobností $P(A) = p$ nastává k -krát a jev B s $P(B) = 1 - p$ nastane $(n - k)$ -krát vede k vyjádření pravděpodobnosti

$$P(A_k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Poznámky

Pravděpodobnost opakovaných pokusů

Je-li pravděpodobnost nastoupení jevu A v pokusu $P(A) = p$, pak pravděpodobnost jevu A_k , že jev A nastane v sérii n opakovaných nezávislých pokusů právě k -krát je

$$P(A_k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

109 - Opakované pokusy - nezávislé

Zadání Směrná pracnost zdění je $0,98 \text{ hod/m}^2$ a spotřeba cihel 16 ks/m^2 . Pravděpodobnost, že zedník cihlu špatně uloží, je 5 %. Po jaké době bude pravděpodobnost jedné špatně uložené cihly větší než $P = 0,9$?

Řešení **Teorie: 22 Příklady: 187**

Odhlédněme nejprve od času a soustředíme se pouze na kusy cihel. A označme $P(i)$ pravděpodobnost toho, že i -tá použitá cihla je špatně uložena. Takto máme $P(1) = 0,05$, tj. první cihla je špatně uložena. Dále $P(2) = 0,95 \cdot 0,05$ odpovídá správnému použití první cihly a špatnému uložení druhé cihly. Odtud už můžeme odvodit pravděpodobnost špatného uložení i -té cihly, tedy $P(i) = 0,95^{i-1} \cdot 0,05$. Opět jsme využili toho, že jde o jevy nezávislé.

Dále pravděpodobnost toho, že špatně bude uložena první nebo druhá cihla, lze určit jako $P(1) + P(2) = 0,0975$. Hledáme proto n takové, aby $P(1) + P(2) + \dots + P(n) \geq 0,9$. Upravujeme tedy tento součet

$$0,05 + 0,95 \cdot 0,05 + 0,95^2 \cdot 0,05 + \dots + 0,95^n \cdot 0,05 = 0,05(1 + 0,95 + 0,95^2 + \dots + 0,95^n).$$

Protože součet v závorce předchozího vztahu představuje prvních n členů geometrické posloupnosti, použijeme vzorec pro její částečný součet, dosadíme a výsledek porovnáme s požadovanou hodnotou 0,9.

$$\begin{aligned} 0,05 \left(\frac{1 - 0,95^{n+1}}{1 - 0,95} \right) &= 0,9 \\ 1 - 0,95^{n+1} &= 0,9 \\ \ln 0,1 &= \ln 0,95^{n+1} \\ n + 1 &= \frac{\ln 0,1}{\ln 0,95} \\ n &\doteq 43,8 \end{aligned}$$

Po uložení 44 cihel, tedy již bude pravděpodobnost, toho, že jedna z nich je uložena špatně větší než 0,9. Z technického listu zdícího materiálu dodané výrobcem si dopočteme, že 44 cihel odpovídá $2,75 \text{ m}^2$. A to dle směrné pracnosti zdění odpovídá $2,75 \cdot 0,98 = 2,695$ hodinám práce, tj. 2 hodiny a 42 minut.

Poznámky

Pravděpodobnost opakovaných pokusů

Je-li pravděpodobnost nastoupení jevu A v pokusu $P(A) = p$, pak pravděpodobnost jevu A_k , že jev A nastane v sérii n opakovaných nezávislých pokusů právě k -krát je

$$P(A_k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

110 - Opakované pokusy - závislé

Zadání V dodávce 1000 cihel je 53, které neodpovídají normě. Jaká je pravděpodobnost, že ve 20 vybraných budou nejvýše tři vadné?

Řešení

Teorie: 22 Příklady: 188

Vybereme-li nyní jednu vadnou cihlu, zůstane z dodávky 999 cihel a v nich 52 špatných cihel. Pokud je prvním výběrem dobrá cihla, pak z dodávky zůstává 999 cihel a v nich 53 vadných. Z toho je jasné patrné, že nejde o nezávislé jevy a že tedy musíme použít jiný přístup než v příkladech ze stran 108 a 109. Určeme si nejprve například, kolika způsoby lze vybrat jednu vadnou cihlu a pravděpodobnost $P(A_1)$ takového výběru.

Protože nejde o pořadí v jakém výběr provedeme, budeme proto počítat kombinace. Nejprve $\binom{53}{1}$, tj. výběr jedné zmetkové cihly, dále $\binom{947}{19}$ představující výběr 19ti dobrých cihel. Celkově je tedy $\binom{53}{1} \cdot \binom{947}{19}$ možností, jak vybrat dvacítku cihel, v nichž je jedna vadná. Dále určíme $\binom{1000}{20}$, což je počet všech možností výběru dvaceti cihel.

Pravděpodobnost výběru, v němž je jedna vadná, jako obvykle spočteme podílem příznivých a všech možných jevů,

$$P(A_1) = \frac{\binom{53}{1} \cdot \binom{947}{19}}{\binom{1000}{20}} \doteq 0,380.$$

Pro zbývající jevy A_i , kde $i = 0, 2, 3$ značí počet vadných cihel ve výběru, máme

$$P(A_i) = \frac{\binom{53}{i} \cdot \binom{947}{20-i}}{\binom{1000}{20}}.$$

Dosazením do tohoto vztahu dostáváme pravděpodobnosti $P(A_0) \doteq 0,333$, $P(A_1) \doteq 0,380$, $P(A_2) \doteq 0,202$ a $P(A_3) \doteq 0,066$. Konečně pravděpodobnost jevu A , že ve výběru 20 cihel budou nejvýše tři vadné, určíme jako součet spočtených pravděpodobností

$$P(A) = P(A_0 \cup A_1 \cup A_2 \cup A_3) = P(A_0) + P(A_1) + P(A_2) + P(A_3) \doteq 0,981.$$

Poznámky

Pravděpodobnost závislých opakovaných pokusů

V souboru n prvků, má $k, 0 \leq k \leq n$ určitou vlastnost a $n - k$ tuto vlastnost nemá. Ze souboru vybíráme postupně j prvků, které nevracíme. Označme A_i jev, odpovídající tomu, že v tomto výběru má právě i prvků sledovanou vlastnost. Pravděpodobnost $P(A_i)$ vypočteme podle vzorce

$$P(A_i) = \frac{\binom{k}{i} \cdot \binom{n-k}{j-i}}{\binom{n}{j}}.$$

Řešené příklady – Náhodná veličina

112 - Rozdělení diskretní náhodné veličiny - část 1.

Zadání V divadelním foyer jsou dva automaty na nápoje. Pravděpodobnost, že automat po vhození mince přestane vydávat nápoje je u prvního z nich 7 % a u druhého 5 %. Náhodná veličina bude označovat počet automatů nevydávající nápoj. Určete $p(x)$, $F(x)$ náhodné veličiny X . Dále určete $E(x)$, $D(x)$, σ a modus.

Řešení

Teorie: 25,26 Příklady: 190,191,192,193

náhodná veličina X ... počet automatů nevydávající nápoj

Náhodná veličina X může nabývat tří hodnot 0, 1, 2, takže se jedná o diskretní náhodnou veličinu (DNV).

označíme jevy:

jev A_1 ... první automat nevydává nápoj, $P(A_1) = 0,07$

jev A_2 ... druhý automat nevydává nápoj, $P(A_2) = 0,05$

Určíme pravděpodobnostní funkci NVX:

$$P(X = 0) = P(\bar{A}_1 \cap \bar{A}_2) = P(\bar{A}_1) \cdot P(\bar{A}_2) = 0,93 \cdot 0,95 = 0,8835$$

$$P(X = 1) = P((\bar{A}_1 \cap A_2) \cup (A_1 \cap \bar{A}_2)) = P(\bar{A}_1) \cdot P(A_2) + P(A_1) \cdot P(\bar{A}_2) = 0,93 \cdot 0,05 + 0,07 \cdot 0,95 = 0,0465 + 0,0665 = 0,113$$

$$P(X = 2) = P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = 0,07 \cdot 0,05 = 0,0035$$

Zapíšeme do tabulky:

x	0	1	2
$p(x)$	0,8835	0,113	0,0035

Jedna z vlastností pravděpodobnostní funkce je, že $\sum_i p(x_i) = 1$, ověříme: $\sum_i p(x_i) = 0,8835 + 0,113 + 0,0035 = 1$

Určení distribuční funkce:

$$F(x) = \sum_{x_i < x} p(x_i)$$

$$\forall x \in (-\infty; 0) : F(x) = P(X < x) = 0$$

$$\forall x \in (0; 1) : F(x) = P(X < x) = P(X = 0) = 0,8835$$

$$\forall x \in (1; 2) : F(x) = P(X < x) = P(X = 0) + P(X = 1) = 0,8835 + 0,113 = 0,9965$$

$$\forall x \in (2; \infty) : F(x) = P(X < x) = P(X = 0) + P(X = 1) + P(X = 2) = 0,8835 + 0,113 + 0,0035 = 1$$

Poznámky

Diskretní náhodná veličina

Pravděpodobnostní funkce

$$p(x) = P(X = x)$$

Vlastnosti

- $p(x_i) \geq 0$

- $\sum_{i=1}^n p(x_i) = 1$

Distribuční funkce

$$F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$$

Vlastnosti

- $0 \leq F(x) \leq 1$

- $F(x)$ je neklesající funkce

- $\lim_{x \rightarrow -\infty} F(x) = 0$

- $\lim_{x \rightarrow \infty} F(x) = 1$

- $F(x)$ je zleva spojitá

113 - Rozdělení diskretní náhodné veličiny - část 2.

Zadání V divadelním foyer jsou dva automaty na nápoje. Pravděpodobnost, že automat po vhození mince přestane vydávat nápoje je u prvního z nich 7 % a u druhého 5 %. Náhodná veličina bude označovat počet automatů nevýdávající nápoj. Určete $p(x)$, $F(x)$ náhodné veličiny X . Dále určete $E(X)$, $D(X)$, σ a modus.

Řešení

Teorie: 25,26 Příklady: 190,191,192,193

Vyjádříme tabulkou:

$x \in$	$(-\infty, 0)$	$(0, 1)$	$(1, 2)$	$(2, \infty)$
$F(x)$	0	0,8835	0,9965	1

Nyní budeme určovat další číselné charakteristiky náhodné veličiny X . Vycházíme z předdefinovaných vztahů pro střední hodnotu, rozptyl, směrodatnou odchylku a modus:

$$E(X) = \mu = \sum_{i=1}^3 x_i p(x_i) = 0 \cdot 0,8835 + 1 \cdot 0,113 + 2 \cdot 0,0035 = 0,12$$

$$D(X) = \sigma^2 = E(X^2) - [E(X)]^2 = \sum_{i=1}^3 x_i^2 p(x_i) - [E(X)]^2 = 0^2 \cdot 0,8835 + 1^2 \cdot 0,113 + 2^2 \cdot 0,0035 - (0,12)^2 = 0,1126$$

$$\begin{aligned} \text{(nebo } D(X) = \sigma^2 &= \sum_{i=1}^3 (x_i - \mu)^2 p(x_i) = (0 - 0,12)^2 \cdot 0,8835 + (1 - 0,12)^2 \cdot 0,113 + (2 - 0,12)^2 \cdot 0,0035 \\ &= 0,0127224 + 0,0875072 + 0,0123704 = 0,1126) \end{aligned}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{D(X)} = \sqrt{0,1126} \doteq 0,3356$$

$Mo = 0 \dots$ je to hodnota, kterou DNV nabývá s největší pravděpodobností (tj. pro $x = 0$)

Poznámky

Střední hodnota

$$E(X) = \sum_i x_i p(x_i)$$

Rozptyl

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \sum_i x_i^2 p(x_i) - [E(X)]^2 \end{aligned}$$

Směrodatná odchylka

$$\sigma = \sqrt{D(X)}$$

114 - Číselné charakteristiky

Zadání Z kamerového záznamu na křižovatce v jistém městě byl zjištěn počet aut, která projely na červenou během jednoho dne. Hodnoty byly zaznamenány do následující tabulky:

počet aut, které projely na červenou/den	0	1	2	3	4	5
počet dnů	2	3	10	6	3	6

Určete $p(x)$, $F(x)$, $E(X)$, $D(X)$ náhodné veličiny X , která představuje počet motorových vozidel, které projely na červenou během jednoho dne.

Řešení **Teorie: 24,26 Příklady: 190,191,192,193**

Celkový počet sledovaných dnů: $2+3+10+6+3+6=30$ dnů

X	0	1	2	3	4	5
$p(x)$	$\frac{2}{30}$	$\frac{3}{30}$	$\frac{10}{30}$	$\frac{6}{30}$	$\frac{3}{30}$	$\frac{6}{30}$

X	$(-\infty, 0)$	$(0, 1)$	$(1, 2)$	$(2, 3)$	$(3, 4)$	$(4, 5)$	$(5, \infty)$
$F(x)$	0	$\frac{2}{30}$	$\frac{5}{30}$	$\frac{15}{30}$	$\frac{21}{30}$	$\frac{24}{30}$	$\frac{30}{30} = 1$

Výpočet střední hodnoty a rozptylu provedeme pomocí daných vztahů:

$$E(X) = \mu = \sum_{i=1}^6 x_i p(x_i) = 0 \cdot \frac{2}{30} + 1 \cdot \frac{3}{30} + 2 \cdot \frac{10}{30} + 3 \cdot \frac{6}{30} + 4 \cdot \frac{3}{30} + 5 \cdot \frac{6}{30} = \frac{3 + 20 + 18 + 12 + 30}{30} = \frac{83}{30} \doteq 2,7667$$

$$D(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum_{i=1}^6 x_i^2 p(x_i) = 0^2 \cdot \frac{2}{30} + 1^2 \cdot \frac{3}{30} + 2^2 \cdot \frac{10}{30} + 3^2 \cdot \frac{6}{30} + 4^2 \cdot \frac{3}{30} + 5^2 \cdot \frac{6}{30} = 0 + \frac{3}{30} + \frac{40}{30} + \frac{54}{30} + \frac{48}{30} + \frac{150}{30} = \frac{295}{30} \doteq 9,8333$$

$$D(X) = \frac{295}{30} - \left(\frac{83}{30}\right)^2 \doteq 2,1571$$

Poznámky

Diskrétní náhodná veličina

Pravděpodobnostní funkce
 $p(x) = P(X = x)$

Vlastnosti

- $p(x_i) \geq 0$
- $\sum_{i=1}^n p(x_i) = 1$

Distribuční funkce
 $F(x) = P(X < x)$
 $= \sum_{x_i < x} P(X = x_i)$

Vlastnosti

- $0 \leq F(x) \leq 1$

Střední hodnota
 $E(X) = \sum_i x_i p(x_i)$

Rozptyl
 $D(X) = E(X^2) - [E(X)]^2$
 $= \sum_i x_i^2 p(x_i) - [E(X)]^2$

115 - Rozdělení spojité náhodné veličiny

Zadání Náhodná veličina X má funkci hustoty

$$f(x) = \begin{cases} Ce^{-2x} & \text{pro } 0 < x < 2 \\ 0 & \text{pro ostatní } x \end{cases}.$$

Určete konstantu C , $P(X < 1)$, $P(X > \frac{1}{2})$.

Řešení

Teorie: 25,27 Příklady: 194,195

Poněvadž $\int_{-\infty}^{\infty} f(x) dx = 1$, je $\int_{-\infty}^0 0 dx + \int_0^2 f(x) dx + \int_2^{\infty} 0 dx = 1$

takže $\int_0^2 f(x) dx = 1 \Rightarrow C \int_0^2 e^{-2x} dx = C \left[\frac{e^{-2x}}{-2} \right]_0^2 = C \left(\frac{e^{-4}}{-2} + \frac{1}{2} \right) = 1 \Rightarrow C \cdot \frac{1 - e^{-4}}{2} = 1$, odtud

$$C = \frac{2}{1 - e^{-4}}$$

$$\begin{aligned} P(X < 1) &= \int_0^1 \frac{2}{1 - e^{-4}} \cdot e^{-2x} dx = \frac{2}{1 - e^{-4}} \left[\frac{e^{-2x}}{-2} \right]_0^1 = \frac{2}{1 - e^{-4}} \left(\frac{e^{-2}}{-2} + \frac{1}{2} \right) = \frac{2}{1 - e^{-4}} \cdot \frac{1 - e^{-2}}{2} = \frac{\frac{e^2 - 1}{e^2}}{\frac{e^4 - 1}{e^4}} \\ &= \frac{(e^2 - 1)e^2}{(e^2 - 1)(e^2 + 1)} = \frac{e^2}{e^2 + 1} \end{aligned}$$

$$\begin{aligned} P(X > \tfrac{1}{2}) &= \int_{\frac{1}{2}}^2 \frac{2}{1 - e^{-4}} \cdot e^{-2x} dx = \frac{2}{1 - e^{-4}} \left[\frac{e^{-2x}}{-2} \right]_{\frac{1}{2}}^2 = -\frac{1}{1 - e^{-4}} \left[e^{-2x} \right]_{\frac{1}{2}}^2 = -\frac{1}{1 - e^{-4}} (e^{-4} - e^{-1}) \\ &= -\frac{1}{1 - \frac{1}{e^4}} \left(\frac{1}{e^4} - \frac{1}{e} \right) = -\frac{1}{\frac{e^4 - 1}{e^4}} \cdot \frac{1 - e^3}{e^4} = \frac{e^3 - 1}{e^4 - 1} \end{aligned}$$

Poznámky

Spojité náhodné veličiny

Vlastnosti funkce hustoty

- $f(x) \geq 0$
- $f(x) = F'(x)$
- $\lim_{x \rightarrow -\infty} f(x) = 0$
- $\lim_{x \rightarrow \infty} f(x) = 0$
- $\int_{-\infty}^{\infty} f(x) dx = 1$

Vlastnosti distribuční funkce

- $0 \leq F(x) \leq 1$
- $F(x)$ je neklesající funkce
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow \infty} F(x) = 1$
- $P(X < x) = F(x)$
- $P(X > x) = 1 - F(x)$

116 - Číselné charakteristiky

Zadání Náhodná veličina X má distribuční funkci

$$F(x) = \begin{cases} 0 & x \leq -1, \\ \frac{x^2}{2} + x + \frac{1}{2} & -1 < x \leq 0 \\ -\frac{x^2}{2} + x + \frac{1}{2} & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}.$$

Určete střední hodnotu a směrodatnou odchylku.

Řešení

Teorie: 25,27,29 Příklady: 194,195

K výpočtu je potřeba znát funkci hustoty NV X . Z vlastností funkce hustoty platí:

$$f(x) = F'(x) = \begin{cases} 0 & x \leq -1, \\ x+1 & -1 < x \leq 0 \\ -x+1 & 0 < x \leq 1 \\ 0 & x > 1 \end{cases}.$$

Odtud

$$\begin{aligned} E(X) = \mu &= \int_{-1}^0 x f(x) dx + \int_0^1 x f(x) dx = \int_{-1}^0 x(x+1) dx + \int_0^1 x(-x+1) dx = \left[\frac{x^3}{3} + \frac{x^2}{2} \right]_{-1}^0 + \left[-\frac{x^3}{3} + \frac{x^2}{2} \right]_0^1 \\ &= 0 + \frac{1}{3} - \frac{1}{2} - \frac{1}{3} + \frac{1}{2} - 0 = 0 \\ \sigma^2 = D(X) &= E(X^2) - [E(X)]^2 = \int_{-1}^0 x^2 f(x) dx + \int_0^1 x^2 f(x) dx - [E(X)]^2 = \int_{-1}^0 x^2(x+1) dx + \int_0^1 x^2(-x+1) dx - 0 \\ &= \left[\frac{x^4}{4} + \frac{x^3}{3} \right]_{-1}^0 + \left[-\frac{x^4}{4} + \frac{x^3}{3} \right]_0^1 = -\frac{1}{4} + \frac{1}{3} - \frac{1}{4} + \frac{1}{3} = -\frac{1}{2} + \frac{2}{3} = \frac{1}{6} \Rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{D(X)} = \sqrt{\frac{1}{6}} = \frac{\sqrt{6}}{6} \end{aligned}$$

nebo

$$\sigma^2 = [v_2] = \int_{-1}^0 (x - \mu)^2 f(x) dx + \int_0^1 (x - \mu)^2 f(x) dx = \int_{-1}^0 (x - 0)^2 (x+1) dx + \int_0^1 (x - 0)^2 (-x+1) dx = \dots = \frac{1}{6}$$

Poznámky

Spojité náhodné veličiny

Vlastnosti $f(x)$ a $F(x)$

- $f(x) \geq 0$
- $f(x) = F'(x)$
- $\lim_{x \rightarrow -\infty} f(x) = 0$
- $\lim_{x \rightarrow \infty} f(x) = 0$
- $P(x_1 \leq X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$

Střední hodnota

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Rozptyl

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - [E(X)]^2 \end{aligned}$$

Směrodatná odchylka

$$\sigma = \sqrt{\sigma^2} = \sqrt{D(X)}$$

117 - Náhodná veličina - část 1.

Zadání Určete střední hodnotu, rozptyl, šikmost a špičatost náhodné veličiny X s funkcí hustoty

$$f(x) = \begin{cases} \frac{3}{7x^4} & \text{pro } x \in \langle \frac{1}{2}; 1 \rangle \\ 0 & \text{pro ostatní } x \end{cases}.$$

Řešení**Teorie: 29 Příklady: 195**

Výpočet střední hodnoty:

$$E(X) = \int_{\frac{1}{2}}^1 x f(x) dx = \int_{\frac{1}{2}}^1 x \cdot \frac{3}{7x^4} dx = \frac{3}{7} \int_{\frac{1}{2}}^1 \frac{1}{x^3} dx = \frac{3}{7} \left[\frac{x^{-2}}{-2} \right]_{\frac{1}{2}}^1 = \frac{3}{7} \cdot \left(-\frac{1}{2}\right) \left[\frac{1}{x^2} \right]_{\frac{1}{2}}^1 = -\frac{3}{14} \cdot (1 - 4) = \frac{9}{14} = 0,642857$$

Výpočet rozptylu:

$$E(X^2) = \int_{\frac{1}{2}}^1 x^2 f(x) dx = \int_{\frac{1}{2}}^1 x^2 \cdot \frac{3}{7x^4} dx = \frac{3}{7} \int_{\frac{1}{2}}^1 \frac{1}{x^2} dx = \frac{3}{7} \left[\frac{x^{-1}}{-1} \right]_{\frac{1}{2}}^1 = -\frac{3}{7} \left[\frac{1}{x} \right]_{\frac{1}{2}}^1 = -\frac{3}{7} \cdot (1 - 2) = \frac{3}{7} = 0,428571$$

Nyní vše dosadíme do vztahu pro výpočet rozptylu:

$$D(X) = E(X^2) - [E(X)]^2 = 0,4286 - 0,6429^2 \doteq 0,0153$$

Pro zjištění dalších charakteristik je potřeba vypočítat počáteční momenty $\mu_1, \mu_2, \mu_3, \mu_4$ a směrodatnou odchylku σ :

$$\mu_1 = \mu = E(X) = \int_{\frac{1}{2}}^1 x f(x) dx = \dots = 0,642857$$

$$\mu_2 = E(X^2) = \int_{\frac{1}{2}}^1 x^2 f(x) dx = \dots = 0,428571 \quad (\text{z předešlých výpočtů})$$

$$\mu_3 = \int_{\frac{1}{2}}^1 x^3 f(x) dx = \int_{\frac{1}{2}}^1 x^3 \cdot \frac{3}{7x^4} dx = \frac{3}{7} \int_{\frac{1}{2}}^1 \frac{1}{x} dx = \frac{3}{7} [\ln x]_{\frac{1}{2}}^1 = \frac{3}{7} (\ln 1 - \ln 2^{-1}) = \frac{3}{7} \ln 2 = 0,297063$$

$$\mu_4 = \int_{\frac{1}{2}}^1 x^4 f(x) dx = \int_{\frac{1}{2}}^1 x^4 \cdot \frac{3}{7x^4} dx = \int_{\frac{1}{2}}^1 \frac{3}{7} dx = \frac{3}{7} [x]_{\frac{1}{2}}^1 = \frac{3}{7} (1 - \frac{1}{2}) = \frac{3}{7} \cdot \frac{1}{2} = \frac{3}{14} = 0,214286$$

Poznámky

Střední hodnota

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Rozptyl

$$D(X) = E(X^2) - [E(X)]^2 \\ = \int_{-\infty}^{\infty} x^2 f(x) dx - [E(X)]^2$$

Směrodatná odchylka

$$\sigma = \sqrt{\sigma^2} = \sqrt{D(X)}$$

Koeficient šikmosti (asymetrie)

$$A = \bar{v}_3 = \frac{\nu_3}{\sigma^3} = \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{\sigma^3}$$

Koeficient špičatosti (excesu)

$$\bar{e} = \bar{v}_4 = \frac{\nu_4}{\sigma^4} - 3 \\ = \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{\sigma^4} - 3$$

118 - Náhodná veličina - část 2.

