

Vysoká škola technická a ekonomická v Českých Budějovicích

Katedra informatiky a přírodních věd

Studijní materiály

Základy popisné statistiky – sbírka příkladů

Ing. Martin Telecký, Ph.D.

2021

Učební materiály poskytují základní vymezené praktické příklady, včetně řešení v rámci problematiky popisné statistiky. Učební materiály slouží pro studenty technických a ekonomických oborů.

Za odbornou a jazykovou stránku odpovídá autor práce.

Zpracováno na základě interní grantové soutěže: IGS 8210 - Podpora online vzdělávání v základních kurzech – fyzika, informatika, matematika a statistika.

Kapitola 1: Problematika třídění dat v rámci absolutních a relativních četností

Kapitola se zabývá funkcí:

- třídění primárních či sekundárních dat do tzv. absolutních a relativních četností,
- vymezení správného počtu disjunktních intervalů,
- vymezení vhodné šířky/délky jednoho disjunktního intervalu,
- tvorby grafického vyjádření absolutních a relativních četností prostřednictvím histogramu a polygonu.

Příklad 1

U 6 studentů na Vysoké škole technické a ekonomické, oboru Technologie dopravy a přepravy, jsme zjistili jejich počet chyb za hodinu během zkouškového testu. Statistický soubor obsahuje údaje o počtu chyb v testu za hodinu. Zjistěte počet a délku disjunktních intervalů. Zjistěte absolutní a relativní četnosti, kumulované absolutní a relativní četnosti. Dále vytvořte histogram a polygon u absolutních četností. Graficky znázorněte relativní četnosti tohoto příkladu.

Student	počet chyb v testu za hodinu
1	25
2	29
3	10
4	27
5	14
6	8

Zdroj: autor

Výsledek

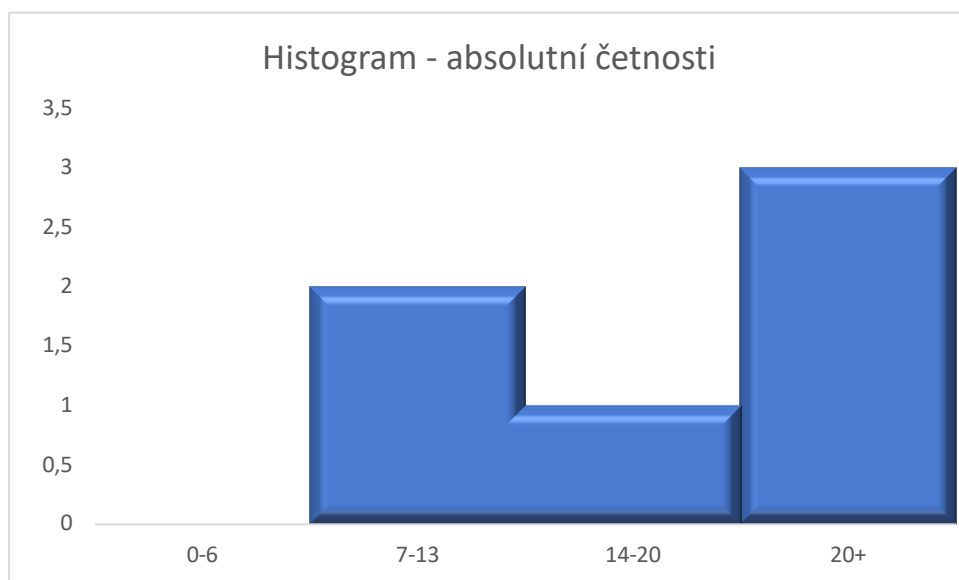
- Sturgesovo pravidlo = $1 + 3,3 * \log(n) = 1 + 3,3 * \log(6) =$ **cca 4 disjunktní intervaly**
- Šíře/délka intervalu = **cca po 6**