Zadání Určete střední hodnotu, rozptyl, šikmost a špičatost náhodné veličiny X s funkcí hustoty

$$f(x) = \begin{cases} \frac{3}{7x^4} & \text{pro } x \in \langle \frac{1}{2}; 1 \rangle \\ 0 & \text{pro ostatní } x \end{cases}.$$

Řešení

Teorie: 29 Příklady: 195

$$\sigma = \sqrt{D(X)} = \sqrt{0,0153} \doteq 0,1237$$

Tedy

šikmost

$$A = \bar{v}_3 = \frac{\nu_3}{\sigma^3} = \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{\sigma^3} = \frac{0,297063 - 3 \cdot 0,428571 \cdot 0,642857 + 2 \cdot 0,642857^3}{0,1237^3} = \dots = 0,98996$$

špičatost

$$\begin{aligned} \bar{e} = \bar{v}_4 &= \frac{\nu_4}{\sigma^4} - 3 = \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{\sigma^4} - 3 \\ &= \frac{0,214286 - 4 \cdot 0,297063 \cdot 0,642857 + 6 \cdot 0,428571 \cdot 0,642857^2 - 3 \cdot 0,642857^4}{0,1237^4} - 3 = \dots = 0,1026 \end{aligned}$$

Poznámky

Střední hodnota

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Rozptyl

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - [E(X)]^2 \end{aligned}$$

Směrodatná odchylka

$$\sigma = \sqrt{\sigma^2} = \sqrt{D(X)}$$

Koeficient šikmosti (asymetrie)

$$A = \bar{v}_3 = \frac{\nu_3}{\sigma^3} = \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{\sigma^3}$$

Koeficient špičatosti (excesu)

$$\begin{aligned} \bar{e} = \bar{v}_4 &= \frac{\nu_4}{\sigma^4} - 3 \\ &= \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{\sigma^4} - 3 \end{aligned}$$

119 - Kvantilové charakteristiky - část 1.

Zadání Určete medián, horní a dolní kvartil náhodné veličiny X s distribuční funkcí

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - \frac{1}{x^3} & x \geq 1 \end{cases}.$$

Řešení

Teorie: 30 Příklady: 196

Pro zjištění hodnoty kvartilů $x_{0,25}; x_{0,5}; x_{0,75}$ vycházíme z toho, že $F(x_p) = p$

Medián $x_{0,5} = Me$, dosadíme

$$\begin{aligned} F(x_{0,5}) &= 0,5 \\ 1 - \frac{1}{x_{0,5}^3} &= 0,5 \\ 1 - 0,5 &= \frac{1}{x_{0,5}^3} \\ x_{0,5}^3 &= \frac{1}{0,5} \\ x_{0,5} &= \sqrt[3]{\frac{1}{0,5}} = \frac{1}{0,7937} \doteq 1,260 \end{aligned}$$

Horní kvartil $x_{0,75}$:

$$\begin{aligned} F(x_{0,75}) &= 0,75 \\ 1 - \frac{1}{x_{0,75}^3} &= 0,75 \\ \frac{1}{x_{0,75}^3} &= 1 - 0,75 \\ x_{0,75}^3 &= \frac{1}{0,25} \\ x_{0,75} &= \sqrt[3]{\frac{1}{0,25}} = \frac{1}{0,62996} \doteq 1,5874 \end{aligned}$$

Poznámky

Spojité náhodné veličiny

$F(x_p) = p$, kde $p \in \langle 0, 1 \rangle$

x_p ... p -kvantil

$F(x)$... distribuční funkce

Kvantily:

- kvantily: $x_{0,25}; x_{0,5}; x_{0,75}$
- decily: $x_{0,1}; x_{0,2}; \dots; x_{0,9}$

120 - Kvantilové charakteristiky - část 2.

Zadání Určete medián, horní a dolní kvartil náhodné veličiny X s distribuční funkcí

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - \frac{1}{x^3} & x \geq 1 \end{cases}.$$

Řešení

Teorie: 30 Příklady: 196

Dolní kvartil $x_{0,25}$:

$$\begin{aligned} F(x_{0,25}) &= 0,25 \\ 1 - \frac{1}{x_{0,25}^3} &= 0,25 \\ 0,75 &= \frac{1}{x_{0,25}^3} \\ x_{0,25}^3 &= \frac{1}{0,75} \\ x_{0,25} &= \sqrt[3]{\frac{1}{0,75}} = \frac{1}{0,90856} \doteq 1,10064 \end{aligned}$$

Poznámky

Spojité náhodné veličiny

$F(x_p) = p$, kde $p \in \langle 0, 1 \rangle$

x_p ... p -kvantil

$F(x)$... distribuční funkce

Kvantily:

- kvantily: $x_{0,25}; x_{0,5}; x_{0,75}$
- decily: $x_{0,1}; x_{0,2}; \dots; x_{0,9}$

Řešené příklady – Rozdělení diskrétní náhodné veličiny

122 - Binomické rozdělení

Zadání Průměrná zmetkovitost výroby sledovaného výrobku je 2 %. Jaká je pravděpodobnost, že se mezi 300 výrobky vyskytne a) právě pět zmetků, b) více než tři zmetky? Dále určete střední hodnotu a rozptyl počtu zmetků mezi těmito 300 výrobky.

Řešení

Teorie: 33 Příklady: 199,202

Pravděpodobnost zmetkovitosti jednoho výrobku je $p = 0,02$. Uvedených 300 výrobků pak představuje 300 nezávislých pokusů, přičemž v každém je pravděpodobnost „úspěchu“ (výrobek je zmetek) rovna p .

Náhodná veličina X popisující počet zmetků mezi sledovanými výrobky má tedy binomické rozdělení s parametry $n = 300$ a $p = 0,02$, tj. $X \sim Bi(300; 0,02)$.

ad a)

Máme určit pravděpodobnost toho, že náhodná veličina X nabude hodnoty právě 5:

$$P(X = 5) = p(5) = \binom{300}{5} \cdot 0,02^5 \cdot 0,98^{295}.$$

EXCEL:

$$P(X = 5) = \text{BINOM.DIST}(5; 300; 0,02; 0) \doteq 0,162$$

ad b)

Určíme, s jakou pravděpodobností bude hodnota náhodné veličiny X větší než 3 (tj. $3 < X < +\infty$):

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^3 P(X = x) = 1 - \sum_{x=0}^3 p(x) = 1 - \sum_{x=0}^3 \binom{300}{x} \cdot 0,02^x \cdot 0,98^{300-x}.$$

EXCEL:

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \text{BINOM.DIST}(3; 300; 0,02; 1) \doteq 0,851$$

Dále

$$E(X) = np = 300 \cdot 0,02 = 6$$

$$D(X) = np(1 - p) = 300 \cdot 0,02 \cdot 0,98 = 5,88.$$

Poznámky

Binomické rozdělení $Bi(n, p)$

n - počet nezávislých pokusů

p - pravděpodobnost úspěchu při jednom pokusu

Pravděpodobnostní funkce

$$p(x) = \binom{n}{x} p^x (1 - p)^{n-x},$$

$$x = 0, 1, \dots, n$$

Vlastnosti

$$E(X) = np$$

$$D(X) = np(1 - p)$$

EXCEL:

$$P(X = x) = \text{BINOM.DIST}(x; n; p; 0)$$

$$P(X \leq x) = \text{BINOM.DIST}(x; n; p; 1)$$

123 - Hypergeometrické rozdělení

Zadání Do krabice, která obsahuje 150 žárovek se závitem E27, přidáme 50 žárovek se závitem E14. Z krabice pak náhodně vybereme 30 žárovek. Jaká je pravděpodobnost, že mezi nimi bude alespoň 25 se závitem E27? A jaký je průměrný počet žárovek E27 v takovém výběru?

Řešení **Teorie: 34 Příklady: 200,202**

Celkem máme v krabici $150 + 50 = 200$ žárovek, ze kterých náhodně a bez vracení vybíráme 30. Tento výběr tedy představuje 30 závislých pokusů - pravděpodobnost, že vyberete žárovku E27 v každém pokusu závisí na tom, jaké žárovky jste vybrali v předchozích pokusech.

Náhodná veličina X popisující počet žárovek E27 ve 30ti prvkovém výběru má tedy hypergeometrické rozdělení s parametry $N = 200$ (počet všech žárovek), $M = 150$ (počet žárovek E27) a $n = 30$ (počet žárovek ve výběru), tj. $X \sim H(200, 150, 30)$.

Máme určit pravděpodobnost toho, že náhodná veličina X nabude hodnoty alespoň 25 (tj. $25 \leq X \leq 30$):

$$P(X \geq 25) = \sum_{x=25}^{30} P(X = x) = \sum_{x=25}^{30} p(x) = \sum_{x=25}^{30} \frac{\binom{150}{x} \binom{50}{30-x}}{\binom{200}{30}}.$$

EXCEL:

$$P(X \geq 25) = 1 - P(X < 25) = 1 - P(X \leq 24) = 1 - \text{HYPGEOM.DIST}(24; 30; 150; 200; 1) \doteq 0,181$$

Průměrný počet žárovek E27 ve výběru určíme jako střední hodnotu náhodné veličiny X :

$$E(X) = \frac{nM}{N} = \frac{30 \cdot 150}{200} = 22,5.$$

Poznámky

Hypergeometrické rozdělení $H(N, M, n)$

N - počet prvků základního souboru

M - počet prvků základního souboru se sledovanou vlastností

n - počet prvků ve výběru

Pravděpodobnostní funkce

$$p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, \dots, n$$

Vlastnosti

$$E(X) = \frac{nM}{N}$$

$$D(X) = \frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$$

EXCEL:

$$P(X = x) = \text{HYPGEOM.DIST}(x; n; M; N; 0)$$

$$P(X \leq x) = \text{HYPGEOM.DIST}(x; n; M; N; 1)$$

124 - Poissonovo rozdělení

Zadání Během dvanáctihodinové pracovní směny přijede na myčku s jednou mycí linkou průměrně 144 automobilů.

- Jaká je pravděpodobnost, že jich během 40 min. nepřijede více než 6?
- Jaká je pravděpodobnost, že během 10 min. přijede alespoň jeden automobil?

Řešení

Teorie: 35 Příklady: 201,202

ad a)

Průměrný počet automobilů, které přijedou na myčku během 1 min. je $\frac{144}{12 \cdot 60}$, v časovém úseku délky 40 min. jich pak bude průměrně $\frac{144}{12 \cdot 60} \cdot 40 = 8$.

Náhodná veličina X popisující počet automobilů, které přijedou na myčku během 40 min., má tedy Poissonovo rozdělení s parametrem $\lambda = 8$, tj. $X \sim Po(8)$.

Určíme pravděpodobnost, že náhodná veličina X nabude hodnot nejvýše 6 (tj. $0 \leq X \leq 6$):

$$P(X \leq 6) = \sum_{x=0}^6 P(X = x) = \sum_{x=0}^6 p(x) = e^{-8} \sum_{x=0}^6 \frac{8^x}{x!}.$$

EXCEL:

$$P(X \leq 6) = \text{POISSON.DIST}(6; 8; 1) \doteq 0,313$$

ad b)

Analogicky, náhodná veličina Y popisující počet automobilů, které přijedou na myčku během 10 min., bude mít Poissonovo rozdělení s parametrem $\lambda = \frac{144}{12 \cdot 60} \cdot 10 = 2$, tj. $Y \sim Po(2)$, a určíme pravděpodobnost, že tato náhodná veličina nabude hodnot alespoň 1 (tj. $1 \leq Y \leq +\infty$):

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - p(0) = 1 - \frac{2^0}{0!} e^{-2} = 1 - \frac{1}{e^2} \doteq 0,865.$$

EXCEL:

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \text{POISSON.DIST}(0; 2; 0) \doteq 0,865$$

Poznámky

Poissonovo rozdělení $Po(\lambda)$

λ - průměrný počet výskytů sledovaného jevu v úseku jednotkové délky

Pravděpodobnostní funkce

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

$$x = 0, 1, \dots$$

Vlastnosti

$$E(X) = \lambda$$

$$D(X) = \lambda$$

EXCEL:

$$P(X = x) = \text{POISSON.DIST}(x; \lambda; 0)$$

$$P(X \leq x) = \text{POISSON.DIST}(x; \lambda; 1)$$

Řešené příklady – Rozdělení spojité náhodné veličiny

126 - Rovnoměrné rozdělení

Zadání Z konečné zastávky Studentské koleje odjíždí během dne autobusy pravidelně každých 20 min. Jaká je pravděpodobnost, že při náhodném příchodu na zastávku nebudete čekat déle než 12 min.? Dále určete střední hodnotu a směrodatnou odchylku náhodné veličiny popisující dobu čekání na odjezd autobusu při náhodném příchodu na uvedenou zastávku.

Řešení

Teorie: 37 Příklady: 204,207

Hodnoty náhodné veličiny X popisující dobu čekání při náhodném příchodu leží v intervalu $\langle 0, 20 \rangle$ a všechny mají stejnou možnost výskytu, tato náhodná veličina má tedy rovnoměrné rozdělení s parametry $a = 0$ a $b = 20$, tj. $X \sim R(0, 20)$. Její hustota pravděpodobnosti a distribuční funkce jsou pak ve tvaru:

$$f(x) = \begin{cases} \frac{1}{20} & x \in \langle 0, 20 \rangle \\ 0 & x \notin \langle 0, 20 \rangle \end{cases}, \quad F(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ \frac{x}{20} & x \in \langle 0, 20 \rangle \\ 1 & x \in (20, +\infty) \end{cases}.$$

Pravděpodobnost, že doba čekání nepřesáhne 12 min.:

$$P(X \leq 12) = P(X < 12) = F(12) = \frac{12}{20} = 0,6.$$

U spojitě náhodné veličiny je $P(X = x) = 0$, tj. $P(X \leq x) = P(X < x) = F(x)$.

Střední hodnota:

$$E(X) = \frac{a+b}{2} = \frac{0+20}{2} = 10.$$

Rozptyl:

$$D(X) = \sigma^2 = \frac{(b-a)^2}{12} = \frac{(20-0)^2}{12} = \frac{100}{3} \doteq 33,33.$$

Směrodatná odchylka:

$$\sigma = \sqrt{D(X)} = \sqrt{\frac{100}{3}} = \frac{10\sqrt{3}}{3} \doteq 5,77.$$

Poznámky

Rovnoměrné rozdělení $R(a, b)$
 a, b - krajní meze intervalu

Hustota pravděpodobnosti

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in \langle a, b \rangle \\ 0 & x \notin \langle a, b \rangle \end{cases}$$

Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, a) \\ \frac{x-a}{b-a} & x \in \langle a, b \rangle \\ 1 & x \in (b, +\infty) \end{cases}$$

Vlastnosti

$$E(X) = \frac{a+b}{2}$$

$$D(X) = \frac{(b-a)^2}{12}$$

127 - Exponenciální rozdělení

Zadání

- a) Průměrná doba mezi příjezdy automobilů na hraniční přechod je 7 min. Jaká je pravděpodobnost, že mezi příjezdy nebude větší prodleva než 10 min.?
- b) Průměrná životnost sledovaného výrobku byla experimentálně stanovena na 1150 hodin. Jakou životnost má uvádět výrobce ve svých materiálech, jestliže chce, aby deklarované životnosti dosáhlo minimálně 80 % výrobků?

Řešení

Teorie: 38 Příklady: 205,207

ad a)

Náhodná veličina X popisující dobu mezi příjezdy automobilů na hraniční přechod má exponenciální rozdělení s parametrem $\lambda = \frac{1}{7}$. Máme určit pravděpodobnost, že hodnota této náhodné veličiny bude nejvýše 10 (tj. $0 \leq X \leq 10$).

$$P(X \leq 10) = P(X < 10) = F(10) = 1 - e^{-\frac{10}{7}} \doteq 0,76$$

EXCEL:

$$P(X \leq 10) = P(X < 10) = F(10) = \text{EXPON.DIST}(10; 1/7; 1) \doteq 0,76$$

ad b)

Náhodná veličina X popisující životnost sledovaného výrobku má exponenciální rozdělení s parametrem $\lambda = \frac{1}{1150}$. Hledáme takovou hodnotu životnosti x , aby platilo

$$\begin{aligned} P(X \geq x) &= 0,8 \\ 1 - P(X < x) &= 0,8 \\ 1 - F(x) &= 0,8 \\ F(x) &= 0,2 \\ 1 - e^{-\frac{x}{1150}} &= 0,2 \\ x &= -1150 \cdot \ln(0,8) \\ x &\doteq 257 \end{aligned}$$

Poznámky

Exponenciální rozdělení $E(\lambda)$
 λ - převrácená hodnota střední hodnoty

Hustota pravděpodobnosti

$$f(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ \lambda e^{-\lambda x} & x \in \langle 0, +\infty \end{cases}$$

Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ 1 - e^{-\lambda x} & x \in \langle 0, +\infty \end{cases}$$

Vlastnosti

$$\begin{aligned} E(X) &= \frac{1}{\lambda} \\ D(X) &= \frac{1}{\lambda^2} \end{aligned}$$

EXCEL:

$$\begin{aligned} f(x) &= \text{EXPON.DIST}(x; \lambda; 0) \\ F(x) &= \text{EXPON.DIST}(x; \lambda; 1) \end{aligned}$$

128 - Normální rozdělení - část 1.

Zadání Náhodná veličina popisující skutečnou hmotnost expedovaných 25 kg pytlů cementu má při dodržení standardních výrobních podmínek normální rozdělení se střední hodnotou 24,8 kg a směrodatnou odchylkou 0,6 kg.

- S jakou pravděpodobností vyhoví náhodně vybraný pytel normě, která předepisuje hmotnost v rozmezí 24 až 25,5 kg?
- Jaká je pravděpodobnost, že námi zakoupený pytel cementu bude vážit více než 25 kg?
- Určete, jakou hmotnost překročí 85 % všech expedovaných pytlů.
- V jakém rozsahu (symetrickém kolem střední hodnoty) můžeme předpokládat hmotnost 90 % všech expedovaných pytlů?

Řešení

Teorie: 39 Příklady: 206,207

Náhodná veličina X popisující hmotnost expedovaných pytlů cementu má dle zadání normální rozdělení s parametry $\mu = 24,8$ (střední hodnota) a $\sigma^2 = (0,6)^2 = 0,36$ (rozptyl), tj. $X \sim N(24,8; 0,36)$.

ad a)

Hledáme pravděpodobnost toho, že hodnoty náhodné veličiny X leží mezi 24 a 25,5 (tj. $24 \leq X \leq 25,5$):

EXCEL:

$$\begin{aligned} P(24 \leq X \leq 25,5) &= F(25,5) - F(24) \\ &= \text{NORM.DIST}(25,5; 24,8; 0,6; 1) - \text{NORM.DIST}(24; 24,8; 0,6; 1) \doteq 0,787 \end{aligned}$$

ad b)

Nyní máme určit pravděpodobnost, že náhodná veličina X nabývá hodnot větších než 25 (tj. $25 < X < +\infty$):

EXCEL:

$$P(X > 25) = 1 - P(X \leq 25) = 1 - F(25) = 1 - \text{NORM.DIST}(25; 24,8; 0,6; 1) \doteq 0,369$$

ad c)

Hledáme takovou hodnotu x náhodné veličiny X , aby pravděpodobnost toho, že $X \geq x$ byla 85 %:

$$\begin{aligned} P(X \geq x) &= 0,85 \\ 1 - P(X < x) &= 0,85 \\ P(X < x) &= 0,15 \\ F(x) &= 0,15 \end{aligned}$$

Poznámky

Normální rozdělení $N(\mu, \sigma^2)$

μ - střední hodnota

σ^2 - rozptyl

Hustota pravděpodobnosti

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Distribuční funkce

$$F(x) = \int_{-\infty}^x f(t)dt$$

Vlastnosti

$$\begin{aligned} E(X) &= \mu \\ D(X) &= \sigma^2 \end{aligned}$$

EXCEL:

$$f(x) = \text{NORM.DIST}(x; \mu; \sigma; 0)$$

$$F(x) = \text{NORM.DIST}(x; \mu; \sigma; 1)$$

$$x_p = \text{NORM.INV}(p; \mu; \sigma)$$

129 - Normální rozdělení - část 2.

Zadání Náhodná veličina popisující skutečnou hmotnost expedovaných 25 kg pytlů cementu má při dodržení standardních výrobních podmínek normální rozdělení se střední hodnotou 24,8 kg a směrodatnou odchylkou 0,6 kg.

- S jakou pravděpodobností vyhoví náhodně vybraný pytel normě, která předepisuje hmotnost v rozmezí 24 až 25,5 kg?
- Jaká je pravděpodobnost, že námi zakoupený pytel cementu bude vážit více než 25 kg?
- Určete, jakou hmotnost překročí 85 % všech expedovaných pytlů.
- V jakém rozsahu (symetrickém kolem střední hodnoty) můžeme předpokládat hmotnost 90 % všech expedovaných pytlů?

Řešení Teorie: 39 Příklady: 206,207

Hledáme tedy hodnotu 0,15-quantilu $x_{0,15}$ náhodné veličiny X , tj. hodnotu, která rozdělí plochu pod grafem hustoty pravděpodobnosti v poměru 15 : 85.

EXCEL:

$$x_{0,15} = \text{NORM.INV}(0,15;24,8;0,6) \doteq 24,18$$

ad d)

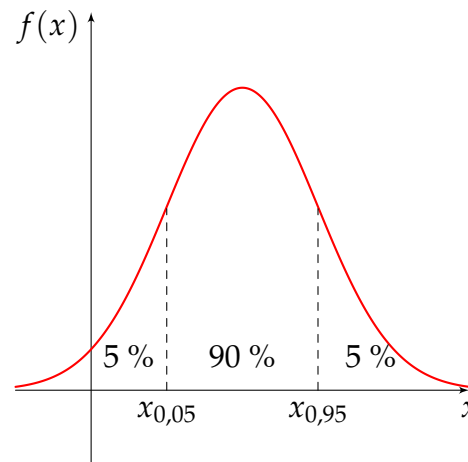
Jak je patrné z obrázku, hledáme takové hodnoty náhodné veličiny X , které rozdělí plochu pod grafem hustoty pravděpodobnosti v poměru 5 : 90 : 5, tj. kvantily $x_{0,05}$ a $x_{0,95}$.

EXCEL:

$$x_{0,05} = \text{NORM.INV}(0,05;24,8;0,6) \doteq 23,81$$

$$x_{0,95} = \text{NORM.INV}(0,95;24,8;0,6) \doteq 25,79$$

Dá se tedy předpokládat, že hmotnost 90 % všech expedovaných pytlů cementu bude v rozmezí od 23,81 do 25,79 kg.



Poznámky

Normální rozdělení $N(\mu, \sigma^2)$

μ - střední hodnota

σ^2 - rozptyl

Hustota pravděpodobnosti

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Distribuční funkce

$$F(x) = \int_{-\infty}^x f(t)dt$$

Vlastnosti

$$E(X) = \mu$$

$$D(X) = \sigma^2$$

EXCEL:

$$f(x) = \text{NORM.DIST}(x; \mu; \sigma; 0)$$

$$F(x) = \text{NORM.DIST}(x; \mu; \sigma; 1)$$

$$x_p = \text{NORM.INV}(p; \mu; \sigma)$$

Řešené příklady – Náhodný vektor

131 - Sdružené rozdělení pravděpodobnosti - část 1.

Zadání Máme rodinu se třemi dětmi. Zavedeme náhodnou veličinu X , která určuje počet synů a NV Y , která vyjadřuje, kolik má prostřední dítě mladších bratrů. Určete sdruženou pravděpodobnostní a distribuční funkci.

Řešení

Teorie: 42 Příklady: 209,213

realizace NV $X \dots X = \{0, 1, 2, 3\}$

realizace NV $Y \dots Y = \{0, 1\}$

Sestavíme sdružené pravděpodobnostní funkce pro všechny možné kombinace. Možností, jak se mohou sourozenci narodit je 8: HHH, HHK, HKH, KHH, KKH, KHK, HKK, KKK.

Při výpočtu pravděpodobností jednotlivých možností využijeme klasické definice (příznivé/všem).

$p(0,0) = P(X=0, Y=0) = \text{žádný kluk a nejmladší je holka} - \text{jedině HHH} = 1/8$

$p(1,0) = P(X=1, Y=0) = \text{jeden kluk a nejmladší je holka} - \text{HHK, HKH} = 2/8, \text{ atd.}$

Sestavíme pravděpodobnostní tabulku:

$X \setminus Y$	0	1
0	0,125	0
1	0,25	0,125
2	0,125	0,25
3	0	0,125

Provedeme kontrolu - součet všech hodnot v tabulce musí být 1 $\left(\sum_i \sum_j p(x_i, y_j) = 1 \right)$.

Pro určení hodnot sdružené distribuční funkce využijeme vztahu v Poznámkách.

$F(0,0) = P(X < 0, Y < 0) = 0$

$F(1,0) = P(X < 1, Y < 0) = 0$

Poznámky

Pravděpodobnostní funkce
 $p(x, y) = P(X = x, Y = y)$

Distribuční funkce
 $F(x, y) = \sum_{x_i < x} \sum_{y_j < y} p(x_i, y_j)$

132 - Sdružené rozdělení pravděpodobnosti - část 2.

Zadání Máme rodinu se třemi dětmi. Zavedeme náhodnou veličinu X , která určuje počet synů a NV Y , která vyjadřuje, kolik má prostřední dítě mladších bratrů. Určete sdruženou pravděpodobnostní a distribuční funkci.

Řešení

$F(1,1) = P(X < 1, Y < 1) =$ počet synů je menší jak 1, tj. 0 a počet mladších bratrů prostředního dítěte je menší jak 1, tj. 0 $= p(0,0) = 1/8$

$F(2,1) = P(X < 2, Y < 1) =$ počet synů je menší jak 2, tj. 0 nebo 1 a počet mladších bratrů prostředního dítěte je menší jak 1, tj. 0 $= p(0,0) + p(1,0) = 3/8$

$F(3,1) = P(X < 3, Y < 1) =$ počet synů je menší jak 3, tj. 0,1 nebo 2 a počet mladších bratrů prostředního dítěte je menší jak 1, tj. 0 $= p(0,0) + p(1,0) + p(2,0) = 4/8$

$F(1,2) = P(X < 1, Y < 2) =$ počet synů je menší jak 1, tj. 0 a počet mladších bratrů prostředního dítěte je menší jak 2, tj. 0 nebo 1 $= p(0,0) + p(0,1) = 1/8$, atd.

Sestavíme tabulku:

$X \setminus Y$	$(-\infty, 0)$	$(0, 1)$	$(1, \infty)$
$(-\infty, 0)$	0	0	0
$(0, 1)$	0	0,125	0,125
$(1, 2)$	0	0,375	0,5
$(2, 3)$	0	0,5	0,875
$(3, \infty)$	0	0,5	1

Poznámky

Pravděpodobnostní funkce
 $p(x, y) = P(X = x, Y = y)$

Distribuční funkce
 $F(x, y) = \sum_{x_i < x} \sum_{y_j < y} p(x_i, y_j)$

133 - Marginální rozdělení, nezávislost složek

Zadání Máme rodinu se třemi dětmi. Zavedeme náhodnou veličinu X , která určuje počet synů a NV Y , která vyjadřuje, kolik má prostřední dítě mladších bratrů. Určete marginální pravděpodobnostní a distribuční funkci. Rozhodněte, zda jsou náhodné veličiny X, Y nezávislé.

Řešení

Teorie: 43 Příklady: 209,212,213

Marginální rozdělení popisuje jednotlivé složky náhodného vektoru, tzn. marginální pravděpodobnost $p_X(x)$ je pravděpodobnostní funkcí náhodné veličiny X .

$p_X(0) = p(0,0) + p(0,1) = 1/8$ - sečteme hodnoty v prvním řádku tabulky sdružené pravděpodobnosti

$p_X(1) = p(1,0) + p(1,1) = 3/8$ - sečteme hodnoty ve druhém řádku tabulky, atd.

Analogicky získáme marginální pravděpodobnost $p_Y(y)$ - sčítáme hodnoty ve sloupcích.

$p_Y(0) = p(0,0) + p(1,0) + p(2,0) + p(3,0) = 4/8$ - sečteme hodnoty v prvním sloupci tabulky sdružené pravděpodobnosti

Sestavíme pravděpodobnostní tabulku rozšířenou i o marginální pravděpodobnosti:

$X \setminus Y$	0	1	$p_X(x)$
0	0,125	0	0,125
1	0,25	0,125	0,375
2	0,125	0,25	0,375
3	0	0,125	0,125
$p_Y(y)$	0,5	0,5	1

Aby byly složky náhodného vektoru nezávislé, musela by být každá z hodnot sdružené pravděpodobnosti rovna součinu příslušných marginálních pravděpodobností. Toto zřejmě neplatí, například:

$$0,25 = p(2,1) \neq p_X(2)p_Y(1) = 0,1875.$$

Náhodné veličiny proto nejsou nezávislé.

Poznámky

Marginální pravděpodobnostní funkce

$$p_X(x) = P(X = x) = \sum_y p(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_x p(x, y)$$

Marginální distribuční funkce

$$F_X(x) = P(X < x) = F(x, \infty)$$

$$F_Y(y) = P(Y < y) = F(\infty, y)$$

Nezávislost složek (X, Y)

X, Y jsou nezávislé, pokud:

$$F(x, y) = F_X(x)F_Y(y)$$

$$p(x, y) = p_X(x)p_Y(y)$$

134 - Marginální rozdělení pravděpodobnosti

Zadání Máme rodinu se třemi dětmi. Zavedeme náhodnou veličinu X , která určuje počet synů a NV Y , která vyjadřuje, kolik má prostřední dítě mladších bratrů. Určete marginální pravděpodobnostní a distribuční funkci.

Řešení **Teorie: 43 Příklady: 209,213**

K nalezení marginálních distribučních funkcí využijeme marginálních pravděpodobnostních funkcí a znalostí z popisu náhodné veličiny.

Marginální distribuční funkce $F_X(x)$.

X	$p_X(x)$		X	$F_X(x)$
0	0,125	$\Rightarrow$	$(-\infty, 0)$	0
1	0,375		$(0, 1)$	0,125
2	0,375		$(1, 2)$	0,5
3	0,125		$(2, 3)$	0,875
			$(3, \infty)$	1

Analogicky určíme marginální distribuční funkci pro NV Y .

Y	$p_Y(y)$		Y	$F_Y(y)$
0	0,5	$\Rightarrow$	$(-\infty, 0)$	0
1	0,5		$(0, 1)$	0,5
			$(1, \infty)$	1

Poznámky

Marginální pravděpodobnostní funkce

$$p_X(x) = P(X = x) = \sum_y p(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_x p(x, y)$$

Marginální distribuční funkce

$$F_X(x) = P(X < x) = F(x, \infty)$$

$$F_Y(y) = P(Y < y) = F(\infty, y)$$

135 - Podmíněné rozdělení pravděpodobnosti

Zadání Pokračujme v příkladu s dětmi. Určete podmíněné pravděpodobnosti.

Řešení **Teorie: 44 Příklady: 210,213**

K určení podmíněných pravděpodobností využijeme vztahů v Poznámkách. Například pravděpodobnost, že rodina má dva syny ($x = 2$), pokud víme, že prostřední dítě má jednoho bratra ($y = 1$) je

$$P(2|Y = 1) = \frac{p(2, 1)}{p_Y(1)} = 0,5.$$

Sestavíme tabulky podmíněných pravděpodobností $P(x|y), P(y|x)$:

$X \setminus Y$		0	1
$P(x y) :$	0	0,25	0
	1	0,5	0,25
	2	0,25	0,5
	3	0	0,25

$X \setminus Y$		0	1
$P(y x) :$	0	1	0
	1	2/3	1/3
	2	1/3	2/3
	3	0	1

Sestavíme tabulky podmíněných distribučních funkcí $F(x|y), F(y|x)$:

$X \setminus Y$		0	1
$F(x y) :$	$(-\infty, 0)$	0	0
	$(0, 1)$	0,25	0
	$(1, 2)$	0,75	0,25
	$(2, 3)$	1	0,75
	$(3, \infty)$	1	1

$X \setminus Y$		$(-\infty, 0)$	$(0, 1)$	$(1, \infty)$
$F(y x) :$	0	0	1	1
	1	0	2/3	1
	2	0	1/3	1
	3	0	0	1

Poznámky

Podmíněná pravděpodobnostní funkce

$$P(x|y) = \frac{p(x, y)}{p_Y(y)}, p_Y(y) \neq 0$$

$$P(y|x) = \frac{p(x, y)}{p_X(x)}, p_X(x) \neq 0$$

Podmíněná distribuční funkce

$$F(x|y) = \frac{\sum_{x < x_i} p(x_i, y)}{p_Y(y)}, p_Y(y) \neq 0$$

$$F(y|x) = \frac{\sum_{y < y_i} p(x, y_i)}{p_X(x)}, p_X(x) \neq 0$$

Pravděpodobnostní tabulka k zadání

$X \setminus Y$	0	1	$p_X(x)$
0	0,125	0	0,125
1	0,25	0,125	0,375
2	0,125	0,25	0,375
3	0	0,125	0,125
$p_Y(y)$	0,5	0,5	1

136 - Číselné charakteristiky

Zadání Pokračujeme v řešeném příkladu, máme náhodný vektor, který je popsán sdruženou i marginální pravděpodobnostní funkcí. Určete:

- a) $E(X), E(Y), D(X), D(Y)$,
- b) $E(\mathbf{X})$ a $D(\mathbf{X})$,
- c) $E(X|Y = 1)$.

Řešení

Teorie: 45 Příklady: 211

Za a) máme určit střední hodnotu a rozptyl náhodných veličin X, Y - marginální charakteristiky, využijeme tedy marginálních pravděpodobností.

$$E(X) = \sum_i x_i p_X(x_i) = 0 \cdot 0,125 + 1 \cdot 0,375 + 2 \cdot 0,375 + 3 \cdot 0,125 = 1,5$$

$$E(Y) = \sum_j y_j p_Y(y_j) = 0 \cdot 0,5 + 1 \cdot 0,5 = 0,5$$

Při určení rozptylu je výhodné použít výpočetní vztah $D(X) = E(X^2) - [E(X)]^2$.

$$D(X) = 0^2 \cdot 0,125 + 1^2 \cdot 0,375 + 2^2 \cdot 0,375 + 3^2 \cdot 0,125 - (1,5)^2 = 0,75$$

$$D(Y) = 0^2 \cdot 0,5 + 1^2 \cdot 0,5 - (0,5)^2 = 0,25$$

ad b)

$$E(\mathbf{X}) = (E(X), E(Y)) = (1,5; 0,5)$$

$$D(\mathbf{X}) = (D(X), D(Y)) = (0,75; 0,25)$$

ad c)

Pro výpočet $E(X|Y = 1)$ potřebujeme znát podmíněnou pravděpodobnostní funkci $P(x|y)$, kterou jsme již určili (135).

$$E(X|Y = 1) = \sum_i x_i p(x_i|1) = 0 \cdot 0 + 1 \cdot 0,25 + 2 \cdot 0,5 + 3 \cdot 0,25 = 2$$

Poznámky

Marginální charakteristiky

$$E(X) = \sum_i x_i p_X(x_i)$$

$$E(Y) = \sum_j y_j p_Y(y_j)$$

$$D(X) = \sum_i (x_i - E(X))^2 p_X(x_i)$$

$$D(Y) = \sum_j (y_j - E(Y))^2 p_Y(y_j)$$

Podmíněná střední hodnota

$$E(X|Y = y) = \sum_i x_i p(x_i|y)$$

Pravděpodobnostní tabulka k zadání

$X \setminus Y$	0	1	$p_X(x)$
0	0,125	0	0,125
1	0,25	0,125	0,375
2	0,125	0,25	0,375
3	0	0,125	0,125
$p_Y(y)$	0,5	0,5	1

137 - Kovariance a korelační koeficient

Zadání Pokračujeme v řešeném příkladu, máme náhodný vektor, který je popsán sdruženou pravděpodobnostní funkcí. Určete kovarianci a korelační koeficient.

Řešení **Teorie: 46 Příklady: 211,212**

Podobně jako u výpočtu rozptylu je výhodnější pro výpočet kovariance využít tzv. výpočetní vztah uvedený v Poznámkách.

$$E(XY) = \sum_i \sum_j x_i y_j p(x_i, y_j) = 0 \cdot 0 \cdot 0,125 + 0 \cdot 1 \cdot 0 + 1 \cdot 0 \cdot 0,25 + 1 \cdot 0 \cdot 0,125 + \dots + 3 \cdot 1 \cdot 0,125 = 1$$

$$\text{cov}(X, Y) = 1 - 1,5 \cdot 0,5 = 0,25$$

$$\text{cov}(Y, X) = \text{cov}(X, Y)$$

Kovarianční matice je:

$$\begin{pmatrix} 0,75 & 0,25 \\ 0,25 & 0,25 \end{pmatrix}.$$

S využitím kovarianční matice určíme hodnotu korelačního koeficientu.

$$\rho(X, Y) = \frac{0,25}{\sqrt{0,75 \cdot 0,25}} = 0,577$$

$$\rho(Y, X) = \rho(X, Y)$$

Korelační matice je:

$$\begin{pmatrix} 1 & 0,577 \\ 0,577 & 1 \end{pmatrix}.$$

Na základě korelačního koeficientu lze říci, že mezi náhodnými veličinami existuje středně silná pozitivní korelace, tj. že pravděpodobně s růstem X bude lineárně růst Y .

Poznámky

Kovariance

$$\begin{aligned} \text{cov}(X, Y) &= E(XY) - E(X)E(Y) \\ &= \sum_i \sum_j x_i y_j p(x_i, y_j) - E(X)E(Y) \end{aligned}$$

Kovarianční matice

$$\begin{pmatrix} D(X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & D(Y) \end{pmatrix}$$

Koeficient korelace

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}}$$

Korelační matice

$$\begin{pmatrix} 1 & \rho(X, Y) \\ \rho(Y, X) & 1 \end{pmatrix}$$

Pravděpodobnostní tabulka k zadání

$X \setminus Y$	0	1
0	0,125	0
1	0,25	0,125
2	0,125	0,25
3	0	0,125

Řešené příklady – Popisná analýza

139 - Třídění, četnosti

Zadání Provádíme výzkum o počtu televizí v domácnostech. Náhodně jsme oslovili 16 domácností a zjistili následující údaje. Proveďte prosté třídění a určete četnosti pro každou třídu, sestavte histogram.

2 1 0 1 2 4 3 1 2 4 2 1 0 2 3 2

Řešení

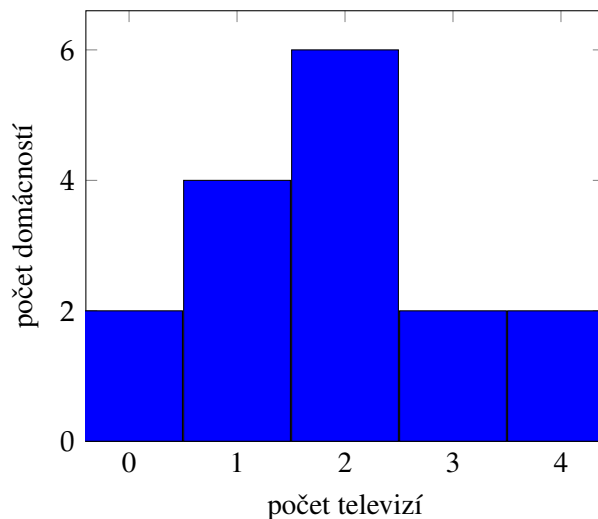
Teorie: 49,50 Příklady: 215

Sestavíme tabulku četností:

Počet televizí	Absolutní četnost	Kumulativní abs. četnost	Relativní četnost	Kumulativní rel. četnost
0	2	2	0,125	0,125
1	4	6	0,25	0,375
2	6	12	0,375	0,75
3	2	14	0,125	0,875
4	2	16	0,125	1

Např. ve 25 % domácností mají pouze jednu televizi a v 75 % domácností jsou maximálně 2 televize

Histogram:



Postup s využitím nástroje Analýza dat v **EXCEL**:

1. do sloupce zadáme horní hranice tříd (v případě prostého třídění přímo hodnoty kategorií - 0,1,2,3,4)
2. Data → Analýza dat → Histogram
3. Vstupní oblat - data; Hranice tříd - Horní hranice tříd; Výstupní oblast - libovolná buňka
4. Zaškrtneme volbu Kumulativní procentuální podíl, Vytvořit graf

Poznámky

Absolutní četnost

f_i ... počet prvků patřící do i -té třídy

Relativní četnost

$\varphi_i = \frac{f_i}{n}$, n je rozsah souboru

Kumulativní absolutní četnost

$$F_i = \sum_{l=1}^i f_l$$

Kumulativní relativní četnost

$\phi_i = \frac{F_i}{n}$, n je rozsah souboru

EXCEL:

=ČETNOSTI(data;horní hranice tříd)

- vzorec je maticový

- tzn. musí se označit počet buněk odpovídající počtu tříd, pak zadáme vzorec, stiskneme F2 a poté CTRL+Shift+Enter

140 - Charakteristiky polohy

Zadání Provádíme výzkum o počtu televizí v domácnostech. Náhodně jsme oslovili 16 domácností a zjistili následující údaje. Na základě charakteristik polohy proveďte závěr o počtu televizí v domácnostech.

2	1	0	1	2	4	3	1	2	4	2	1	0	2	3	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Řešení

Teorie: 51 Příklady: 216,218,219,220

Aritmetický průměr

$$\bar{x} = \frac{1}{16} (2 + 1 + 0 + 1 + 2 + 4 + 3 + 1 + 2 + 4 + 2 + 1 + 0 + 2 + 3 + 2) = \frac{30}{16} = 1,88$$

Modus

- z tabulky četností je zřejmé, že největší absolutní četnost je 6 a odpovídá hodnotě $2 \Rightarrow \hat{x} = 2$

Kvartily

1. hodnoty seřadíme: 0,0,1,1,1,1,2,2,2,2,2,3,3,4,4
2. vypočteme pořadí hodnoty reprezentující příslušný kvartil:

$$z_{0,25} = 16 \cdot 0,25 + 0,5 = 4,5 \quad z_{0,5} = 16 \cdot 0,5 + 0,5 = 8,5 \quad z_{0,75} = 16 \cdot 0,75 + 0,5 = 12,5$$

3. určíme hodnoty kvartilů:

$$x_{0,25} = \frac{1+1}{2} = 1 \quad \tilde{x} = \frac{2+2}{2} = 2 \quad x_{0,75} = \frac{2+3}{2} = 2,5$$

Průměrný počet televizí v domácnosti je 1,88, nejčastěji mají lidé 2 televize. Ve čtvrtině domácnostech mají maximálně jednu televizi.

Poznámky

Výběrový průměr

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$$

EXCEL: =PRŮMĚR(číslo1;číslo2;...)

Modus

EXCEL: =MODE(číslo1;číslo2;...)

Kvantily

EXCEL: =PERCENTIL(oblast;p)

EXCEL: =MEDIAN(číslo1;číslo2;...)

EXCEL: =QUARTIL.EXC(matice;1)

EXCEL: =QUARTIL.EXC(matice;3)

Využití analytického nástroje Popisná statistika v Excelu.

1. Data → Analýza dat → Popisná statistika
2. Vstupní oblast - data; Výstupní oblast - libovolná buňka
3. Zaškrtneme volbu Celkový přehled

141 - Charakteristiky variability

Zadání Provádíme výzkum o počtu televizí v domácnostech. Náhodně jsme oslovili 16 domácností a zjistili následující údaje. Určete základní charakteristiky variability.

2	1	0	1	2	4	3	1	2	4	2	1	0	2	3	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Řešení **Teorie: 52 Příklady: 216-220**

Variační rozpětí

$$R = 4 - 0 = 4$$

Interkvartilové rozpětí

$$IQR = 2,5 - 1 = 1,5$$

Výběrový rozptyl

$$s^2 = \frac{1}{15} \left((2 - 1,88)^2 + (1 - 1,88)^2 + \dots + (2 - 1,88)^2 \right) = \frac{21,75}{15} = 1,45$$

Výběrová směrodatná odchylka

$$s = \sqrt{1,45} = 1,2$$

Variační koeficient

$$V_x = \frac{1,2}{1,88} = 0,64$$

Poznámky

Výběrový rozptyl

$$s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

EXCEL: =VAR.S(číslo1;číslo2;...)

Výběrová směrodatná odchylka

$$s = \sqrt{s^2}$$

EXCEL: =SMODCH.VÝBĚR.S(číslo1;číslo2;...)

Variační koeficient

$$V_x = \frac{s}{\bar{x}}$$

Využití analytického nástroje Popisná statistika v Excelu.

1. Data → Analýza dat → Popisná statistika
2. Vstupní oblat - data; Výstupní oblast - libovolná buňka
3. Zaškrtneme volbu Celkový přehled

142 - Charakteristiky tvaru

Zadání Provádíme výzkum o počtu televizí v domácnostech. Náhodně jsme oslovili 16 domácností a zjistili následující údaje. Určete základní charakteristiky tvaru rozdělení.