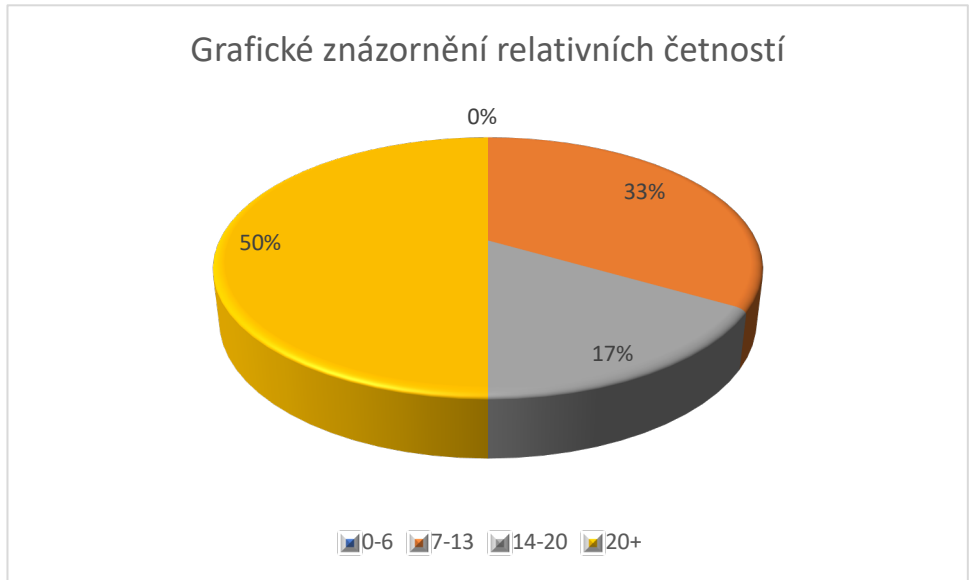
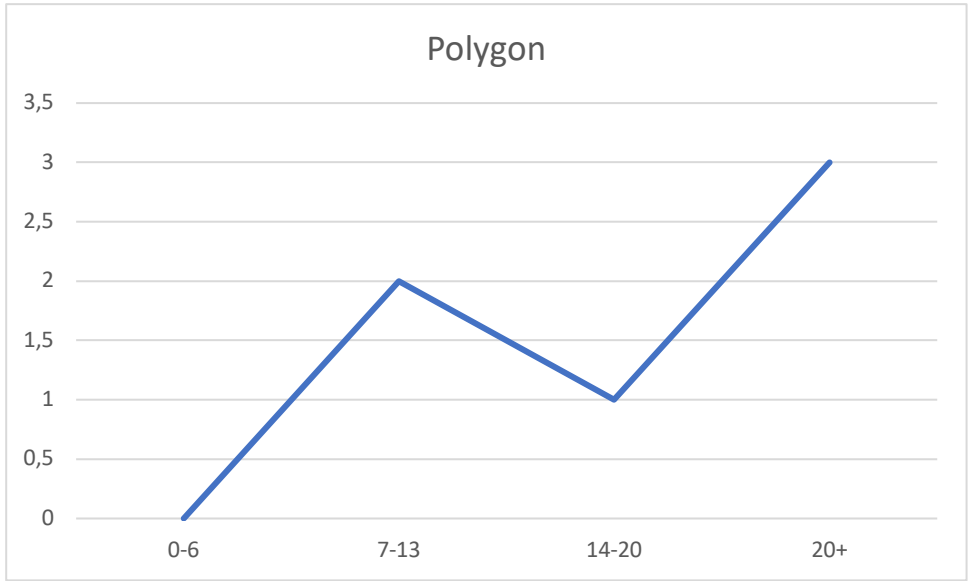
Výsledná tabulka absolutních a relativních četností:

intervaly		absolutní četnosti	relativní četnosti
0-	6	0	0%
7-	13	2	33%
14-	20	1	17%
20+		3	50%
Celkem		6	100%

Výsledná tabulka kumulovaných absolutních a relativních četností:

intervaly		kumulované absolutní četnosti	kumulované relativní četnosti
0-	6	0	0%
7-	13	2	33%
14-	20	3	50%
20+		6	100%





Kapitola 2: Základní charakteristiky úrovně a variability

Kapitola vymezuje základní charakteristiky úrovně či polohy a dále variability. Cílem kapitoly je efektivně aplikovat vhodný statistický ukazatel jako je například:

- prostý aritmetický průměr,
- medián,
- modus,
- vážený aritmetický průměr,
- geometrický průměr,
- harmonický průměr,
- populační rozptyl,
- výběrový rozptyl,
- populační směrodatná odchylka,
- výběrová směrodatná odchylka,
- variační koeficient.