2	1	0	1	2	4	3	1	2	4	2	1	0	2	3	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Řešení **Teorie: 53 Příklady: 217,218,219,220**

Šikmost

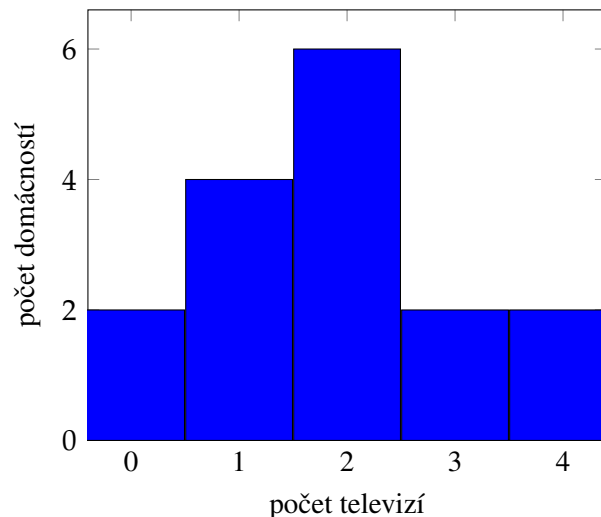
$$A = \frac{16}{15 \cdot 14 \cdot (1,2)^3} \left((2 - 1,88)^3 + (1 - 1,88)^3 + \dots + (2 - 1,88)^3 \right) = \frac{93,78}{362,88} = 0,26$$

Šikmost vyšla kladná ($A > 0$) $\Rightarrow$ ve většině domácností je méně televizí než je průměrná hodnota (1,88).

Špičatost

$$\tilde{e} = \frac{16 \cdot 17}{15 \cdot 14 \cdot 13 \cdot (1,2)^4} \left((2 - 1,88)^4 + (1 - 1,88)^4 + \dots + (2 - 1,88)^4 \right) - \frac{3 \cdot 15^2}{14 \cdot 13} = -0,348$$

Špičatost vyšla záporná ($\tilde{e} < 0$) $\Rightarrow$ jedná se o ploché rozdělení, lze říci, že v domácnostech se vyskytují všechny varianty (0-4).



Poznámky

Šikmost

$$A = \frac{n}{(n-1)(n-2)s^3} \sum_{j=1}^n (x_j - \bar{x})^3$$

EXCEL: =SKEW(číslo1;číslo2;...)

Špičatost - exces

$$\tilde{e} = \frac{n(n+1)}{(n-1)(n-2)(n-3)s^4} \sum_{j=1}^n (x_j - \bar{x})^4 - 3 \frac{(n-1)^2}{(n-2)(n-3)}$$

EXCEL: =KURT(číslo1;číslo2;...)

Využití analytického nástroje Popisná statistika v Excelu.

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Popisná statistika
2. Vstupní oblast - data; Výstupní oblast - libovolná buňka
3. Zaškrtneme volbu Celkový přehled

Řešené příklady – Odhady parametrů

144 - Interval spolehlivosti pro střední hodnotu

Zadání Experimentem byly získány doby obsluhy v systému hromadné obsluhy. Určili jsme výběrové charakteristiky:

$$n = 100, \bar{x} = 20,52, s = 7,76.$$

Určete 95% interval spolehlivosti pro odhad střední doby obsluhy.

Řešení

Teorie: 56,58 Příklady: 222,223

Hledáme 95% oboustranný interval spolehlivosti pro střední hodnotu při neznámém rozptylu $\Rightarrow$ použijeme t-rozdělení.

Potřebujeme znát hodnotu výběrového průměru, směrodatné odchylky a rozsah výběru:
ze zadání: $n = 100$; $\bar{x} = 20,52$; $s = 7,76$.

Určíme hodnotu $(1 - \frac{\alpha}{2})$ -tého kvantilu t-rozdělení s $(n - 1)$ stupni volnosti:
 $\alpha = 0,05 \Rightarrow t_{0,975}(99) = (T.INV(0,975;99)) = 1,98$.

Vypočítáme hodnotu δ :

$$\delta = \frac{7,76}{\sqrt{100}} \cdot 1,98 = 1,54.$$

Meze IS:

$$t_D = \bar{x} - \delta = 20,52 - 1,54 = 18,98,$$

$$t_H = \bar{x} + \delta = 20,52 + 1,54 = 22,06.$$

Závěr: S pravděpodobností 95% leží střední doba obsluhy v intervalu $\langle 18,98; 22,06 \rangle$.

Poznámky

Interval spolehlivosti pro μ

$$\mu \in \langle \bar{x} \pm \delta \rangle$$

1. známe σ^2

$$\delta = \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}$$

2. neznáme σ^2

$$\delta = \frac{s}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2}}(n-1)$$

EXCEL:

kvantil normovaného normálního rozdělení $z_{1-\frac{\alpha}{2}}$

$$=NORM.S.INV\left(1 - \frac{\alpha}{2}\right)$$

kvantil t-rozdělení s $n - 1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n - 1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

Využití analytického nástroje Popisná analýza v Excelu.

- určení hodnoty δ z výběrového souboru, neznám σ^2

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Popisná analýza

2. Vstupní oblat - data; Výstupní oblast - libovolná buňka

3. Zaškrtneme Hladina spolehlivosti pro stř. hodnotu a zvolíme hladinu spolehlivosti $(1 - \alpha) \cdot 100$

145 - Interval spolehlivosti pro střední hodnotu

Zadání Na základě provedeného měření průřezové plochy [mm²] Ap předpínacího lana Ls15.7

148,6	148	148,3	148,2	148,3	148,5	148,9	148,5	148,4	148,7	148,6	148,5
-------	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

a) stanovte 90% interval spolehlivosti pro odhad střední hodnoty,

b) určete na 10% hladině významnosti pravostranný interval spolehlivosti pro odhad střední hodnoty.

Řešení

Teorie: 56,58 Příklady: 222,223

a) hledáme 90% oboustranný interval spolehlivosti pro střední hodnotu při neznámém rozptylu. Jelikož máme k dispozici výběrový soubor, využijeme k nalezení výběrového průměru a hodnoty δ nástroje Popisná statistika v Analýze dat v Excelu:

$$\bar{x} = 148,46; \delta = 0,12$$

Meze IS:

$$t_D = \bar{x} - \delta = 148,46 - 0,12 = 148,34,$$

$$t_H = \bar{x} + \delta = 148,46 + 0,12 = 148,58.$$

Závěr: S pravděpodobností 90% leží střední hodnota velikosti průřezové plochy v intervalu $\langle 148,34; 148,58 \rangle$.

b) potřebujeme určit pouze horní mez, proto hledáme 90% jednostranný interval spolehlivosti pro střední hodnotu (musíme použít kvantil pro jednostranný test, tzn. $(1 - \alpha)$ kvantil t-rozdělení). Opět k nalezení výběrového průměru a hodnoty δ použijeme nástroje Popisné statistiky v Analýze dat, hodnota výběrového průměru je stejná jako za a), ale v případě výpočtu δ zvolíme: Hladina spolehlivosti pro stř. hodnotu $(1 - 2\alpha) \cdot 100$, tzn. zadáme číslo 80:

$$\bar{x} = 148,46; \delta = 0,09$$

Horní mez IS:

$$t_H = \bar{x} + \delta = 148,46 + 0,09 = 148,55.$$

Závěr: S 90% spolehlivostí velikost průřezové plochy nepřekročí hodnotu 148,55mm², tzn. 90% jednostranný IS: $\mu \in (-\infty; 148,55)$.

Poznámky

Interval spolehlivosti pro μ

$$\mu \in \langle \bar{x} \pm \delta \rangle$$

1. známe σ^2

$$\delta = \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}$$

2. neznáme σ^2

$$\delta = \frac{s}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2}}(n-1)$$

EXCEL:

kvantil normovaného normálního rozdělení $z_{1-\frac{\alpha}{2}}$

$$= \text{NORM.S.INV}\left(1 - \frac{\alpha}{2}\right)$$

kvantil t-rozdělení s $n - 1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$= \text{T.INV}\left(1 - \frac{\alpha}{2}; n - 1\right)$$

Využití analytického nástroje Popisná analýza v Excelu.

- určení hodnoty δ z výběrového souboru, neznám σ^2

1. Data → Analýza dat → Popisná analýza

2. Vstupní oblat - data; Výstupní oblast - libovolná buňka

3. Zaškrtneme Hladina spolehlivosti pro stř. hodnotu a zvolíme hladinu spolehlivosti $(1 - \alpha) \cdot 100$

146 - Interval spolehlivosti pro střední hodnotu

Zadání Na základě provedeného měření průřezové plochy [mm²] Ap předpínacího lana Ls15.7 určete meze, ve kterých lze s 99% spolehlivostí očekávat velikost plochy, pokud z dřívějších analýz víme, že hodnota směrodatné odchylky je 0,28.

148,6	148	148,3	148,2	148,3	148,5	148,9	148,5	148,4	148,7	148,6	148,5
-------	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Řešení

Teorie: 56,57 Příklady: 222

Hledáme 99% oboustranný interval spolehlivosti pro střední hodnotu při známém rozptylu $\Rightarrow$ budeme používat normované normální rozdělení.

Určíme hodnotu výběrového průměru a rozsah výběru:

$$n = 12; \bar{x} = 148,46.$$

Určíme hodnotu $(1 - \frac{\alpha}{2})$ -tého kvantilu normovaného normálního rozdělení:

$$\alpha = 0,01 \Rightarrow z_{0,995} = (\text{NORM.S.INV}(0,995)) = 2,58.$$

Vypočítáme hodnotu δ :

$$\delta = \frac{0,28}{\sqrt{12}} \cdot 2,58 = 0,21.$$

Meze IS:

$$t_D = \bar{x} - \delta = 148,46 - 0,21 = 148,25,$$

$$t_H = \bar{x} + \delta = 148,46 + 0,21 = 148,67.$$

Závěr: S pravděpodobností 99% leží střední velikost plochy v intervalu $\langle 148,25; 148,67 \rangle$.

Poznámky

Interval spolehlivosti pro μ

$$\mu \in \langle \bar{x} \pm \delta \rangle$$

1. znám σ^2

$$\delta = \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}$$

2. neznám σ^2

$$\delta = \frac{s}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2}}(n-1)$$

EXCEL:

kvantil normovaného normálního rozdělení $z_{1-\frac{\alpha}{2}}$

$$=\text{NORM.S.INV}\left(1 - \frac{\alpha}{2}\right)$$

kvantil t -rozdělení s $n-1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=\text{T.INV}\left(1 - \frac{\alpha}{2}; n-1\right)$$

Využití analytického nástroje Popisná analýza v Excelu.

- určení hodnoty δ z výběrového souboru, neznám σ^2

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Popisná analýza

2. Vstupní oblat - data; Výstupní oblast - libovolná buňka

3. Zaškrtneme Hladina spolehlivosti pro stř. hodnotu a zvolíme hladinu spolehlivosti $(1 - \alpha) \cdot 100$

147 - Interval spolehlivosti pro rozptyl, směrodatnou odchylku

Zadání Zajímá nás objemová tíha betonu [kN/m³]. Náhodně jsme provedli 10 měření:

18	20	22	30	28	25	31	21	26	30
----	----	----	----	----	----	----	----	----	----

Určete interval spolehlivosti pro rozptyl a směrodatnou odchylku.

Řešení **Teorie: 56,58 Příklady: 224**

Zvolíme hladinu významnosti: $\alpha = 0,05$ a určíme 95% oboustranný interval spolehlivosti pro rozptyl a směrodatnou odchylku.

Určíme rozsah výběru a hodnotu výběrového rozptylu (=VAR.S(data)): $n = 10$; $s^2 = 21,66$.

S využitím funkcí v Excelu vypočteme hodnoty obou kvantilů

1. $(1 - \frac{\alpha}{2})$ kvantil chí-kvadrát rozdělení: $\chi^2_{0,975}(9) = (\text{CHISQ.INV}(0,975;9)) = 19,02$.

2. $\frac{\alpha}{2}$ kvantil chí-kvadrát rozdělení: $\chi^2_{0,025}(9) = (\text{CHISQ.INV}(0,025;9)) = 2,70$.

Meze IS pro rozptyl:

$$t_D = \frac{9 \cdot 21,66}{19,02} = 10,25,$$

$$t_H = \frac{9 \cdot 21,66}{2,70} = 72,17.$$

Závěr: 95 % IS: $\sigma^2 \in \langle 10,25; 72,17 \rangle$.

Meze IS pro směrodatnou odchylku:

$$t_D = \sqrt{10,25} = 3,20,$$

$$t_H = \sqrt{72,17} = 8,50.$$

Závěr: 95% IS: $\sigma \in \langle 3,20; 8,50 \rangle$.

Poznámky

Interval spolehlivosti pro σ^2

$$\sigma^2 \in \left\langle \frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}; \frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right\rangle$$

Interval spolehlivosti pro σ

$$\sigma \in \left\langle \sqrt{\frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}; \sqrt{\frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} \right\rangle$$

EXCEL:

kvantil chí-kvadrát rozdělení s $n - 1$ stupni volnosti

$$1. \chi^2_{1-\frac{\alpha}{2}}(n-1) \\ = \text{CHISQ.INV}\left(1 - \frac{\alpha}{2}; n-1\right)$$

$$2. \chi^2_{\frac{\alpha}{2}}(n-1) \\ = \text{CHISQ.INV}\left(\frac{\alpha}{2}; n-1\right)$$

Řešené příklady – Testování hypotéz

149 - Hypotéza o střední hodnotě

Zadání Měřením součinitele tření f v ložisku byly získány následující údaje

0,0148	0,0124	0,0102	0,0085	0,0071	0,0059	0,0051
--------	--------	--------	--------	--------	--------	--------

Ověřte s 95% spolehlivostí, že průměrná hodnota součinitele tření je 0,01.

Řešení

Teorie: 61 Příklady: 226

Zvolíme nulovou a alternativní hypotézu:

$$H_0 : \mu = 0,01,$$

$$H_1 : \mu \neq 0,01.$$

Pro výpočet testové statistiky potřebujeme určit hodnotu výběrového průměru, směrodatné odchylky a rozsah výběru (využijeme nástroje Popisná statistika v Analýze dat v Excelu):

$$n = 7; \bar{x} = 0,0091; s = 0,0035$$

$$T = \frac{0,0091 - 0,01}{0,0035} \cdot \sqrt{7} = -0,68$$

Určíme kritickou hodnotu:

$$\alpha = 0,05 \Rightarrow t_{0,975}(6) = (\text{T.INV}(0,975;6)) = 2,45.$$

Závěr: $|-0,68| < 2,45 \Rightarrow H_0$ nezamítáme a lze s pravděpodobností 95 % tvrdit, že průměrná hodnota součinitele tření f v ložisku je 0,01.

Poznámky

Jednovýběrový t -test

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

výpočet testové hodnoty

$$T = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}$$

výpočet kritické hodnoty

$$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n-1)$$

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

EXCEL:

kvantil t -rozdělení s $n-1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=\text{T.INV}\left(1 - \frac{\alpha}{2}; n - 1\right)$$

150 - Hypotéza o rozptylu

Zadání Přesnost nastavení stroje se určí z rozptylu délky vyráběných součástek. Pokud je rozptyl větší než 380, je potřeba stroj znovu nastavit. Z 15 náhodně vybraných součástek jsme zjistili hodnotu $s^2 = 680$. Je stroj dostatečně přesný, nebo je potřeba stroj přenastavit?

Řešení

Teorie: 62 Příklady: 227

Zvolíme nulovou hypotézu, alternativní hypotézu a hladinu významnosti:

$$H_0 : \sigma^2 = 380,$$

$$H_1 : \sigma^2 > 380,$$

$$\alpha = 0,05.$$

Výpočet testové statistiky:

$$\chi^2 = \frac{14 \cdot 680}{380} = 25,05.$$

Určíme kritickou hodnotu $\left(\chi_{1-\alpha}^2(n-1)\right)$:

$$\chi_{\text{krit}}^2 = 23,68.$$

Závěr: $25,05 > 23,68 \Rightarrow H_0$ zamítáme, stroj je potřeba znovu nastavit.

Poznámky

Test o rozptylu

$$H_0 : \sigma^2 = \sigma_0^2$$

$$H_1 : \sigma^2 \neq \sigma_0^2 (\sigma^2 > \sigma_0^2; \sigma^2 < \sigma_0^2)$$

výpočet testové hodnoty

$$\chi^2 = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

výpočet kritických hodnot

oboustranný test:

$$\chi_{\frac{\alpha}{2}}^2(n-1); \chi_{1-\frac{\alpha}{2}}^2(n-1)$$

jednostranný test:

$$\chi_{\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 < \sigma_0^2$$

$$\chi_{1-\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 > \sigma_0^2$$

Závěr:

oboustranný test:

$$\chi^2 \in \left\langle \chi_{\frac{\alpha}{2}}^2(n-1); \chi_{1-\frac{\alpha}{2}}^2(n-1) \right\rangle \Rightarrow H_0 \text{ nezamítáme}$$

jednostranný test:

$$\chi^2 > \chi_{\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 < \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

$$\chi^2 < \chi_{1-\alpha}^2(n-1) \text{ pro } H_1 : \sigma^2 > \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil Chí-kvadrát rozdělení s $(n-1)$ stupni volnosti

$$\chi_{1-\frac{\alpha}{2}}^2(n-1)$$

$$=\text{CHISQ.INV}\left(1 - \frac{\alpha}{2}; n-1\right)$$

151 - Hypotéza o rovnosti rozptylů

Zadání Chceme porovnat dva typy betonu z hlediska součinitele tepelné vodivosti λ [W/mK]. Jsou srovnatelné rozptyly u obou typů betonů?

Beton z keramzitu	0,28	0,31	0,34	0,40	0,48	0,56	0,63	0,75	1,30		
Beton ze škváry	0,52	0,54	0,67	0,69	0,73	0,74	0,79	0,82	0,90	0,97	1,01

Řešení

Teorie: 63 Příklady: 228

Zvolíme nulovou hypotézu, alternativní hypotézu a hladinu významnosti:

$$H_0 : \sigma_1^2 = \sigma_2^2,$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2,$$

$$\alpha = 0,05.$$

Pro výpočet testové statistiky využijeme nástroje v Analýze dat v Excelu (hodnota u F):

$$F = 4,03.$$

Určíme kritickou hodnotu (F krit):

$$F_{\text{krit}} = 3,85.$$

Závěr: $4,03 > 3,85 \Rightarrow H_0$ zamítáme a s 95% pravděpodobností můžeme tvrdit, že variabilita součinitele tepelné vodivosti se u vybraných typů betonu liší.

Poznámky

Dvouvýběrový F-test

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

výpočet testové hodnoty

$$F = \frac{s_1^2}{s_2^2}$$

Indexy 1,2 volíme, aby $F > 1$.

výpočet kritické hodnoty

$$F_{\text{krit}} = F_{1-\frac{\alpha}{2}}(n_1 - 1; n_2 - 1)$$

Závěr: $F < F_{\text{krit}} \Rightarrow H_0$ nezamítáme

EXCEL:

kvantil F-rozdělení s $(n_1 - 1; n_2 - 1)$ stupni volnosti

$$=F.INV\left(1 - \frac{\alpha}{2}; n_1 - 1; n_2 - 1\right)$$

Využití analytického nástroje v Excelu.

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Dvouvýběrový F-test pro rozptyl

2. Vstup - data (1. soubor zvolíme ten s větším rozptylem); Výstupní oblast - libovolná buňka

3. Alfa: zadáme $\frac{\alpha}{2}$

152 - Hypotéza o rovnosti středních hodnot

Zadání Rozhodněte na hladině významnosti 0,05, že hodnota součinitele tepelné vodivosti [W/mK] u betonu ze škváry je větší než u betonu z keramzitu.

Beton z keramzitu	0,28	0,31	0,34	0,40	0,48	0,56	0,63	0,75	1,30		
Beton ze škváry	0,52	0,54	0,67	0,69	0,73	0,74	0,79	0,82	0,90	0,97	1,01

Řešení

Teorie: 64 Příklady: 228

Zvolíme nulovou a alternativní hypotézu:

$$H_0 : \mu_1 = \mu_2,$$

$$H_1 : \mu_1 < \mu_2.$$

První musíme otestovat, zda lze považovat rozptyly u obou souborů za srovnatelné (F-test). Tento test jsme provedli na 151, zjistili jsme, že rozptyly nejsou srovnatelné $\Rightarrow$ použijeme t-test s nerovností rozptylů.

Pro výpočet testové statistiky a kritické hodnoty využijeme nástroje v Analýze dat v Excelu (hodnota u t Stat a t krit (1)):

$$T = -1,73,$$

$$t_{\text{krit}} = 1,80.$$

Závěr: $|1,73| < 1,80 \Rightarrow H_0$ nezamítáme, tedy s 95% pravděpodobností lze tvrdit, že oba druhy betonu jsou srovnatelné z hlediska hodnoty součinitele tepelné vodivosti.

Poznámky

Dvouvýběrový t-test

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

výpočet testové hodnoty a kritické hodnoty

1. platí $\sigma_1^2 = \sigma_2^2$, použijeme t-test s rovností rozptylů

2. neplatí $\sigma_1^2 = \sigma_2^2$, použijeme t-test s nerovností rozptylů

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

EXCEL:

kvantil t-rozdělení s $n_1 + n_2 - 2$ stupni volnosti
 $t_{1-\alpha}(n_1 + n_2 - 2)$
 =T.INV($1 - \alpha; n_1 + n_2 - 2$)

Využití analytického nástroje v Excelu.

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Dvouvýběrový t-test s rovností (nerovností) rozptylů
2. Vstup - data; Výstupní oblast - libovolná buňka
3. Alfa: zadáme hladinu významnosti

153 - Párový t-test

Zadání Při určování mechanických vlastností konstrukční oceli (teplota 20 °C) jsme provedli orientační přepočet tvrdosti k hodnotám pevnosti v tahu. Přepočet jsme provedli podle Vickerse a podle Brinella. Dávají oba postupy srovnatelné hodnoty tvrdosti?

Pevnost v tahu	350	370	385	400	415	430	450	465	480	495	510	530
Tvrдост dle Brinella	105	109	114	119	124	128	133	138	143	147	152	156
Tvrдост dle Vickerse	110	115	120	125	130	135	140	145	150	155	160	165

Řešení

Teorie: 66 Příklady: 229

Zvolíme nulovou a alternativní hypotézu:

$$H_0 : \mu_1 = \mu_2,$$

$$H_1 : \mu_1 \neq \mu_2.$$

Pro výpočet testové statistiky využijeme nástroje v Analýze dat v Excelu (hodnota u t Stat):

$$T = 21, 23.$$

Určíme kritickou hodnotu (t krit (2)):

$$t_{\text{krit}} = 2, 20.$$

Závěr: $|21, 23| > 2, 20 \Rightarrow H_0$ zamítáme a lze s 95% pravděpodobností tvrdit, že vypočtená hodnota tvrdosti závisí na použitém postupu.

Poznámky

Párový t-test

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

výpočet testové hodnoty

$$t = \frac{|\bar{x}_d|}{s_d} \cdot \sqrt{n},$$

kde $\bar{x}_d$, s_d jsou výběrové charakteristiky souboru tvořených z rozdílů párových hodnot.

výpočet kritické hodnoty

$$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n-1)$$

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

EXCEL:

kvantil t-rozdělení s $n-1$ stupni volnosti $t_{1-\frac{\alpha}{2}}(n-1)$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

Využití analytického nástroje v Excelu.

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Dvouvýběrový párový t-test na střední hodnotu
2. Vstup - data; Výstupní oblast - libovolná buňka
3. Alfa: zadáme hladinu významnosti

154 - Testy dobré shody - část 1.

Zadání Otestujte hypotézu, že náhodné množství kyseliny benzoové v balíčcích nakládacího přípravku je rozloženo normálně, tedy zkoumaná náhodná veličina má normální rozdělení pravděpodobnosti. Změřili jsme obsah ve 150 balíčcích.

Třída	$(0,3;0,5)$	$(0,5;0,7)$	$(0,7;0,9)$	$(0,9;1,1)$	$(1,1;1,3)$	$(1,3;1,5)$
Četnost	5	10	30	55	40	10

Řešení

Teorie: 67,68 Příklady: 230

Zvolíme nulovou a alternativní hypotézu:

H_0 : náhodná veličina má normální rozdělení,

H_1 : $\neg H_0$.

1) χ^2 – test

Odhadneme parametry normálního rozdělení $\mu = \bar{x}$; $\sigma = s$:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k f_{ei} \cdot x_i = \frac{1}{150} \cdot (5 \cdot 0,4 + 10 \cdot 0,6 + 30 \cdot 0,8 + 55 \cdot 1 + 40 \cdot 1,2 + 10 \cdot 1,4) = 0,99$$

$$s^2 = \frac{1}{n-1} \left(\sum_{i=1}^k (x_i - \bar{x})^2 \cdot f_{ei} \right) = \frac{1}{149} \cdot ((0,4 - 0,99)^2 \cdot 5 + (0,6 - 0,99)^2 \cdot 10 + \dots + (1,4 - 0,99)^2 \cdot 10) = 0,052$$

$$s = \sqrt{s^2} = 0,23$$

Určíme očekávané četnosti:

$$f_{o1} = 150 \cdot F(0,5) = (150 * \text{NORM.DIST}(0,5;0,99;0,23;1)) = 2,49$$

$$f_{o2} = 150 \cdot (F(0,7) - F(0,5)) = (150 * (\text{NORM.DIST}(0,7;0,99;0,23;1) - \text{NORM.DIST}(0,5;0,99;0,23;1))) = 13,07$$

Stejně pokračujeme pro všechny třídy, výsledky dáme do tabulky.

Poznámky

χ^2 – test

Nulová hypotéza:

ZS má očekávané rozdělení pravděpodobnosti, tj. četnosti f_{ei} a f_{oi} pro $i = 1, \dots, n$ se liší pouze náhodně.

výpočet testové hodnoty

$$\chi^2 = \sum_{i=1}^k \frac{(f_{ei} - f_{oi})^2}{f_{oi}}$$

výpočet kritické hodnoty

$$t_{\text{krit}} = \chi_{1-\alpha}^2(k - s - 1)$$

Závěr: $\chi^2 < \chi_{\text{krit}}^2 \Rightarrow H_0$ nezamítáme

155 - Testy dobré shody - část 2.

Zadání Pokračujeme v řešení z předcházejícího listu (154).**Řešení****Teorie: 67,68 Příklady: 230**

f_{ei}	5	10	30	55	40	10
f_{oi}	2,49	13,07	36,62	50,40	34,11	11,33

Vypočteme testovou statistiku:

$$\chi^2 = \frac{(5 - 2,49)^2}{2,49} + \dots + \frac{(10 - 11,33)^2}{11,33} = 6,06.$$

Určíme kritickou hodnotu ($\chi^2(k - s - 1)$):

$$\chi_{\text{krit}}^2 = \chi_{1-0,05}^2(6 - 2 - 1) = (\text{CHISQ.INV}(0,95;3)) = 7,81.$$

Závěr: $6,06 < 7,81 \Rightarrow H_0$ nezamítáme a lze s 95% pravděpodobností tvrdit, že rozdělení obsahu kyseliny v balíčcích se statisticky neliší od normálního.

2) Kolmogorovův-Smirnovův test dobré shody

Tabulku s teoretickými a očekávanými absolutními četnostmi doplníme o kumulativní četnosti.

F_{ei}	5	15	45	100	140	150
F_{oi}	2,49	15,56	52,18	102,58	136,69	148,02

Vypočteme testovou statistiku:

$$D_1 = \frac{1}{150} \cdot |-7,17| = 0,048.$$

Určíme kritickou hodnotu:

$$D_{1;0,05}(150) = \frac{1,36}{\sqrt{150}} = 0,11.$$

Závěr: $0,048 < 0,11 \Rightarrow H_0$ nezamítáme a lze s 95% pravděpodobností tvrdit, že rozdělení obsahu kyseliny v balíčcích se statisticky neliší od normálního.**Poznámky***Nulová hypotéza:*ZS má očekávané rozdělení pravděpodobnosti, tj. četnosti f_{ei} a f_{oi} pro $i = 1, \dots, n$ se liší pouze náhodně. χ^2 - test

výpočet testové hodnoty

$$\chi^2 = \sum_{i=1}^k \frac{(f_{ei} - f_{oi})^2}{f_{oi}}$$

výpočet kritické hodnoty

$$\chi_{\text{krit}}^2 = \chi_{1-\alpha}^2(k - s - 1)$$

EXCEL:kvantil chí-kvadrát rozdělení
s $k - s - 1$ stupni volnosti
=CHISQ.INV($1 - \alpha; k - s - 1$)Závěr: $\chi^2 < \chi_{\text{krit}}^2 \Rightarrow H_0$ nezamítáme*Kolmogorovův-Smirnovův test*

výpočet testové hodnoty

$$D_1 = \frac{1}{n} \max |F_{ei} - F_{oi}|$$

výpočet kritické hodnoty

$$\alpha = 0,05 \Rightarrow D_{1;0,05}(n) = \frac{1,36}{\sqrt{n}}$$

Závěr: $D_1 < D_{1;\alpha}(n) \Rightarrow H_0$ nezamítáme

156 - Testy extrémních hodnot

Zadání Při rozboru křemičitanu byl nalezen tento obsah SiO₂: 52,44 %; 53,82 %; 52,91 %; 50,10 %; 54,03 %; 53,89 %. Je hodnota 50,10 % odlehlá na hladině významnosti $\alpha = 0,05$?

Řešení

Teorie: 69,70 Příklady: 231

Zvolíme nulovou hypotézu:

H_0 : Hodnota $x_1 = 50,10$ se neliší významně od ostatních hodnot souboru.

Hodnoty seřadíme: 50,10; 52,44; 52,91; 53,82; 53,89; 54,03.

1) Dixonův test

Vypočteme testovou statistiku:

$$Q_1 = \frac{52,44 - 50,10}{54,03 - 50,10} = 0,595.$$

Určíme kritickou hodnotu : Z tabulek pro $n = 6$ a $\alpha = 0,05$ dostáváme

$$Q_{1;0,05} = 0,560.$$

Závěr: $0,595 > 0,56 \Rightarrow H_0$ zamítáme a lze s 95% spolehlivostí tvrdit, že hodnota 50,10 % je odlehlou.

2) Grubbsův test

Určíme hodnotu výběrového průměru a směrodatné odchylky:

$$\bar{x} = 52,865, s = 1,363.$$

Vypočteme testovou statistiku:

$$T_1 = \frac{52,865 - 50,10}{1,363} = 2,029.$$

Určíme kritickou hodnotu : Z tabulek pro $n = 6$ a $\alpha = 0,05$ dostáváme

$$T_{1;0,05} = 1,996.$$

Závěr: $2,029 > 1,996 \Rightarrow H_0$ zamítáme a i podle Dixonova testu lze s 95% pravděpodobností tvrdit, že hodnota 50,10 % je odlehlou hodnotou.

Poznámky

Nulová hypotéza:

H_0 : Hodnota x_1 , resp. x_n (největší hodnota) se neliší významně od ostatních hodnot souboru.

Dixonův test

výpočet testové hodnoty

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1}$$

$$Q_n = \frac{x_n - x_{n-1}}{x_n - x_1}$$

výpočet kritické hodnoty

Kritické hodnoty $Q_{1;\alpha}$, resp. $Q_{n;\alpha}$ jsou tabelovány.

Závěr:

$Q_1 < Q_{1;\alpha} \Rightarrow H_0$ nezamítáme

$Q_n < Q_{n;\alpha} \Rightarrow H_0$ nezamítáme

Grubbsův test

výpočet testové hodnoty

$$T_1 = \frac{\bar{x} - x_1}{s}$$

$$T_n = \frac{x_n - \bar{x}}{s}$$

výpočet kritické hodnoty

Kritické hodnoty $T_{1;\alpha}$, resp. $T_{n;\alpha}$ jsou tabelovány.

Závěr:

$T_1 < T_{1;\alpha} \Rightarrow H_0$ nezamítáme

$T_n < T_{n;\alpha} \Rightarrow H_0$ nezamítáme

Řešené příklady – Lineární regrese

158 - Regresní model

Zadání Známe výsledky dvou testů u 8 studentů,

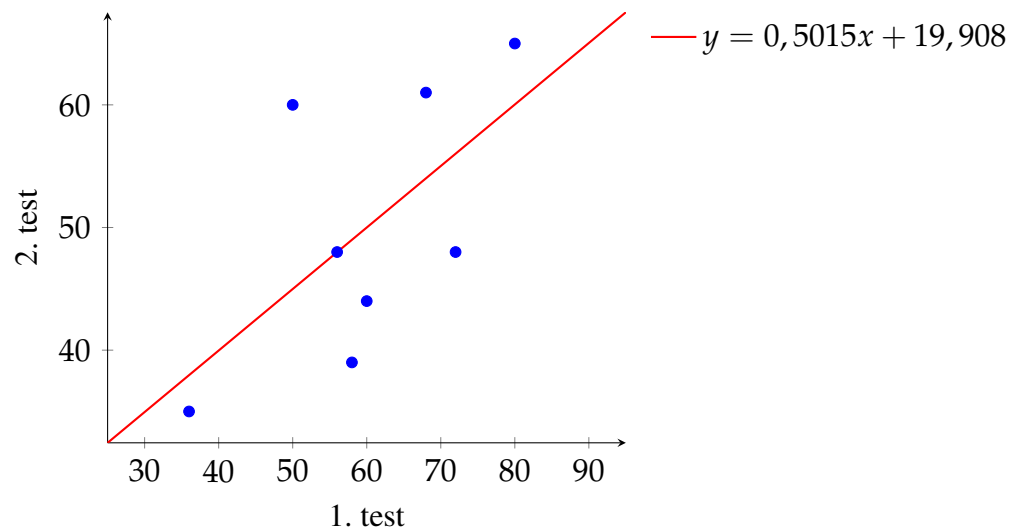
1. test	80	50	36	58	72	60	56	68
2. test	65	60	35	39	48	44	48	61

odhadněte parametry regresní přímky a určete odhad pro počet bodů z druhého testu u studenta, který dosáhl z 1. testu 90 bodů.

Řešení **Teorie: 74,75 Příklady: 233,235,236**

Proměnné vyneseme do bodového grafu, závislou (body z 2. testu) na osu y a nezávislou (body z 1. testu) na osu x .

S využitím metody nejmenších čtverců odhadneme regresní přímku: $y = 0,5x + 19,9$.



Odhad pro počet bodů z druhého testu:

$$\text{počet bodů z 2. testu} = 0,5015 \cdot \text{počet bodů z 1. testu} + 19,908$$

$$y = 0,5015 \cdot 90 + 19,908 = 65,043 \Rightarrow 65 \text{ bodů}$$

Poznámky

Regresní přímka

$$y = ax + b$$

Soustava normálních rovnic

$$a \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i$$

$$a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i$$

EXCEL:

Grafické znázornění + rovnice regresní přímky
Vložení -> Bodový graf -> Přidat spojnicí trendu

zvolíme Lineární

zaškrtneme Zobrazit rovnici v grafu

Nalezení koeficientů

=LINREGRESE(data y;data x)

- vzorec je maticový

- tzn. musíme označit 2 buňky vedle sebe, pak zadáme vzorec, stiskneme F2 a poté CTRL+Shift+Enter

Využití analytického nástroje Regrese v Excelu.

1. Data -> Analýza dat -> Regrese

2. Vstupní oblast Y - závislá proměnná; Vstupní oblast X - nezávislá proměnná; Výstupní oblast - libovolná buňka

159 - Regresní model - verifikace modelu

Zadání Pokračujeme v příkladu se studenty,

1. test	80	50	36	58	72	60	56	68
2. test	65	60	35	39	48	44	48	61

posuďte kvalitu nalezeného modelu.

Řešení **Teorie: 76 Příklady: 233,234,235,236**

$$S_R^2 = \sum_{i=1}^8 (\tilde{y}_i - 50)^2 = 328,003$$

$$SSE = \sum_{i=1}^8 (y_i - \tilde{y}_i)^2 = 507,997$$

$$S_T^2 = \sum_{i=1}^8 (y_i - 50)^2 = 836$$

Hodnota koeficientu determinace je: $R^2 = 1 - \frac{507,997}{836} = 0,3923$

pouze 39 % změn ve výsledku druhého testu je vysvětleno výsledkem testu prvního. Zbylých 61 % je ovlivněno jinými vlivy. Z toho vyplývá, že použití regresní přímky není příliš vhodné a model není kvalitní.

Tento závěr lze ještě otestovat pomocí celkového F-testu (na hladině významnosti 0,05).

H_0 : model není vhodný

H_1 : $\neg H_0$

$F = \frac{328,003}{2-1} : \frac{507,997}{8-2} = 3,874 < F_{0,95}(1;6) = F.INV(0,95;1;6) = 5,987 \Rightarrow H_0$ nezamítáme, tzn., že model je chybně specifikován.

Lze využít analytického nástroje Regrese, z tabulky ANOVA dostáváme: významnost $F = 0,097 > 0,05 \Rightarrow H_0$ nezamítáme.

Poznámky

Kvalita modelu

R^2 ... koeficient determinace

Celkový F-test

$H_0 : a = b = 0$

$H_1 : \neg H_0$

EXCEL:

Tvorba modelu pomocí Excelu

Vložení -> Bodový graf -> Přidat spojnicí trendu

zaškrtneme Zobrazit rovnici v grafu

zaškrtneme Zobrazit hodnotu spolehlivosti R (koeficient determinace)

Využití analytického nástroje Regrese v Excelu.

1. Data -> Analýza dat -> Regrese

2. Vstupní oblast Y - závislá proměnná; Vstupní oblast X - nezávislá proměnná; Výstupní oblast - libovolná buňka

3. zaškrtneme Hladina spolehlivosti

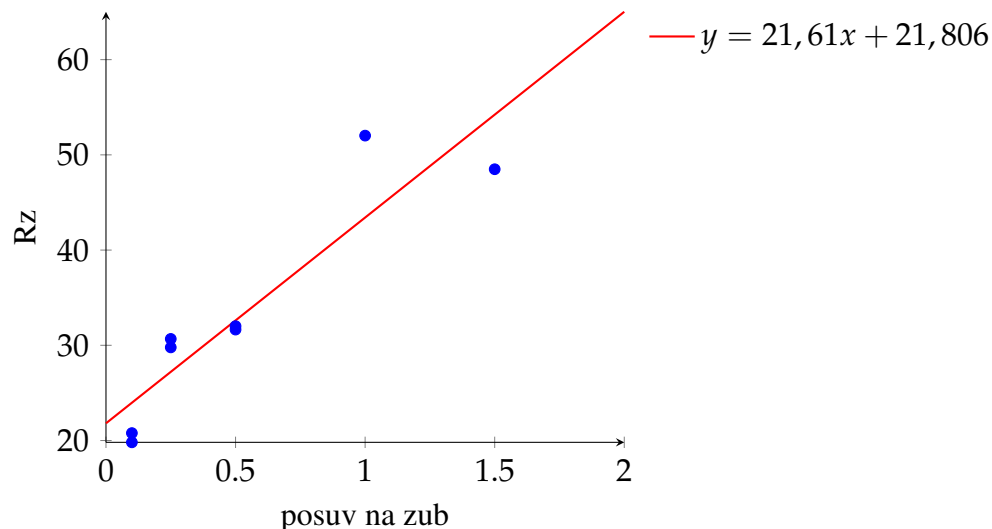
160 - Regresní model - korelační analýza

Zadání Zjišťovalo se, zda existuje závislost mezi hodnotou příčné drsnosti ocele Rz a nastavenou hodnotou posuvu na zub. Výsledky jsou zaznamenány v tabulce.

drsnost Rz	20,78	19,81	30,67	29,78	31,65	32,01	52,02	48,49
posuv na zub	0,1	0,1	0,25	0,25	0,5	0,5	1	1,5

Řešení **Teorie: 77 Příklady: 233,234,235,236**

Proměnné vyneseme do bodového grafu, závislou (drsnost Rz) na osu y a nezávislou (posuv na zub) na osu x . S využitím metody nejmenších čtverců odhadneme regresní přímku: $y = 21,61x + 21,806$.



Hodnota koeficientu determinace:

$R^2 = 0,837 \Rightarrow$ model je vhodně zvolený.

Koeficient korelace:

$R = 0,915 \Rightarrow$ značí silnou lineární závislost mezi sledovanými veličinami.

S rostoucí hodnotou posuvu na zub roste i hodnota příčné drsnosti ocele.

Poznámky

Kvalita modelu

R^2 ... koeficient determinace

Míra závislosti

$R = \sqrt{R^2}$... korelační koeficient

EXCEL:

=CORREL(matice1;matice1)

Dílčí t-test

$H_0 : a = 0$

$H_1 : a \neq 0$

Celkový F-test

$H_0 : a = b = 0$

$H_1 : \neg H_0$

EXCEL:

Tvorba modelu pomocí Excelu

Vložení $\rightarrow$ Bodový graf $\rightarrow$ Přidat spojnicí trendu

zaškrtneme Zobrazit rovnici v grafu

zaškrtneme Zobrazit hodnotu spolehlivosti R (koeficient determinace)

Využití analytického nástroje Regrese v Excelu.

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Regrese

2. Vstupní oblast Y - závislá proměnná; Vstupní oblast X - nezávislá proměnná; Výstupní oblast - libovolná buňka

3. zaškrtneme Hladina spolehlivosti

Matematika III: Pracovní listy do cvičení

Viktor Dubovský, Marcela Jarošová, Jiří Krček, Jitka Krčková, Petra Schreiberová, Petr Volný

Katedra matematiky a deskriptivní geometrie

VŠB - Technická univerzita Ostrava

Příklady – Kombinatorika

163 - Kombinatorické pravidlo součtu

Zadání

- a) Ve skupině je 10 studentů a 7 studentek. Kolik členů má skupina celkem?
- b) Ve skupině je 15 studentů s vykonanou zkouškou z Matematiky, 12 se zkouškou z Fyziky a 10 se zkouškou z obou předmětů. Kolik členů skupina má?
- c) Ve skupině 17 studentů je 15 studentů, kteří mají zkušenosti s alkoholem, 9 vyzkoušelo marihuanu a 4 extázi. Dále 8 přiznalo zkušenost s alkoholem a marihuanou, 1 s extází a marihuanou a 3 s alkoholem a extází. Žádný z nich nevyzkoušel všechny tři zmíněné drogy.
- i) Je ve skupině student, který neokusil žádnou z těchto drog?
- ii) Kolik studentů jen pije alkohol?

Řešení

Teorie: 11 Řešené příklady: 80

Tahák

Kombinatorické pravidlo součtu

Necht' $A_1, A_2, \dots, A_k$ jsou konečné množiny, které mají $p_1, p_2, \dots, p_k$ prvků, a které jsou po dvou vzájemně disjunktní, tj. nemají společný prvek, pak počet prvků sjednocení $A_1 \cup A_2 \cup \dots \cup A_k$ je roven součtu

$$p_1 + p_2 + \dots + p_k.$$

164 - Kombinatorické pravidlo součinu

Zadání

- a) V rovině je dáno 7 bodů, z nichž žádné tři neleží na jedné přímce. Kolik je těmito body určeno
- i) polopřímek?
 - ii) přímek?
- b) Kolik čtyřciferných přirozených čísel větších než 5000 lze sestavit z číslic 1, 3, 5, 7,
- i) jestliže se žádná z cifer neopakuje?
 - ii) jestliže se cifry mohou opakovat?
 - iii) jestliže má být číslo sudé?

Řešení

Teorie: 11 Řešené příklady: 81

Tahák

Kombinatorické pravidlo součinu

Počet všech uspořádaných k -tic, jejichž první člen lze vybrat z n_1 možných prvků, druhý člen z n_2 prvků, atd. až k -tý člen z n_k prvků, je rovný součinu

$$n_1 \cdot n_2 \cdot \dots \cdot n_k.$$

165 - Kombinatorická pravidla součtu a součinu

Zadání

- a) V menze si lze vybrat ze dvou polévek, pěti hlavních jídel a tří salátů. Kolik různých obědových menu lze sestavit, pokud víme, že
- i) byly objednány všechny položky, tj. polévka, hlavní jídlo i salát.
 - ii) právě jedna z položek byla vynechána.
 - iii) alespoň jedna z položek byla vynechána.
- b) V menze si lze vybrat ze dvou polévek, pěti hlavních jídel a tří salátů. Cena polévek je 10Kč, dvě hlavní jídla stojí 30 Kč, dvě 34 Kč a jedno 43 Kč. Saláty jsou za 7 Kč, 11 Kč a 13 Kč. Kolik různých cen lze zaplatit, pokud víme, že
- i) byly objednány všechny položky, tj. polévka, hlavní jídlo i salát.
 - ii) právě jedna z položek byla vynechána.
 - iii) alespoň jedna z položek byla vynechána.

Řešení

Teorie: 11 Řešené příklady: 80,81

Tahák

Kombinatorické pravidlo součinu

Počet všech uspořádaných k -tic, jejichž první člen lze vybrat z n_1 možných prvků, druhý člen z n_2 prvků, atd. až k -tý člen z n_k prvků, je rovný součinu

$$n_1 \cdot n_2 \cdot \dots \cdot n_k.$$

Kombinatorické pravidlo součtu

Necht' $A_1, A_2, \dots, A_k$ jsou konečné množiny, které mají $p_1, p_2, \dots, p_k$ prvků, a které jsou po dvou vzájemně disjunktní, tj. nemají společný prvek, pak počet prvků sjednocení $A_1 \cup A_2 \cup \dots \cup A_k$ je roven součtu

$$p_1 + p_2 + \dots + p_k.$$

166 - Permutace bez opakování

Zadání

- a) i) Kolika způsoby lze srovnat 6 knih do police?
ii) V kolika různých pořadích může být přehráno 13 písní z CD?
iii) Kolik možných výsledků může mít závod, kterého se účastní 24 běžců?
iv) Kolik možností obsazení všech 49 míst autobusu Karosa C954?
- b) Porovnejte výsledky části a).

Řešení

Teorie: 12 Řešené příklady: 82

Tahák

Permutace bez opakování

Permutace z n prvků je uspořádaná n -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje právě jednou. Značí se $P(n)$ a platí, že

$$P(n) = n!.$$

Faktoriál

$$n! = \prod_{i=1}^n i$$
$$0! = 1$$

167 - Variace bez opakování

Zadání

- a) Kolik signálů lze vyslat, použijeme-li dva z pěti různobarevných praporků?
- b) Kolik tříkopečkových zmrzlin si jde objednat, pokud si vybíráme ze sedmi různých příchutí a žádnou nezvolíme vícekrát?
- c) Kolika způsoby lze sestavit denní pětihodinový rozvrh, v němž se žádný z jedenácti předmětů neopakuje?