Příklad 1

Čtyři společnosti vyrábějí identický výrobek. První společnosti trvá vyrobení výrobku 5 minut, druhé společnosti 4 minuty, třetí společnosti 6 minut a čtvrté společnosti 4 minuty.

Stanovte průměrnou dobu zhotovení jednoho výrobku jednou společností? Kolik výrobků vyrobí společnosti celkově za jednu hodinu?

Výsledek

- **4,61538 min**

Příklad 2

K dispozici máme hodnoty 2, 3, 5, 8, 6. Stanovte prostý aritmetický průměr, geometrický průměr a harmonický průměr.

Výsledek

- Prostý aritmetický průměr = 4,8
- Geometrický průměr = 4,2823

- Harmonický průměr = 3,7736

Příklad 3

Byla stanovena průměrná mzda 26355 Kč. Určete průměrnou mzdu, pokud

- a) každý pracovník dostane přidáno 600 Kč,
- b) každý pracovník dostane 2násobek platu,
- c) každý pracovník dostane přidáno 15 % ze stávajícího platu.

Výsledek

- a) $26355 + 600 = 26955$ Kč
- b) $26355 * 2 = 52710$ Kč
- c) $26355 * 1,15 = 30308,25$ Kč

Příklad 4

Stanovte medián z následujících dat: 18, 21, 22, 17, 15, 23, 12, 14.

Výsledek

- 17,5

Příklad 5

Stanovte modus z následujících dat: 18, 21, 22, 17, 15, 23, 12, 14.

Výsledek

- nelze stanovit

Příklad 6

Byla stanovena průměrná mzda 26355 Kč a rozptyl mezd 810000. Určete směrodatnou odchylku mezd, pokud

- a) každý pracovník dostane přidáno 600 Kč,
- b) každý pracovník dostane 2násobek platu,
- c) každý pracovník dostane přidáno 15 % ze stávajícího platu.

Výsledek

- a) směrodatná odchylka se nezmění
- b) 1800
- c) 1035

Příklad 7

K dispozici máme dva projekty s různými variantami příjmů.

U projektu A:

optimistický odhad	1.000.000 Kč	s pravděpodobností 0,6
pesimistický odhad	700.000 Kč	s pravděpodobností 0,3
neutrální odhad	850.000 Kč	s pravděpodobností 0,1

U projektu B očekáváme:

optimistický odhad	1.200.000 Kč	s pravděpodobností 0,4
pesimistický odhad	800.000 Kč	s pravděpodobností 0,5
neutrální odhad	900.000 Kč	s pravděpodobností 0,1

Jaký projekt je pro Vás výhodnější z hlediska očekávaných příjmů? Jaký projekt je pro Vás výhodnější z hlediska zahrnutí faktoru rizika? Stanovte variační koeficient a rozhodněte, který projekt vyberete.

Výsledek

Nezapomeňte, že součet pravděpodobností u jednotlivých projektů musí dát součet 1.

Pravděpodobnosti se stanoví na základě subjektivního či objektivního principu stanovení pravděpodobnosti, například metodou kvantilů.

Projekt A:

- očekávaný příjem: 895000 Kč
- očekávané riziko: 135000
- variační koeficient: 0,150838 (15,0838 %)

Projekt B:

- očekávaný příjem: 970000Kč
- očekávané riziko: 190000
- variační koeficient: 0,19588 (19,588 %)

Projekt B je výhodnější z hlediska očekávaného příjmu. Projekt A je výhodnější z hlediska zohlednění faktoru rizika a zároveň výhodnější z hlediska variačního koeficientu.

Popisná statistika dokáže zohlednit v určitých mezích hodnocení efektivnosti investic. Avšak ve spojení s finančním řízením projektů a statistiky dokážeme získat velmi cenné informace a efektivně tak rozhodnout o budoucím vývoji investice v krátkodobém, střednědobém a dlouhodobém horizontu.

V rámci finančního řízení je vhodné aplikovat:

- čistou současnou hodnotu,
- metodu doby návratnosti,
- ekonomickou přidanou hodnotu (EVA),
- tržní přidanou hodnotu,
- metodu průměrných nákladů celkového kapitálu.