Řešení

Teorie: 12 Řešené příklady: 83

Tahák

Variace bez opakování

k -členná variace z n prvků je uspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše jednou. Značí se $V_k(n)$ a platí, že

$$V_k(n) = \frac{n!}{(n-k)!}.$$

168 - Kombinace bez opakování

Zadání

- Na projektu pracuje 6 lidí. Kolik lze utvořit dvojic, které budou prezentovat výsledky?
- Na hřišti se sešlo deset kluků. Kolika způsoby se mohou rozdělit do pětičlenných týmů?
- Do turnaje se přihlásilo 12 týmů. Organizátoři se rozhodli pro systém kvalifikačních skupin, jejichž vítězové postoupí do skupiny finálové. Ve všech skupinách hraje „každý s každým.“ Odehraje se více zápasů v případě 4 kvalifikačních skupin nebo v případě 3 kvalifikačních skupin?

Řešení

Teorie: 12 Řešené příklady: 84

Tahák

Kombinace bez opakování

k -členná kombinace z n prvků je ne-uspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše jednou. Značí se $C_k(n)$ a platí, že

$$C_k(n) = \binom{n}{k} = \frac{n!}{k! (n - k)!}.$$

Základní vlastnosti kombinačních čísel

$$\binom{n}{0} = 1$$

$$\binom{n}{n} = 1$$

$$\binom{n}{1} = n$$

169 - Variace s opakováním

Zadání

- a) Kolik existuje čtyřmístných číselných PINů, které jsou složeny z číslic 1, 3, 7?
- b) Kolik stavů systému může signalizovat řídicí panel s šesti diodami, může-li každá z nich svítit, nebo blikat červeně, zeleně, anebo být zhasnuta?
- c) Kolo 1. fotbalové ligy sestává z 8 zápasů. Kolik možných tipů na tyto zápasy lze podat, pokud tipujeme
- i) jednoduché tipy 1,0,2 (domácí, remíza, hosté),
 - ii) dvojtipy,
 - iii) jednoduché tipy na jednotlivé poločasy,
 - iv) přesný počet gólů v zápase.

Řešení

Teorie: 13 Řešené příklady: 85

Tahák

Variace s opakováním

k -členná variace s opakováním z n prvků je uspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše k -krát. Pro jejich počet platí

$$V'_k(n) = n^k.$$

170 - Permutace s opakováním

Zadání

- a) Kolik existuje PINů složených z číslic 1, 1, 2, 3?
- b) V kolika různých pořadích lze sníst 3 jahodové, 4 borůvkové a 2 bílé jogurty?
- c) Kolika způsoby můžeme rozdělit 10 růží mezi 4 přítelkyně?
- d) Kolik je možných rozestavení figurek na šachovnici 6×6 , máme-li k dispozici 8 bílých, 7 žlutých, 6 oranžových, 5 červených, 4 modré, 3 zelené, 2 fialové a jednu černou figurku?

Řešení

Teorie: 13 Řešené příklady: 86

Tahák

Permutace s opakováním

Základní množina má n prvků, které jsou k různých typů. Přitom $n_1, \dots, n_k$ značí počet prvků daného typu a platí $n = n_1 + n_2 + \dots + n_k$. Permutace s opakováním je uspořádaná n -tice, z těchto prvků sestavená. Značí se $P'(n_1, n_2, \dots, n_k)$ a jejich počet je

$$P'(n_1, n_2, \dots, n_k) = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}.$$

171 - Kombinace s opakováním

Zadání

- a) Útrata byla uhrazena třemi bankovkami v hodnotě 100 a 200Kč.
- i) Kolik různých částek mohlo být zaplaceno?
 - ii) O jaké částky jde?
- b) Kolika způsoby lze vybavit 3 obchodní zástupce služebním telefonem, vybíráme-li z 6 druhů přístrojů?
- c) Parta 7 dělníků si vybírá ze 4 obědových menu. Kolika způsoby si mohou objednat?

Řešení

Teorie: 13 Řešené příklady: 87

Tahák

Kombinace s opakováním

k -členná kombinace s opakováním z n prvků je neuspořádaná k -tice sestavená z těchto prvků tak, že každý se v ní vyskytuje nejvýše k -krát. Značí se $C'_k(n)$ a jejich počet je

$$\begin{aligned} C'_k(n) &= \binom{n+k-1}{k} \\ &= \frac{(n+k-1)!}{k!(n-1)!}. \end{aligned}$$

172 - Úprava výrazů s kombinačními čísly

Zadání Vyjádřete jedním kombinačním číslem a vyčíslete

a) $\binom{3}{3} + \binom{7}{6},$

b) $\binom{8}{3} + \binom{8}{6} + \binom{9}{2},$

c) $\binom{15}{3} + \binom{15}{11} + \binom{16}{11} + \binom{17}{13}.$

Řešení

Teorie: 14 Řešené příklady: 88,89

Tahák

Kombinační čísla

základní

$$\binom{n}{0} = 1$$

$$\binom{n}{n} = 1$$

$$\binom{n}{1} = n$$

symetrie

$$\binom{n}{k} = \binom{n}{n-k}$$

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

173 - Rovnice s kombinačními čísly

Zadání Vyřešte rovnici

a) $\binom{x+1}{2} + \binom{x+2}{2} = \binom{x+4}{2},$

b) $\binom{x+2}{x+1} + \binom{x+2}{x} = \binom{x+3}{x},$

c) $\binom{x}{2} + \binom{x}{3} = \binom{x+1}{4}.$

Řešení

Teorie: 14 Řešené příklady: 92

Tahák

Kombinační čísla

základní

$$\binom{n}{0} = 1$$

$$\binom{n}{n} = 1$$

$$\binom{n}{1} = n$$

symetrie

$$\binom{n}{k} = \binom{n}{n-k}$$

součet

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

174 - Směs kombinatorických postupů

Zadání

a) ARMATURA

- i) Kolik přesmyček lze sestavit z písmen slova ARMATURA?
- ii) Kolik z nich neobsahuje RUM?
- iii) V kolika z nich se střídají souhlásky se samohláskami?

b) Ve skupině je 15 studentů a 12 studentek.

- i) Lze sestavit více trojic, které obsahují právě dva chlapce, nebo trojic, které obsahují právě dvě dívky?
- ii) Kolik sestavíme čtveřic, v nichž jsou alespoň dvě studentky?
- iii) Kolik je pětic s nejvýše jedním studentem?

Řešení

Řešené příklady: 93

175 - Směs kombinatorických postupů

Zadání

- a) Výběrového řízení se účastní 7 firem, kritériem výběru je nejnižší cena.
- i) Kolik je možných výsledků soutěže?
 - ii) Kolik je možných výsledků, pokud víme, že *firma č.2* a *firma č.5* jsou v pořadí „vedle sebe“ a navíc, že *firma č.5* soutěž nevyhrála?
 - iii) Kolika způsoby mohou být obsazena první tři místa, víme-li, že *firma č.3* byla vyřazena a *firma č.6* nepodala nejvyšší ani nejnižší nabídku?
- b) Kolem stolu má být rozmístěno 6 židlí, 2 bílé, 2 červené a 2 černé.
- i) Kolik je možných rozestavení, je-li stůl kulatý?
 - ii) Kolik je možností pro obdélníkový stůl, v jehož čele je červená židle a proti ní židle bílá?

Řešení

Řešené příklady: 93

176 - Směs kombinatorických postupů

Zadání

a) Kolik lze z čísel $1, 2, 3, 4, 5, 6, 7, 8, 9$ sestavit trojic,

i) takových, že jejich součet je menší než 10?

ii) takových, že jejich součet je lichý?

iii) takových, že jejich součin je sudý?

b) Kolika způsoby lze na šachovnici 8×8 umístit

i) dva bílé pěšce tak, aby nestáli ve stejném sloupci ani řádku?

ii) dva bílé pěšce tak, aby stáli na různobarevných polích, ale nestáli ve stejném sloupci?

iii) 6 různých figur tak, aby 4 stály na černých a 2 bílých polích?

Řešení

Řešené příklady: 93

Příklady – Pravděpodobnost

178 - Náhodný pokus, náhodný jev

Zadání

- a) Náhodným pokusem jsou dva hody mincí.
- i) Vypište všechny elementární jevy ω_i , základní prostor Ω a všechny náhodné jevy A_i .
 - ii) Určete jejich pravděpodobnosti.
- b) Náhodným pokusem jsou tři hody mincí. Jev A je padnutí panny v prvním hodu, jev B padnutí dvou orlů a jev C padnutí panny ve třetím hodu. Zapište jevy a určete jejich pravděpodobnost.
- i) $A \cup B$, $A \cup C$, $B \cup C$ a $A \cup B \cup C$,
 - ii) $A \cap B$, $A \cap C$, $B \cap C$ a $A \cap B \cap C$,
 - iii) $\bar{A}$, $\bar{B}$ a $\bar{C}$,
 - iv) $A - B$, $B - A$ a $A - C$.

Řešení

Teorie: 16,17 Řešené příklady: 95,96,97

Tahák

Sjednocení (součet) jevů A, B

$$A \cup B = \{\omega \in \Omega : (\omega \in A) \vee (\omega \in B)\}$$

Průnik (součin) jevů A, B

$$A \cap B = \{\omega \in \Omega : (\omega \in A) \wedge (\omega \in B)\}$$

Rozdíl jevů A, B

$$A - B = \{\omega \in \Omega : (\omega \in A) \wedge (\omega \notin B)\}$$

Jev opačný k A

$$\bar{A} = \Omega - A = \{\omega \in \Omega : \omega \notin A\}$$

179 - Náhodný pokus, náhodný jev

Zadání Náhodným pokusem je hod červenou a černou kostkou. Sledujeme rozdíl červené a černé hodnoty.

- Jev A nastává, pokud je rozdíl záporný. Jev B nastává, pokud je rozdíl kladný. Jsou jevy A , B neslučitelné? Pokud ano, tvoří úplný systém neslučitelných jevů? Pokud ne, doplňte jev C , tak aby se o úplný systém jednalo. Určete pravděpodobnosti $P(A)$, $P(B)$ a $P(C)$.
- Jev A nastává, pokud je rozdíl rovný dvěma. Jev B nastává, pokud je absolutní hodnota rozdílu rovna dvěma. Jev C , pokud je rozdíl menší než tři. Vyjádřete vztahy (podjev, rovnost) jevů A , B , C . Určete pravděpodobnosti $P(A)$, $P(B)$ a $P(C)$.
- Jev A nastává, pokud je rozdíl násobkem 5, jev B násobek 4 a jev C násobek 2. Určete vztahy mezi jevy a dále jevy $A \cup B$, $A \cap B$, $\bar{A}$. Určete pravděpodobnosti všech jevů.

Řešení

Teorie: 16,17 Řešené příklady: 95,96,97

Tahák

Jev A je podjevem B

$$A \subset B \Leftrightarrow \{\forall \omega \in \Omega : (\omega \in A) \Rightarrow (\omega \in B)\}$$

Jevy A , B se rovnají

$$A = B \Leftrightarrow \{\forall \omega \in \Omega : (\omega \in A) \Leftrightarrow (\omega \in B)\}$$

Jevy A , B jsou neslučitelné

$$A \cap B = \emptyset$$

Systém neslučitelných jevů $A_1, A_2, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

Úplný systém neslučitelných jevů $A_1, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

$$A_1 \cup A_2 \cup \dots \cup A_n = \Omega$$

$$0 \leq P(A) \leq 1.$$

181 - Geometrická pravděpodobnost

Zadání

- Vodovodní potrubí je mezi místy A , B , vzdálenými 183 m, přerušeno. Jaká je pravděpodobnost, že porucha je nejvýše 27 m od krajních vstupů?
- V nástěnných hodinách se vybíjí baterie. Jaká je pravděpodobnost, že se zastaví dopoledne?
- Třímetrovou tyč rozdělíme na tři kusy. Jaká pravděpodobnost, že z těchto dílů lze sestavit trojúhelník?
- Mezi 11:11 a 12:34 očekáváme doručení dvou zásilek. Vyložení zásilky A trvá deset minut, zásilku B předají za 3 minuty. Jaká je pravděpodobnost, že nebude žádný z kurýrů čekat?

Řešení

Teorie: 18 Řešené příklady: 101

Tahák

Geometrická pravděpodobnost

Pokud lze základní prostor Ω modelovat jako geometrický útvar $\mathcal{O}$ a náhodný jev A jako útvar $\mathcal{A}$, přičemž platí $\mathcal{A} \subseteq \mathcal{O}$, pak pravděpodobnost nastoupení jevu A lze spočítat jako podíl

$$P(A) = \frac{\mu(\mathcal{A})}{\mu(\mathcal{O})},$$

kde $\mu(\mathcal{A})$ resp. $\mu(\mathcal{O})$ značí *míru* útvaru $\mathcal{A}$, resp. $\mathcal{O}$.

Obvykle je $\mathcal{O}$ konečná, uzavřená oblast na přímce, v rovině či v prostoru a $\mathcal{A}$ její uzavřená podmnožina, a jejich mírou rozumíme délku, plochu či objem.

182 - Podmíněná pravděpodobnost

Zadání

- a) Jaká je pravděpodobnost, že součet dvou hodů kostkou bude větší než 7, víme-li, že padla trojka.
- b) V klobouku je a bílých a b černých koulí. Koule náhodně losujeme a nevracíme. S jakou pravděpodobností vytáhneme ve druhém tahu černou, byla-li černá tažena v tahu prvním?
- c) Podle úmrtnostní tabulky se ze 100 000 narozených chlapců 50 let dožije 93 906, 75 let 53 818 a stovky 428 mužů. Určete pravděpodobnost, že se padesátiletý muž dožije 75 a 100 let.

Řešení

Teorie: 20 Řešené příklady: 102,103

Tahák

Podmíněná pravděpodobnost

Pravděpodobnost nastoupení jevů B za předpokladu, že nastoupil jev A .

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Nezávislé jevy

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

$$P(A|B) = \frac{P(B \cap A)}{P(B)} = P(A)$$

$$P(A \cap B) = P(A) P(B),$$

přičemž je $P(A) > 0$ a $P(B) > 0$.

Závislé jevy

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

183 - Nezávislé jevy

Zadání

- a) Náhodným pokusem jsou tři hody mincí. Popište základní prostor elementárních jevů. Zjistěte, zda jsou následující jevy nezávislé.
- i) v prvním hodu orel, ii) dvakrát orel, iii) aspoň jednou orel, iv) v posledním orel.
- b) Náhodným pokusem je tah jedné karty z pokerového balíčku 52 karet. Zjistěte, zda jsou následující jevy nezávislé.
- i) tažen král, ii) tažen obrázek, iii) tažena srdce, iv) tažena černá.

Řešení

Teorie: 20 Řešené příklady: 102,103

Tahák

Podmíněná pravděpodobnost

Pravděpodobnost nastoupení jevů B za předpokladu, že nastoupil jev A .

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Nezávislé jevy

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

$$P(A|B) = \frac{P(B \cap A)}{P(B)} = P(A)$$

$$P(A \cap B) = P(A) P(B),$$

přičemž je $P(A) > 0$ a $P(B) > 0$.

Závislé jevy

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

184 - Úplná pravděpodobnost

Zadání

- a) Pět procent z produkce výrobní linky jsou zmetky. Výstupní kontrola 98 % zmetků zachytí a vyřadí. Jako zmetek však označí, a vyřadí, i tři procenta dobrých výrobků. Kolik procent výrobků kontrola vyřadí?
- b) Ze šestnácti mincí jsou tři falešné se dvěma pannami. Jaká je pravděpodobnost, že ve třech hodech padne
- i) třikrát panna, ii) třikrát orel, iii) jedna panna, iv) jeden orel.

Řešení

Teorie: 21 Řešené příklady: 104

Tahák

Úplná pravděpodobnost

úplný systém neslučitelných jevů
 $A_1, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

$$\bigcup_{i=1}^n A_i = \Omega$$

úplná pravděpodobnost

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

185 - Bayesův vzorec

Zadání

- Mezi čtyřmi mincemi je jedna falešná se dvěma pannami. Náhodně jednu minci vybereme a dvakrát hodíme. Pokaždé padne panna. Jak velká je pravděpodobnost, že je mince falešná?
- Student zná odpověď na $\frac{2}{3}$ otázek z testu. Ve zbylých případech tipuje jednu ze čtyř možností. Jaká je pravděpodobnost správné odpovědi? Jaká je pravděpodobnost, že správně zodpovězenou otázku student netipoval?
- V populaci je infikováno 5 % jedinců. Test je pozitivní v 90 % případů skutečně nakažených. Avšak pozitivní výsledek dá i v případě 10 % zdravých jedinců. Jaká je pravděpodobnost, že testovaný jedinec je zdravý, přestože je jeho test pozitivní?

Řešení

Teorie: 21 Řešené příklady: 106

Tahák

Bayesův vzorec

neslučitelné jevy

$$A_1 \cap A_2 = \emptyset$$

úplný systém neslučitelných jevů

$$A_1 \cap A_2 = \emptyset$$

$$A_1 \cup A_2 = \Omega$$

úplná pravděpodobnost

$$P(B) = P(B|A_1) P(A_1) + P(B|A_2) P(A_2)$$

Bayesův vzorec

$$\begin{aligned} P(A_1|B) &= \\ &= \frac{P(B|A_1) P(A_1)}{P(B|A_1) P(A_1) + P(B|A_2) P(A_2)} \end{aligned}$$

186 - Bayesův vzorec

Zadání

- a) Tři střelci A , B , C zasáhnou terč s pravděpodobnostmi $P(A) = \frac{5}{6}$, $P(B) = \frac{5}{7}$ a $P(C) = \frac{5}{8}$. O střelci rozhoduje hod kostkou. Padne-li 1 střílí A , padne-li sudé střílí B , jinak střílí C . Náhodně vybraný střelec vystřelil. Jaká je pravděpodobnost zásahu. Pokud byl cíl zasažen, jaká je pravděpodobnost, že střílel střelec A .
- b) Každý ze tří střelců vystřelí jednou na terč. Pravděpodobnosti, že nezasáhnou jsou $P(A) = \frac{1}{5}$, $P(B) = \frac{1}{6}$ a $P(C) = \frac{1}{7}$. Spočítejte pravděpodobnost, že terč zasáhli střelci B a C , když víte, že byl dvakrát zasažen.
- c) V pytlíku je deset mincí. Šest „spravedlivých“, označme je A , pro něž je pravděpodobnost pádu orla $P(O|A) = 0,5$. Tři, typ B jsou falešné s $P(O|B) = 0,7$. Poslední mince, typ C , je taky falešná, ale zvýhodňuje pannu, a to tak, že $P(P|C) = 0,6$. Vytáhneme náhodnou minci, hodíme a padne orel. Jaká je pravděpodobnost, že mince je spravedlivá?

Řešení

Teorie: 21 Řešené příklady: 106,107

Tahák

Bayesův vzorec

úplný systém neslučitelných jevů
 $A_1, \dots, A_n$

$$A_i \cap A_j = \emptyset, \text{ pro } i \neq j$$

$$\bigcup_{i=1}^n A_i = \Omega$$

úplná pravděpodobnost

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

Bayesův vzorec

$$P(A_k|B) = \frac{P(B|A_k) P(A_k)}{\sum_{i=1}^n P(B|A_i) P(A_i)}$$

187 - Opakované pokusy

Zadání

- a) Házíme kostkou a počítáme součin padlých hodnot. Jaká je pravděpodobnost, že po pěti hodech bude hodnota dělitelná
- i) 625, ii) 125, iii) 25, iv) 5.
- b) Z balíčku 52 karet jednu taháme a vracíme zpět. Jaká je pravděpodobnost, že ve třinácti tazích
- i) byla jen třikrát vytažena hodnota menší než 5? ii) bylo aspoň dvakrát vytaženo eso?

Řešení

Teorie: 22 Řešené příklady: 108,109

Tahák

Pravděpodobnost opakovaných pokusů

Je-li pravděpodobnost nastoupení jevů A v pokusu $P(A) = p$, pak pravděpodobnost jevu A_k , že jev A nastane v sérii n opakovaných nezávislých pokusů právě k -krát, je za předpokladu, že nastoupil jev A .

$$P(A_k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

188 - Závislé opakované pokusy

Zadání

- a) V klobouku je 15 bílých a 5 černých koulí. Náhodně taháme a již nevracíme zpět. Jaká je pravděpodobnost, že v sedmi tazích
- i) vytáhneme více černých? ii) vytáhneme více bílých?
- b) Z balíčku 52 karet jednu vytáhneme a již nevracíme zpět. Jaká je pravděpodobnost, že ve třinácti tazích
- i) byla jen třikrát vytažena hodnota menší než 5? ii) bylo aspoň dvakrát vytaženo eso?

Řešení

Teorie: 22 Řešené příklady: 110

Tahák

Pravděpodobnost závislých opakovaných pokusů

V souboru n prvků, má k , $0 \leq k \leq n$ určitou vlastnost a $n - k$ tuto vlastnost nemá. Ze souboru vybíráme postupně j prvků, které nevracíme. Označme A_i jev, odpovídající tomu, že v tomto výběru má právě i prvků sledovanou vlastnost. Pravděpodobnost $P(A_i)$ vypočteme podle vzorce

$$P(A_i) = \frac{\binom{k}{i} \cdot \binom{n-k}{j-i}}{\binom{n}{j}}.$$

Příklady – Náhodná veličina

190 - Náhodná veličina

Zadání Na stavbě domu pracuje 15 zaměstnanců stavební firmy, z nichž 4 mají certifikát pro práci ve výškách. Stavbyvedoucí náhodně vybere 4 zaměstnance. Určete pravděpodobnostní a distribuční funkci náhodné veličiny X , která představuje počet certifikovaných zaměstnanců, kteří mohou provádět výškové práce mezi těmito 4 vybranými.

Řešení

Teorie: 25,26 Řešené příklady: 112,113,114

Tahák

Diskrétní náhodná veličina

Pravděpodobnostní funkce

$$p(x) = P(X = x)$$

Vlastnosti

- $p(x_i) \geq 0$
- $\sum_{i=1}^n p(x_i) = 1$

Distribuční funkce

$$F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$$

Vlastnosti

- $0 \leq F(x) \leq 1$
- $F(x)$ je neklesající funkce
- $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$
- $F(x)$ je zleva spojitá

191 - Náhodná veličina

Zadání Pravděpodobnost poruchy každého ze 4 nezávisle fungujících motorů letadla Boeing 747 je 0,1. Náhodná veličina X udává počet motorů, které mají poruchu.

- Nalezněte pravděpodobnostní funkci náhodné veličiny X .
- Určete distribuční funkci a její graf.

Řešení

Teorie: 25,26 Řešené příklady: 112,113,114

Tahák

Diskrétní náhodná veličina

Pravděpodobnostní funkce
 $p(x) = P(X = x)$

Vlastnosti

- $p(x_i) \geq 0$
- $\sum_{i=1}^n p(x_i) = 1$

Distribuční funkce
 $F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$

Vlastnosti

- $0 \leq F(x) \leq 1$
- $F(x)$ je neklesající funkce
- $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$
- $F(x)$ je zleva spojitá

192 - Náhodná veličina

Zadání Je dána funkce $p(x) = \frac{b}{x}$, kde $x = \{1, 2, 3\}$. Určete konstantu b tak, aby tato funkce byla pravděpodobnostní funkcí nějaké náhodné veličiny.

Řešení

Teorie: 26 Řešené příklady: 112, 113

Tahák

Diskrétní náhodná veličina

Pravděpodobnostní funkce
 $p(x) = P(X = x)$

Vlastnosti

- $p(x_i) \geq 0$
- $\sum_{i=1}^n p(x_i) = 1$

193 - Náhodná veličina

Zadání Na začátku sezóny basketbalové ligy si hráči mezi sebe rozdělují dresy s následujícími čísly:

3, 4, 11, 12, 13, 15, 22, 23, 32, 33, 41, 50.

Určete pravděpodobnostní a distribuční funkci, střední hodnotu a rozptyl náhodné veličiny X , která je dána součtem cifer na náhodně vybraném dresu hráče.

Řešení

Teorie: 25,26 Řešené příklady: 112,113,114

Tahák

Diskrétní náhodná veličina

Pravděpodobnostní funkce

$$p(x) = P(X = x)$$

Vlastnosti

- $p(x_i) \geq 0$
- $\sum_{i=1}^n p(x_i) = 1$

Distribuční funkce

$$F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$$

Vlastnosti

- $0 \leq F(x) \leq 1$

Střední hodnota

$$E(X) = \sum_i x_i p(x_i)$$

Rozptyl

$$\begin{aligned} D(X) &= \sum_i x_i^2 p(x_i) - [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

194 - Náhodná veličina

Zadání Náhodná veličina X má rozdělení popsané funkcí hustoty

$$f(x) = \begin{cases} cx^2(2-x) & x \in \langle 0; 2 \rangle \\ 0 & x \in (-\infty; 0) \cup (2; \infty) \end{cases}.$$

Určete koeficient c , distribuční funkci $F(x)$ a $P(X > 1)$.

Řešení

Teorie: 25,27 Řešené příklady: 115

Tahák

Spojité náhodné veličiny

Vlastnosti funkce hustoty

- $f(x) \geq 0$
- $f(x) = F'(x)$
- $\lim_{x \rightarrow -\infty} f(x) = 0$, $\lim_{x \rightarrow \infty} f(x) = 0$
- $\int_{-\infty}^{\infty} f(x) = 1$

Vlastnosti distribuční funkce

- $0 \leq F(x) \leq 1$
- $F(x)$ je neklesající funkce
- $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$
- $P(X < x) = F(x)$
- $P(X \geq x) = 1 - F(x)$

195 - Náhodná veličina

Zadání Náhodná veličina X má distribuční funkci:

$$F(x) = \begin{cases} 0 & x < 0 \\ 2x/\pi & x \in \langle 0; \pi/2 \rangle \\ 1 & x \geq \frac{\pi}{2} \end{cases}$$

Určete

- | | |
|---|--|
| a) funkci hustoty $f(x)$ a znázorněte graficky, | c) disperzi a směrodatnou odchylku, |
| b) střední hodnotu příslušné náhodné veličiny, | d) pravděpodobnost $P(\frac{1}{2} \leq X < 1)$. |

Řešení

Teorie: 27,28,29 Řešené příklady: 116

Tahák

Spojité náhodné veličiny

Vlastnosti $f(x)$ a $F(x)$

- $f(x) \geq 0$
- $f(x) = F'(x)$
- $\lim_{x \rightarrow -\infty} f(x) = 0$, $\lim_{x \rightarrow \infty} f(x) = 0$
- $P(x_1 \leq X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$

Střední hodnota

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Rozptyl

$$D(X) = \sigma^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - [E(X)]^2 = E(X^2) - [E(X)]^2$$

Směrodatná odchylka

$$\sigma = \sqrt{\sigma^2} = \sqrt{D(X)}$$

196 - Náhodná veličina

Zadání Určete první dva decily $x_{0,1}$; $x_{0,2}$ a třetí kvartil $x_{0,75}$ pro

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{2}x - 1 & x \in \langle 2; 4 \rangle \\ 1 & x > 4 \end{cases}.$$

Řešení

Teorie: 30 Řešené příklady: 119

Tahák

Spojité náhodné veličiny

$F(x_p) = p$, kde $p \in \langle 0, 1 \rangle$

x_p ... p -kvantil

$F(x)$... distribuční funkce

Kvantily:

- kvantily: $x_{0,25}$; $x_{0,5}$; $x_{0,75}$
- decily: $x_{0,1}$; $x_{0,2}$; $\dots$; $x_{0,9}$

197 - Náhodná veličina

Zadání Náhodná veličina X má hustotu:

$$f(x) = \begin{cases} x^3 e^{-x} & x \in (0; \infty) \\ 0 & x \notin (0; \infty) \end{cases}.$$

Určete modus.

Řešení

Teorie: 28,30 Řešené příklady: 119

Tahák

Spojité náhodné veličiny

Modus

M_0 - je hodnota, v níž funkce hustoty $f(x)$ nabývá lokálního maxima

Lokální extrémy

$$f'(x) = 0$$

- v bodě x_0 za předpokladu, že $f'(x_0) = 0$ a $f''(x_0) < 0$, má funkce lokální maximum
- spojitá funkce, jejíž derivace mění v bodě x_0 znaménko, má v bodě x_0 lokální extrém

Příklady – Rozdělení diskrétní náhodné veličiny

199 - Binomické rozdělení

Zadání

- a) Pravděpodobnost toho, že žárovka praskne při jednom přepnutí vypínače je 0,5 %. Jaká je pravděpodobnost, že za jeden rok (nepřestupný) při jednom rozsvícení a zhasnutí denně prasknou nejvýše tři žárovky.
- b) Ve skladu je připraveno k expedici 3000 elektronických součástek stejného typu. Pravděpodobnost, že se součástka spálí při testovacím zapojení je 0,2 %. Jaká je pravděpodobnost, že ze součástek na skladě budou více než tři spáleny při testování?
- c) Dlouhodobým pozorováním byla určena pravděpodobnost jarní povodně na řece na 0,13. Určete pravděpodobnost, že v příštích 30 letech nedojde k více než dvěma povodním. Dále určete střední hodnotu a rozptyl náhodné veličiny, popisující počet povodní v těchto 30 letech.

Řešení

Teorie: 33 Řešené příklady: 122

Tahák

Binomické rozdělení $Bi(n,p)$

n - počet nezávislých pokusů

p - pravděpodobnost úspěchu při jednom pokusu

Pravděpodobnostní funkce

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$x = 0, 1, \dots, n$$

Vlastnosti

$$E(X) = np$$

$$D(X) = np(1-p)$$

EXCEL:

$$P(X = x) = \text{BINOM.DIST}(x; n; p; 0)$$

$$P(X \leq x) = \text{BINOM.DIST}(x; n; p; 1)$$

200 - Hypergeometrické rozdělení

Zadání

- a) V každé dodávce je 1000 kusů výrobků. Při kontrole jich náhodně bez vracení vybereme 20. Dodávka bude přijata, pokud kontrolní výběr obsahuje nejvýše dva zmetky. S jakou pravděpodobností bude odmítnuta dodávka, která obsahuje právě 80 zmetků?
- b) V dodávce 120 polotovarů je 15 vadných. Náhodně vybereme (najednou, tj. „bez vracení“) 6 kusů polotovarů k další kompletaci. Určete střední hodnotu a směrodatnou odchylku náhodné veličiny popisující počet vadných polotovarů v tomto výběru. A jaká je pravděpodobnost, že mezi vybranými prvky bude maximálně jeden vadný?
- c) Ve skladu do krabice, ve které je 1500 šroubů s metrickým závitem, omylem přisypali 180 šroubů s palcovým závitem. Nezaškoleného pomocníka pošlete do této krabice pro 20 šroubů. Jaká je pravděpodobnost, že mezi nimi bude alespoň 18 s metrickým závitem?

Řešení

Teorie: 34 Řešené příklady: 123

Tahák

Hypergeometrické rozdělení $H(N, M, n)$

N - počet prvků základního souboru

M - počet prvků základního souboru se sledovanou vlastností

n - počet prvků ve výběru

Pravděpodobnostní funkce

$$p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$x = 0, 1, \dots, n$$

Vlastnosti

$$E(X) = \frac{nM}{N}$$

$$D(X) = \frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$$

EXCEL:

$$P(X = x) = \text{HYPGEOM.DIST}(x; n; M; N; 0)$$

$$P(X \leq x) = \text{HYPGEOM.DIST}(x; n; M; N; 1)$$

201 - Poissonovo rozdělení

Zadání

- Při psaní 100 stránkové bakalářské práce jste se dopustili celkem 260 gramatických chyb. Jaká je pravděpodobnost, že na náhodně vybrané stránce nebudou více než dvě chyby?
- Během 8 hod. pracovní směny přijede na čerpací stanici s jedním stojanem průměrně 320 automobilů. Jaká je pravděpodobnost, že jich během 45 min. přijede alespoň 25?
- Revizor v ostravské MHD dopadne během jednoho pracovního dne průměrně 14 černých pasažérů. Jaká je pravděpodobnost, že jich za jeden pracovní týden (od pondělí do pátku) dopadne alespoň 65?

Řešení

Teorie: 35 Řešené příklady: 124

Tahák

Poissonovo rozdělení $Po(\lambda)$

λ - průměrný počet výskytů sledovaného jevu v úseku jednotkové délky

Pravděpodobnostní funkce

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$
$$x = 0, 1, \dots$$

Vlastnosti

$$E(X) = \lambda$$

$$D(X) = \lambda$$

EXCEL:

$$P(X = x) = \text{POISSON.DIST}(x; \lambda; 0)$$

$$P(X \leq x) = \text{POISSON.DIST}(x; \lambda; 1)$$

202 - Binomické, hypergeometrické a Poissonovo rozdělení

Zadání

- a) Kapsář pracující na hlavním nádraží okrade průměrně 42 cestujících za týden. Jaká je pravděpodobnost, že se mu dnes podaří okrást alespoň 5 cestujících?
- b) Nově sestavený elektronický přístroj obsahuje 562 součástek stejného typu, u nich výrobce deklaruje pravděpodobnost poruchy při prvním zapojení na 0,8 %. Jaká je pravděpodobnost, že přístroj bude při prvním zapojení fungovat, jestliže porucha tří a více součástek ho vyřadí z provozu?
- c) Do 1. ročníku je zapsáno 212 dívek a 316 chlapců, které zcela náhodně rozdělíme do studijních skupin po 24. Jaká je pravděpodobnost, že ve skupině bude více dívek než chlapců?
- d) Sálový superpočítač má na 500 dní nepřetržitého provozu průměrně 24 poruch. Určete pravděpodobnost poruchy za 100 hodin provozu.
- e) Při výstupní kontrole se z každých 1000 výrobků vybírá 50. Určete střední hodnotu počtu nekvalitních výrobků ve výběru, je-li zmetkovitost výroby 3 %.
- f) V kulometném pásu je 250 nábojů, pravděpodobnost selhání při výstřelu každého z nich je 1,15 %. Jaká je pravděpodobnost, že vystřelíme celý pás bez selhání?

Řešení

Teorie: **33,34,35** Řešené příklady: **122,123,124**

Příklady – Rozdělení spojité náhodné veličiny

204 - Rovnoměrné rozdělení

Zadání

- a) Tramvaje linky č.7 odjíždí během dne ze zastávky Rektorát VŠB pravidelně každých 11 min. Jaká je pravděpodobnost, že při náhodném příchodu na zastávku nebudete čekat déle než 7,5 min? Dále určete střední hodnotu a směrodatnou odchylku náhodné veličiny popisující dobu čekání na tramvaj č.7 při náhodném příchodu na uvedenou zastávku.
- b) Z výrobní linky v automobilce vyjíždí každých 42 min. jeden nový automobil. Jaká je pravděpodobnost, že při čtvrhodinové exkurzi uvidíte, jak nový automobil opouští výrobní linku?
- c) Počítač dodává výsledky numerických výpočtů zaokrouhlené na celá čísla. Jaká je pravděpodobnost, že hodnota reprezentovaná číslem 11 byla ve skutečnosti větší než 11,25?

Řešení

Teorie: 37 Řešené příklady: 126

Tahák

Rovnoměrné rozdělení $R(a, b)$
 a, b - krajní meze intervalu

Hustota pravděpodobnosti

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in \langle a, b \rangle \\ 0 & x \notin \langle a, b \rangle \end{cases}$$

Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, a) \\ \frac{x-a}{b-a} & x \in \langle a, b \rangle \\ 1 & x \in (b, +\infty) \end{cases}$$

Vlastnosti

$$E(X) = \frac{a+b}{2}$$

$$D(X) = \frac{(b-a)^2}{12}$$

205 - Exponenciální rozdělení

Zadání

- Průměrná doba mezi příjezdy automobilů na stanici STK je 12 min. Jaká je pravděpodobnost, že mezi dvěma po sobě jdoucími příjezdy bude prodleva větší než 15 min.?
- Na dvoukilometrovém úseku dálnice D1 je 23 velmi nebezpečných děr. S jakou pravděpodobností narazíte na takovou díru na úseku o délce 120 m?
- Průměrná životnost sledovaného typu žárovky je experimentálně stanovena na 1500 hodin. Jakou životnost má uvádět výrobce ve svých materiálech, jestliže chce, aby deklarované životnosti dosáhlo alespoň 75 % vyrobených žárovek?

Řešení

Teorie: 38 Řešené příklady: 127

Tahák

Exponenciální rozdělení $E(\lambda)$
 λ - převrácená hodnota střední hodnoty

Hustota pravděpodobnosti

$$f(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ \lambda e^{-\lambda x} & x \in \langle 0, +\infty \end{cases}$$

Distribuční funkce

$$F(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ 1 - e^{-\lambda x} & x \in \langle 0, +\infty \end{cases}$$

Vlastnosti

$$E(X) = \frac{1}{\lambda}$$
$$D(X) = \frac{1}{\lambda^2}$$

EXCEL:

$$f(x) = \text{EXPON.DIST}(x; \lambda; 0)$$

$$F(x) = \text{EXPON.DIST}(x; \lambda; 1)$$

206 - Normální rozdělení

Zadání

- a) Hmotnost pytle cementu je hodnocena jako vyhovující, pohybuje-li se v rozmezí 23,20 kg až 25,60 kg. Při dodržení standardních výrobních podmínek podléhá jeho hmotnost normálnímu rozdělení se střední hodnotou 25,00 kg a směrodatnou odchylkou 1,12 kg. Jaká je pravděpodobnost, že hmotnost náhodně kontrolovaného pytle bude v předepsaných mezích?
- b) Hmotnost chlapců v 1. třídě ZŠ je popsána náhodnou veličinou s normálním rozdělením $N(23; 3,61)$. Určete, v jakém intervalu (symetrickém kolem střední hodnoty) se dá předpokládat hmotnost 98 % z nich.
- c) Náhodná veličina popisující hodnotu IQ má pro celou populaci normální rozdělení $N(100, 225)$. Určete, u kolika procent populace můžeme předpokládat hodnotu IQ vyšší než 90.

Řešení

Teorie: 39 Řešené příklady: 128

Tahák

Normální rozdělení $N(\mu, \sigma^2)$

μ - střední hodnota

σ^2 - rozptyl

Hustota pravděpodobnosti

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Distribuční funkce

$$F(x) = \int_{-\infty}^x f(t) dt$$

Vlastnosti

$$E(X) = \mu$$

$$D(X) = \sigma^2$$

EXCEL:

$$f(x) = \text{NORM.DIST}(x; \mu; \sigma; 0)$$

$$F(x) = \text{NORM.DIST}(x; \mu; \sigma; 1)$$

$$x_p = \text{NORM.INV}(p; \mu; \sigma)$$

207 - Rovnoměrné, exponenciální a normální rozdělení

Zadání

- a) Náhodná veličina popisující skutečnou hmotnost 250 g balení dětských piškotů má normální rozdělení se střední hodnotou 247,5 g a směrodatnou odchylkou 7,62 g. Jaká je pravděpodobnost, že námi zakoupený balíček piškotů bude vážit více než 255 g? A v jakém rozsahu můžeme předpokládat hmotnost 95 % všech expedovaných balíčků?
- b) Autobusy z konečné zastávky Studentské koleje odjíždějí pravidelně každých 18 minut. Jaká je pravděpodobnost, že při náhodném příchodu na zastávku budete čekat déle než 5 min?
- c) Průměrná doba, kterou host v nejmenované restauraci čeká na pivo od okamžiku objednávky, je 8 minut. Jaká je pravděpodobnost, že budete čekat na pivo déle než 10 minut? Dále určete dobu čekání, během které budete obslouženi s 95% pravděpodobností.

Řešení

Teorie: **37,38,39** Řešené příklady: **126,127,128**

Příklady – Náhodný vektor

209 - Marginální rozdělení

Zadání U 32 typů ocelí byl zjištěn obsah uhlíku a pevnost příslušné ocele. Byly sledovány 2 kategorie ocele, a to nízkouhlíkové (0 – 0,25 % C) a středněuhlíkové (0,25 – 0,6 % C). Výsledky jsou dané v tabulce.

C [%] \ R _m [MPa]	800 – 1200	1200 – 1600	1600 – 2000
0 – 0,25	6	14	0
0,25 – 0,6	0	4	8

Stanovte pravděpodobnostní rozdělení, distribuční funkci a marginální rozdělení.

Řešení **Teorie: 42,43 Řešené příklady: 131,133,134**

Tahák

Pravděpodobnostní funkce
 $p(x,y) = P(X = x, Y = y)$

Distribuční funkce
 $F(x,y) = \sum_{x_i < x} \sum_{y_j < y} p(x_i, y_j)$

Marginální pravděpodobnostní funkce
 $p_X(x) = P(X = x) = \sum_y p(x,y)$
 $p_Y(y) = P(Y = y) = \sum_x p(x,y)$

Marginální distribuční funkce
 $F_X(x) = P(X < x) = F(x, \infty)$
 $F_Y(y) = P(Y < y) = F(\infty, y)$

210 - Podmíněné pravděpodobnosti

Zadání Pro data z předcházejícího příkladu určete podmíněné pravděpodobnosti a rozhodněte, zda pevnost ocele závisí na obsahu uhlíku.

Řešení **Teorie: 43,44 Řešené příklady: 135**

Tahák

Podmíněná pravděpodobnostní funkce

$$P(x|y) = \frac{p(x,y)}{p_Y(y)}, \quad p_Y(y) \neq 0$$

$$P(y|x) = \frac{p(x,y)}{p_X(x)}, \quad p_X(x) \neq 0$$

Podmíněná distribuční funkce

$$F(x|y) = \frac{\sum_{x < x_i} p(x_i, y)}{p_Y(y)}, \quad p_Y(y) \neq 0$$

$$F(y|x) = \frac{\sum_{y < y_i} p(x, y_i)}{p_X(x)}, \quad p_X(x) \neq 0$$

Nezávislost složek (X, Y)

X, Y jsou nezávislé, pokud:

$$F(x, y) = F_X(x)F_Y(y)$$

$$p(x, y) = p_X(x)p_Y(y)$$

211 - Číselné charakteristiky

Zadání Pro stejná data jako v předchozích příkladech vypočtěte podmíněnou střední hodnotu a rozptyl a určete hodnotu kovariance a koeficientu korelace.

Řešení **Teorie: 45,46 Řešené příklady: 136,137**

Tahák

Podmíněná střední hodnota

$$E(X|Y = y) = \sum_i x_i p(x_i|y)$$

Podmíněný rozptyl

$$D(X|Y = y) = \sum_i [x_i - E(X|y)]^2 p(x_i|y)$$

Kovariance

$$\text{cov}(X, Y) = \sum_i \sum_j x_i y_j p(x_i, y_j) - E(X)E(Y)$$

Koeficient korelace

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}}$$

212 - Nezávislost náhodných veličin

Zadání Rozhodněte, zda jsou náhodné veličiny X a Y nezávislé a vypočítejte kovarianci.

a)

$X \setminus Y$	-2	0	2
0	$\frac{2}{15}$	$\frac{2}{5}$	$\frac{2}{15}$
3	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{15}$

b)

$X \setminus Y$	1	2	3
1	$\frac{1}{8}$	$\frac{2}{9}$	$\frac{1}{9}$
2	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{36}$
3	$\frac{1}{9}$	$\frac{7}{36}$	$\frac{1}{36}$

Řešení Teorie: 43,46 Řešené příklady: 133,137

Tahák

Nezávislost složek (X, Y)
 X, Y jsou nezávislé, pokud:
 $p(x, y) = p_X(x)p_Y(y)$

Kovariance
 $\text{cov}(X, Y) = \sum_i \sum_j x_i y_j p(x_i, y_j) - E(X)E(Y)$

213 - Náhodný vektor

Zadání 16 studentů bylo na zkoušce z matematiky a fyziky s těmito výsledky (první hodnota označuje výsledek z matematiky, druhá z fyziky):

(1,1), (1,2), (1,3), (2,2), (2,3), (2,3), (3,2), (3,3), (3,3), (3,3), (3,3), (3,4), (3,4), (4,3), (4,4), (4,4). Určete

a) sdruženou a marginální pravděpodobnostní a distribuční funkci,

b) $P(x|y)$ a $F(y|x)$,

c) zda jsou X, Y nezávislé.

Řešení

Teorie: 42,43,44 Řešené příklady: 131,133,134,135

Příklady – Popisná analýza

215 - Četnosti, histogram

Zadání V tabulce máme k dispozici údaje o pevnosti oceli v tahu [MPa].

j	x_j	j	x_j	j	x_j
1	608	5	713	9	656
2	728	6	698	10	703
3	663	7	613	11	618
4	668	8	628	12	653

Určete absolutní, relativní, kumulativní a kumulativní relativní četnosti pro čtyři třídy, zobrazte histogram absolutních četností a výsečový graf relativních četností.

Řešení

Teorie: 49,50 Řešené příklady: 139

Tahák

Variační rozpětí

$$R = x_{\max} - x_{\min}$$

Šířka třídy

$$h = \frac{R}{k}, \text{ kde } k \text{ je počet tříd}$$

Absolutní četnost

f_i ... počet prvků patřící do i -té třídy

Histogram

graf absolutních četností f_i

Relativní četnost

$$\varphi_i = \frac{f_i}{n}, \text{ } n \text{ je rozsah souboru}$$

Kumulativní absolutní četnost

$$F_i = \sum_{l=1}^i f_l$$

Kumulativní relativní četnost

$$\phi_i = \frac{F_i}{n}, \text{ } n \text{ je rozsah souboru}$$

Využití analytického nástroje Histogram v Excelu.

216 - Číselné charakteristiky

Zadání V tabulce máme k dispozici údaje o pevnosti oceli v tahu [MPa].

j	x_j	j	x_j	j	x_j
1	608	5	713	9	656
2	728	6	698	10	703
3	663	7	613	11	618
4	668	8	628	12	653

Určete základní charakteristiky polohy a interkvartilové rozpětí.

Řešení **Teorie: 51,52 Řešené příklady: 140,141**

Tahák

Výběrový průměr

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$$

EXCEL: =PRŮMĚR(číslo1;číslo2;...)

Modus

$\hat{x}$... nejčastější hodnota v souboru

EXCEL: =MODE(číslo1;číslo2;...)

Kvantily

Medián $x_{0,5}$; $\tilde{x}$... prostřední hodnota v souboru

EXCEL: =MEDIAN(číslo1;číslo2;...)

Dolní kvartil $x_{0,25}$

EXCEL: =QUARTIL.EXC(matice;1)

Horní kvartil $x_{0,75}$

EXCEL: =QUARTIL.EXC(matice;3)

Interkvartilové rozpětí

$$IQR = x_{0,75} - x_{0,25}$$

Využití analytického nástroje Popisná statistika v Excelu.