V rámci aplikované matematiky jsou vhodné metody analýzy obalu dat (metoda DEA). Tedy:

- model CCR,
- model BCC – v rámci lineárního programování.

Kapitola 3: Základy klasické a statistické pravděpodobnosti

Klasická a statistická pravděpodobnost jsou základními stavebními prvky pro pochopení detailnější statistické analýzy v podobě dalších pravděpodobnostních rozdělení, jejich parametrů a vypovídací schopnosti. Příklady znázorňují aplikaci:

- klasické pravděpodobnosti,
- statistické pravděpodobnosti,
- úplné pravděpodobnosti,
- podmíněné pravděpodobnosti,
- průniku a sjednocení dvou jevů,
- Bayesovy pravděpodobnosti – apriorní a aposteriorní pravděpodobnost.

Příklad 1

Jaká je pravděpodobnost, že na hrací kostce padne číslo 6?

Výsledek

- $1/6$

Příklad 2

Jaká je pravděpodobnost, že na dvou hracích kostkách padne číslo 6?

Výsledek

- $1/36$

Příklad 3

Jaká je pravděpodobnost, že na dvou hracích kostkách padne stejné číslo?

Výsledek

- $6/36 = 1/6$

Příklad 4

Jaká je pravděpodobnost, že na dvou hracích kostkách padne součet 7?

Výsledek

- $6/36 = 1/6$

Příklad 5

Jaká je pravděpodobnost, že na dvou hracích kostkách padne součet 14?

Výsledek

- Na dvou hracích kostkách lze maximálně dosáhnout součtu 12

Příklad 6

Ve váčku je celkově 50 kuliček různých barev. Bílých kuliček je 15, modrých je 10, černých je 20 a 5 žlutých kuliček. Náhodně vybereme tři kuličky (bez vracení). Jaká je pravděpodobnost, že obdržíme kuličky v následujícím barevném pořadí:

- a) P (bílá, bílá, žlutá),
- b) P (černá, bílá, žlutá)?

Výsledek

- a) 0,00898571429 (0,898571429 %)
- b) 0,01275510204 (1,275510204 %)

Příklad 7

V botníku je 13 párů bot. Polovina párů má černou barvu a druhá polovina párů má bílou barvu. Náhodně vytáhneme dvě boty. Jaká je pravděpodobnost, že vybereme boty stejné barvy?

Výsledek

- 0,24 (24 %)

Příklad 8

V garáži v krabici máme 150 šroubů. Víme při kontrole, že mezi 150 šrouby je 10 šroubů špatných. Náhodně vybereme 20 šroubů. Jaká je pravděpodobnost, že výběr nebude obsahovat žádný špatný šroub?

Výsledek

- 0,22778 (22,778 %)

Příklad 9

V rámci logistiky dodávky zboží máme k dispozici dva dodavatele. První dodavatel dodává 30 % zboží a z toho je 60 % první jakosti. Druhý dodavatel dodává 70 % zboží a z toho je 75 % první jakosti. Jaká je pravděpodobnost, že náhodně nakoupený výrobek je druhé jakosti od druhého dodavatele?

Výsledek

- 0,593220339 (59,3220339 %)

Kapitola 4: Pravděpodobnostní rozdělení diskrétní náhodné veličiny

Kapitola se zabývá problematikou pravděpodobnostního rozdělení diskrétní náhodné veličiny.

Věnována je pozornost:

- binomickému rozdělení,
- Poissonovo rozdělení,
- hypergeometrickému rozdělení, včetně jejich základních parametrů,
- grafickému znázornění pravděpodobnostní a distribuční funkce dané diskrétní náhodné veličiny,
- Zákonu velkých čísel, včetně výběrového prostoru pravděpodobnosti,
- základním vlastnostem pravděpodobnostní a distribuční funkce diskrétní náhodné veličiny.

Příklad 1

Jaká je pravděpodobnost, že při 5 hodech hrací kostkou padne:

- žádná šestka,
- jedenkrát šestka,
- dvakrát šestka,
- tříkrát šestka,
- čtyřikrát šestka,
- pětkrát šestka,
- maximálně třikrát šestka,
- maximálně čtyřikrát šestka?