217 - Číselné charakteristiky

Zadání V tabulce máme údaje o objemové hmotnosti betonu [kg/m³].

j	x_j	j	x_j	j	x_j
1	2293	5	2358	9	2328
2	2142	6	2284	10	2547
3	2302	7	2328	11	2370
4	2367	8	2305		

Určete základní charakteristiky variability a tvaru.

Řešení **Teorie: 52,53 Řešené příklady: 141,142**

Tahák

Výběrový rozptyl

$$s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

EXCEL: =VAR.S(číslo1;číslo2;...)

Výběrová směrodatná odchylka

$$s = \sqrt{s^2}$$

EXCEL: =SMODCH.VÝBĚR.S(číslo1;číslo2;...)

Variační koeficient

$$V_x = \frac{s}{\bar{x}}$$

Šikmost

A ... popisuje symetrii rozdělení kolem $\bar{x}$

EXCEL: =SKEW(číslo1;číslo2;...)

Špičatost - exces

$\tilde{e}$... koncentrovanost hodnot kolem středu rozdělení

EXCEL: =KURT(číslo1;číslo2;...)

Využití analytického nástroje Popisná statistika v Excelu.

218 - Popisná analýza

Zadání Porovnejte na základě histogramu a číselných charakteristik objemovou tíhu ocele a betonu [kN/m³].

Obj. tíha ocele [kN/m ³]	72	68	80	76	69	82	81	75	77	68	73	75
Obj. tíha betonu [kN/m ³]	18	20	22	30	28	25	31	21	26	30		

Řešení **Teorie:** 50,51,52,53 **Řešené příklady:** 140,141,142

Tahák

Využití analytického nástroje Histogram a Popisná statistika v Excelu.

- 1. Data -> Analýza dat -> Histogram
Vstupní oblast: Hodnoty
Hranice tříd: Horní hranice tříd
zaškrtneme Vytvořit graf
- 2. Data -> Analýza dat -> Popisná statistika
Vstupní oblast: Hodnoty
zaškrtneme Celkový přehled

219 - Popisná analýza

Zadání Porovnejte na základě číselných charakteristik pevnost v tahu u vybraných vláken a whiskerů.

Vlákno	R_m [MPa]	Whisker	R_m [MPa]
B	3 500	Al_2O_3	21 000
W	2 800	B_4C	14 000
Be	1 400	SiC	21 000
C	2 100	C	20 000
Borsic	3 150	Fe	13 400

Řešení

Teorie: 51,52,53 **Řešené příklady:** 140,141,142

Tahák

Využití analytického nástroje Popisná statistika v Excelu.

1. Data -> Analýza dat -> Popisná statistika
Vstupní oblast: Hodnoty
zaškrtneme Celkový přehled

220 - Popisná analýza

Zadání Termostat je nastaven na 14 stupňů Celsia. Bylo provedeno 11 kontrolních měření a zjištěny následující hodnoty:

13,4	13,2	13,4	13,6	14,5	13,0	14,3	13,8	14,3	14,0	14,1
------	------	------	------	------	------	------	------	------	------	------

Na základě popisné analýzy rozhodněte, zda je třeba provést opravu nastavení termostatu.

Řešení

Teorie: 50,51,52,53 **Řešené příklady:** 140,141,142

Příklady – Odhady parametrů

222 - Interval spolehlivosti pro střední hodnotu

Zadání

1. Měřil se průměr hřídele na 250 součástkách. Předpokládáme normální rozdělení souboru. Z výsledků se určil výběrový průměr a výběrová disperze: $\bar{x} = 995,6$ a $s^2 = 134,7$. Určete interval spolehlivosti pro střední hodnotu na hladině významnosti $\alpha = 0,05$.
2. Byla měřena délka trvání určitého procesu. Z 12 měření byla zjištěna průměrná doba trvání procesu 44 s. Určete interval spolehlivosti na hladině významnosti $\alpha = 0,1$ a $\alpha = 0,05$ pro očekávanou délku procesu za předpokladu normálního rozdělení s hodnotou rozptylu 16.

Řešení

Teorie: 56,57 Řešené příklady: 144,145,146

Tahák

Interval spolehlivosti pro μ

$$\mu \in \langle \bar{x} \pm \delta \rangle$$

1. znám σ^2

$$\delta = \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}$$

2. neznám σ^2

$$\delta = \frac{s}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2}}(n-1)$$

EXCEL:

kvantil normovaného normálního rozdělení

$$z_{1-\frac{\alpha}{2}}$$

$$=\text{NORM.S.INV}\left(1 - \frac{\alpha}{2}\right)$$

kvantil t-rozdělení s $n - 1$ stupni volnosti

$$t_{1-\frac{\alpha}{2}}(n-1)$$

$$=\text{T.INV}\left(1 - \frac{\alpha}{2}; n-1\right)$$

Využití analytického nástroje Popisná analýza v Excelu.

- určení hodnoty δ z výběrového souboru, neznám σ^2

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Popisná analýza

2. Vstupní oblat - data; Výstupní oblast - libovolná buňka

3. Zaškrtneme Hladina spolehlivosti pro stř. hodnotu a zvolíme hladinu spolehlivosti $(1 - \alpha) \cdot 100$

223 - Interval spolehlivosti pro střední hodnotu

Zadání

1. Při měření kapacity sady kondenzátorů bylo provedeno 10 měření s výsledky 152, 156, 148, 153, 150, 156, 140, 155, 145 a 148. Určete interval spolehlivosti na hladině významnosti $\alpha = 0,05$ pro kapacitu těchto kondenzátorů.
2. Hypermarket chce pro zkvalitnění služeb poskytovaných zákazníkům zkrátit dobu jejich čekání u pokladen. Náhodně bylo vybráno 10 zákazníků a byla změřena doba jejich čekání u pokladny. Výsledky šetření (v sekundách): 50, 65, 30, 45, 35, 55, 70, 65, 50, 52. Jaká je horní hranice doby čekání, která nebude s pravděpodobností 0,95 překročena?

Řešení

Teorie: 56,57 Řešené příklady: 144,145,146

Tahák

Oboustranný interval spolehlivosti pro μ
 $\mu \in \langle \bar{x} \pm \delta \rangle$

Jednostranný interval spolehlivosti pro μ
 $\mu \in (-\infty; \bar{x} + \delta)$
 $\mu \in \langle \bar{x} - \delta; \infty \rangle$

EXCEL:

Využití analytického nástroje Popisná analýza v Excelu.

- určení hodnoty δ z výběrového souboru, neznám σ^2

1. Data $\rightarrow$ Analýza dat $\rightarrow$ Popisná analýza
2. Vstupní oblat - data; Výstupní oblast - libovolná buňka
3. Zaškrtneme Hladina spolehlivosti pro stř. hodnotu a zvolíme hladinu spolehlivosti pro oboustranný IS $(1 - \alpha) \cdot 100$ a pro jednostranný $(1 - 2\alpha) \cdot 100$

224 - Interval spolehlivosti pro rozptyl, směrodatnou odchylku

Zadání

- Při zjišťování přesnosti nově zavedené metody pro stanovení obsahu manganu v oceli bylo rozhodnuto provést 4 nezávislá měření. Stanovte odhad pro σ na hladině významnosti 0,1, když výsledky měření byly: 0,31 %, 0,3 %, 0,29 %, 0,32 %.
- Určete intervalový odhad rozptylu a směrodatné odchylky na hladině významnosti 5 % pro následující hodnoty 606, 1249, 267, 44, 510, 340, 109, 1957, 463, 801, 1086, 169, 233, 1734, 1458, 80, 1023, 2736, 917 a 459.

Řešení

Teorie: **58** Řešené příklady: **147**

Tahák

Interval spolehlivosti pro σ^2

$$\sigma^2 \in \left\langle \frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}; \frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right\rangle$$

Interval spolehlivosti pro σ

$$\sigma \in \left\langle \sqrt{\frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}; \sqrt{\frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} \right\rangle$$

EXCEL:

kvantil chí-kvadrát rozdělení s $n - 1$ stupni volnosti

$$1. \chi^2_{1-\frac{\alpha}{2}}(n-1) \\ =\text{CHISQ.INV}\left(1 - \frac{\alpha}{2}; n-1\right)$$

$$2. \chi^2_{\frac{\alpha}{2}}(n-1) \\ =\text{CHISQ.INV}\left(\frac{\alpha}{2}; n-1\right)$$

Příklady – Testování hypotéz

226 - Hypotéza o střední hodnotě

Zadání Pevnost materiálu se zjistila na 9 náhodně vybraných vzorcích materiálu a byly naměřeny hodnoty pevnosti [MPa]:

630	630	660	670	690	700	700	710	710
-----	-----	-----	-----	-----	-----	-----	-----	-----

Ověřte platnost tvrzení, že průměrná pevnost materiálu v dodávkách (v základním souboru) je 660 MPa. Ověření proveďte na hladině významnosti 5 %.

Řešení

Teorie: 61 Řešené příklady: 149

Tahák

Jednovýběrový t-test

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

výpočet testové hodnoty

$$T = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}$$

výpočet kritické hodnoty

$$t_{\text{krit}} = t_{1-\frac{\alpha}{2}}(n-1)$$

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

EXCEL:

kvantil t-rozdělení s $n - 1$ stupni volnosti

$$t_{1-\frac{\alpha}{2}}(n-1)$$

$$=T.INV\left(1 - \frac{\alpha}{2}; n - 1\right)$$

227 - Hypotéza o rozptylu

Zadání U kulečnickových koulí záleží především na tom, aby koule byly stejně těžké. Norma pro směrodatnou odchylku jejich váhy připouští maximální hodnotu této směrodatné odchylky $\sigma = 3$ g. Na základě vah náhodně vybraných 8 koulí z dané sady rozhodněte (na hladině významnosti 1 %), zda celá tato sada jako celek odpovídá normě. Váhy vybraných 8 koulí jsou v gramech:

202	198	203	196	201	203	197	204
-----	-----	-----	-----	-----	-----	-----	-----

Řešení

Teorie: 62 Řešené příklady: 150

Tahák

Test o rozptylu

$$H_0 : \sigma^2 = \sigma_0^2$$

$$H_1 : \sigma^2 \neq \sigma_0^2 (\sigma^2 > \sigma_0^2; \sigma^2 < \sigma_0^2)$$

výpočet testové hodnoty

$$\chi^2 = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

výpočet kritických hodnot

oboustranný test:

$$\chi_{\frac{\alpha}{2}}^2 (n-1); \chi_{1-\frac{\alpha}{2}}^2 (n-1)$$

jednostranný test:

$$\chi_{\alpha}^2 (n-1) \text{ pro } H_1 : \sigma^2 < \sigma_0^2$$

$$\chi_{1-\alpha}^2 (n-1) \text{ pro } H_1 : \sigma^2 > \sigma_0^2$$

Závěr:

oboustranný test:

$$\chi^2 \in \left\langle \chi_{\frac{\alpha}{2}}^2 (n-1); \chi_{1-\frac{\alpha}{2}}^2 (n-1) \right\rangle \Rightarrow H_0 \text{ nezamítáme}$$

jednostranný test:

$$\chi^2 > \chi_{\alpha}^2 (n-1) \text{ pro } H_1 : \sigma^2 < \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

$$\chi^2 < \chi_{1-\alpha}^2 (n-1) \text{ pro } H_1 : \sigma^2 > \sigma_0^2 \Rightarrow H_0 \text{ nezamítáme}$$

EXCEL:

kvantil Chí-kvadrát rozdělení s (n - 1) stupni volnosti $\chi_{1-\frac{\alpha}{2}}^2 (n-1)$

$$=\text{CHISQ.INV} \left(1 - \frac{\alpha}{2}; n-1 \right)$$

228 - Hypotéza o rovnosti středních hodnot

Zadání Výrobek se zhotovuje dvěma různými technologickými postupy. Kontrolním měřením se na náhodných výběrech zjistily parametry zhotovených výrobků dle tabulky. Na hladině významnosti 5 % posuďte, zda se těmito technologickými postupy dosahuje stejné průměrné hodnoty parametru.

postup A	4	16	18	16	14		
postup B	6	17	17	19	18	16	15

Řešení

Teorie: 63,64,65 Řešené příklady: 151,152

Tahák

Dvouvýběrový F-test

EXCEL:

Využití analytického nástroje v Excelu.

1. Data → Analýza dat → Dvouvýběrový F-test pro rozptyl
2. Vstup - data (1. soubor zvolíme ten s větším rozptylem); Výstupní oblast - libovolná buňka
3. Alfa: zadáme $\frac{\alpha}{2}$

Závěr: $F < F_{\text{krit}} \Rightarrow H_0$ nezamítáme

Dvouvýběrový t-test

výpočet testové hodnoty a kritické hodnoty

1. platí $\sigma_1^2 = \sigma_2^2$, použijeme t-test s rovností rozptylů
2. neplatí $\sigma_1^2 = \sigma_2^2$, použijeme t-test s nerovností rozptylů

EXCEL:

Využití analytického nástroje v Excelu.

1. Data → Analýza dat → Dvouvýběrový t-test s rovností (nerovností) rozptylů
2. Vstup - data; Výstupní oblast - libovolná buňka
3. Alfa: zadáme hladinu významnosti

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

229 - Párový t-test

Zadání Svaly horní končetiny byly cyklicky namáhány až do úplného vypovězení funkce. Hmotnost závaží byla konstantní a délka přestávky mezi sériemi byla 30 sekund. Otestujte, zda jsou obě končetiny stejně silné.

končetina P	20	7	3	2	2	2	1	1	1	0	0
končetina L	19	6	3	3	2	2	2	1	1	1	0

Řešení **Teorie: 66 Řešené příklady: 153**

Tahák

Párový t-test

- EXCEL:**
Využití analytického nástroje v Excelu.
1. Data → Analýza dat → Dvouvýběrový párový t-test na střední hodnotu
 2. Vstup - data; Výstupní oblast - libovolná buňka
 3. Alfa: zadáme hladinu významnosti

Závěr: $|T| < t_{\text{krit}} \Rightarrow H_0$ nezamítáme

230 - Testy dobré shody

Zadání Provéřte na hladině významnosti $\alpha = 0,05$, zda má soubor rovnoměrné rozdělení, když pro náhodný výběr byly zjištěny tyto četnosti jednotlivých tříd

10	21	0	8	12	6	8	13	11	11
----	----	---	---	----	---	---	----	----	----

Řešení

Teorie: 67,68 Řešené příklady: 154,155

Tahák

Nulová hypotéza:

ZS má očekávané rozdělení pravděpodobnosti, tj. četnosti f_{ei} a f_{oi} pro $i = 1, \dots, n$ se liší pouze náhodně.

χ^2 – test

výpočet testové hodnoty

$$\chi^2 = \sum_{i=1}^k \frac{(f_{ei} - f_{oi})^2}{f_{oi}}$$

výpočet kritické hodnoty

$$\chi_{\text{krit}}^2 = \chi_{1-\alpha}^2(k - s - 1)$$

EXCEL:

kvantil chí-kvadrát rozdělení s $k - s - 1$ stupni volnosti

$$=\text{CHISQ.INV}(1 - \alpha; k - s - 1)$$

Závěr: $\chi^2 < \chi_{\text{krit}}^2 \Rightarrow H_0$ nezamítáme

Kolmogorovův-Smirnovův test

výpočet testové hodnoty

$$D_1 = \frac{1}{n} \max |F_{ei} - F_{oi}|$$

výpočet kritické hodnoty

$$\alpha = 0,05 \Rightarrow D_{1;0,05}(n) = \frac{1,36}{\sqrt{n}}$$

Závěr: $D_1 < D_{1;\alpha}(n) \Rightarrow H_0$ nezamítáme

231 - Testy extrémních hodnot

Zadání Zjistěte, zda nejmenší hodnota v daném souboru je extrémně odchýlena od ostatních.

111,2	112,4	114,6	95,4	105,6	107,7	108,3	111,8	115,3	109,1
-------	-------	-------	------	-------	-------	-------	-------	-------	-------

Řešení **Teorie: 69,70 Řešené příklady: 156**

Tahák

Nulová hypotéza:

H_0 : Hodnota x_1 , resp. x_n (největší hodnota) se neliší významně od ostatních hodnot souboru.

Dixonův test

výpočet testové hodnoty

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1}$$

$$Q_n = \frac{x_n - x_{n-1}}{x_n - x_1}$$

výpočet kritické hodnoty

Kritické hodnoty $Q_{1;\alpha}$, resp. $Q_{n;\alpha}$ jsou tabelovány.

Závěr:

$$Q_1 < Q_{1;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

$$Q_n < Q_{n;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

Grubbsův test

výpočet testové hodnoty

$$T_1 = \frac{\bar{x} - x_1}{s}$$

$$T_n = \frac{x_n - \bar{x}}{s}$$

výpočet kritické hodnoty

Kritické hodnoty $T_{1;\alpha}$, resp. $T_{n;\alpha}$ jsou tabelovány.

Závěr:

$$T_1 < T_{1;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

$$T_n < T_{n;\alpha} \Rightarrow H_0 \text{ nezamítáme}$$

Příklady – Lineární regrese

233 - Regresní model

Zadání Při zkoumání výtoku kapaliny šterbinou byly získány údaje pro bezrozměrný součinitel výtoku uvedené v tabulce v závislosti na výšce h .

Výška [m]	0,164	0,328	0,656	0,984	1,312	1,640
Součinitel výtoku	0,448	0,432	0,421	0,417	0,414	0,412

Nalezněte parametry regresní přímky. Rozhodněte, zda se jedná o kvalitní model a zda existuje mezi výškou a součinitelem výtoku závislost. Určete odhad hodnoty součinitele výtoku při výšce šterbiny 1 m.

Řešení **Teorie: 74,75,76,77 Řešené příklady: 158,159,160**

Tahák

Regresní přímka

$$y = ax + b$$

Kvalita modelu

R^2 ... koeficient determinace

Míra závislosti

$R = \sqrt{R^2}$... korelační koeficient

EXCEL: =CORREL(matice1;matice1)

Tvorba modelu pomocí Excelu

Vložení -> Bodový graf -> Přidat spojnici trendu

zvolíme Lineární

zaškrtneme Zobrazit rovnici v grafu

zaškrtneme Zobrazit hodnotu spolehlivosti R

234 - Regresní a korelační analýza

Zadání Při měření tvrdosti oceli v závislosti na hmotnostním obsahu uhlíku byly po tepelném zpracování u 9 typů vzorků zjištěny následující střední hodnoty tvrdosti.

Tvrdost [HRC]	35	44	49	57	68	76	82	85	83
Obsah C [%]	0,20	0,35	0,50	0,65	0,80	0,95	1,10	1,25	1,40

Posuďte, zda se jedná o závislé náhodné veličiny a navrhnete vhodnou regresní funkci.

Řešení **Teorie: 76,77 Řešené příklady: 159,160**

Tahák

Kvalita modelu
 R^2 ... koeficient determinace

Míra závislosti
 $R = \sqrt{R^2}$... korelační koeficient
EXCEL: =CORREL(matice1;matice1)

Tvorba modelu pomocí Excelu
Vložení -> Bodový graf -> Přidat spojnicí trendu
zaškrtneme Zobrazit rovnici v grafu
zaškrtneme Zobrazit hodnotu spolehlivosti R

235 - Verifikace modelu

Zadání Odhadněte parametry regresní přímky, na hladině významnosti 5 % otestujte závislost teplotního součinitele na měrném elektrickém odporu a kvalitu lineárního modelu.

Materiál	Měrný el. odpor [$\mu\Omega\text{m}$]	Teplotní součinitel odporu [K^{-1}]
Zlato	0,0230	0,00370
Měď	0,0178	0,00420
Stříbro	0,0163	0,00400
Hliník	0,0285	0,00400
Rtut'	0,9580	0,00090
Železo	0,1000	0,00550
Chrom	1,1000	0,00025
Nikelin	0,4000	0,00011
Wolfram	0,0530	0,00440

Řešení

Teorie: 75,76,77 Řešené příklady: 158,159,160

Tahák

Dílčí t-test

$$H_0 : a_i = 0$$

$$H_1 : a_i \neq 0$$

Celkový F-test

$$H_0 : a_0 = a_2 = \dots = a_k = 0$$

$$H_1 : \neg H_0$$

Tvorba modelu pomocí Excelu

Vložení -> Bodový graf -> Přidat spojnicí trendu

zaškrtneme Zobrazit rovnici v grafu

zaškrtneme Zobrazit hodnotu spolehlivosti R

Testy pro regresní přímku v Excelu

Data -> Analýza dat -> Regrese

zaškrtneme Hladina spolehlivosti

236 - Regresní model

Zadání Vyšetřujeme souvislost znečištění ovzduší oxidy síry a výskyt alergií u dětí do 15 let. V 8 obcích jsme zjistili průměrnou koncentraci (v $\mu\text{g}/\text{m}^3$) a procento dětí s alergií:

koncentrace	0,10	0,12	0,21	0,15	0,05	0,10	0,10	0,12
procento dětí	15	20	28	11	7	18	20	13

Souvisejí spolu tyto dva faktory?

Řešení

Teorie: 75,76,77 **Řešené příklady:** 158,159

Literatura

- [1] ANDĚL, Jiří. *Statistické metody*. Matfyzpress Praha, 1993, 246 s.
- [2] CALDA, Emil a Václav DUPAČ. *Matematika pro gymnázia - Kombinatorika, pravděpodobnost a statistika*. Prometheus Praha, 1994, ISBN 80-85849-10-0.
- [3] DUPAČ, Václav a Marie HUŠKOVÁ. *Pravděpodobnost a matematická statistika*. Karolinum, Praha, 1999.
- [4] HAJKR, Oldřich, Pavel HRADECKÝ, Anna MADRYOVÁ a Matěj TURČAN. *Teorie statistiky*. Ostrava: VŠB - Technická univerzita, 1988, 267 s.
- [5] HEBÁK, Petr a Jiří HUSTOPECKÝ. *Průvodce moderními statistickými metodami*. SNTL Praha, 1990, 293 s., ISBN 80-03-00534-5.
- [6] HEBÁK, Petr a Jana KAHOUNOVÁ. *Počet pravděpodobnosti v příkladech*. Informatorium, Praha, 2005, ISBN 80-733-040-7.
- [7] MAREK, Luboš. *Statistika v příkladech*. 1. vyd. Professional Publishing, Praha, 2013, 404 s., ISBN 978-80-7431-118-5.
- [8] MICHÁLEK, Jaroslav. *Úvod do teorie pravděpodobnosti a matematické statistiky*. Státní pedagogické nakladatelství, Praha, 1984, 204 s.
- [9] OTIPKA, Petr a Vladislav ŠMAJSTRLA. *Pravděpodobnost a statistika*. Ostrava: VŠB - Technická univerzita, 2006, ISBN 80-248-1194-4.
- [10] PAVELKA, Lubomír a Jarmila DOLEŽALOVÁ. *Pravděpodobnost a statistika*. Ostrava: VŠB - Technická univerzita, 2003, ISBN 80-7078-976-X.
- [11] SCHREIBEROVÁ, Petra. *Ukázka využití GeoGebry při řešení základních statistických problémů*. Sborník 3mi, 2013, ISBN 978-80-248-3233-3.
- [12] SCHREIBEROVÁ, Petra. *Výuka náhodných veličin s využitím GeoGebry*. Sborník 3mi, 2014, ISBN 978-80-248-3611-9.
- [13] ZVÁRA, Karel a Josef ŠTĚPÁN. *Pravděpodobnost a matematická statistika*. Matfyzpress, Praha, 2001.