Graficky znázorněte pravděpodobnostní a distribuční funkci této náhodné veličiny.

Výsledek

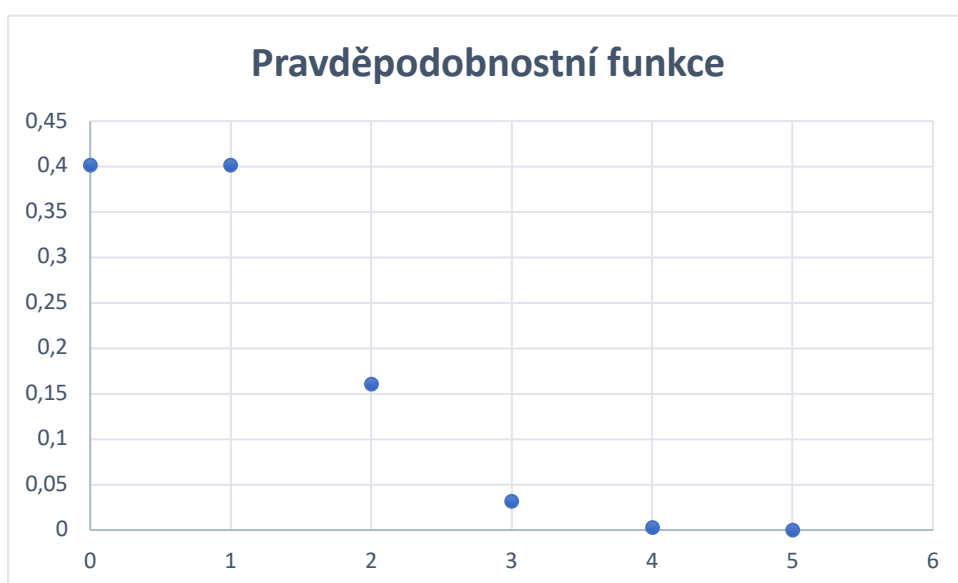
- 0,401878
- 0,401878
- 0,160751
- 0,03215
- 0,003215
- 0,000129
- 0,996656379

h) 0,999871399

Následující tabulka představuje výběrový prostor náhodné veličiny pro potřeby grafického znázornění pravděpodobnostní a distribuční funkce.

x_i	0	1	2	3	4	5
$P(x)$	0,401878	0,401878	0,160751	0,03215	0,003215	0,000129

Pravděpodobnostní funkce:



Distribuční funkce:



Příklad 2

V domácnosti dochází v průběhu času k výpadkům elektřiny. Z minulosti je známo, že v dané domácnosti dochází v průměru ke 3 výpadkům elektřiny za 8 měsíců. Určete, jaká je pravděpodobnost, že za 2 roky dojde v domácnosti maximálně k 6 výpadkům elektřiny.

Výsledek

- 0,20678084 (20,678084 %)

Příklad 3

Obchodník odebírá od dodavatele židle v pravidelných dodávkách po 150 kusech. Po převzetí každé dodávky vybere obchodník 20 židlí ke kontrole kvality. Obchodník předpokládá, že v jedné dodávce židlí se nachází 7 vadných židlí. Určete, jaká je pravděpodobnost, že v dodávce budou maximálně 3 vadné výrobky?

Výsledek

- 0,044028 (4,4028 %)

Kapitola 5: Pravděpodobnostní rozdělení spojité náhodné veličiny

Věnována je pozornost spojité náhodné veličině, která je zakládána na tzv. hustotě pravděpodobnosti vybraných pravděpodobnostních rozdělení, včetně jejich základních parametrů. Právě tato pravděpodobnostní rozdělení tvoří základ většiny statistických testů.

Příklady se zaměřují na problematiku:

- normálního rozdělení,
- normovaného normálního rozdělení,
- Studentova rozdělení,
- stanovení distribuční funkce normálního a normovaného normálního rozdělení,
- určení příslušných kvantilů normálního rozdělení, normovaného normálního rozdělení a Studentova rozdělení.

Výše uvedené faktory slouží jako důležité části odhadu populačního průměru v rámci problematiky teorie odhadu neboli intervalů spolehlivosti.

Pro požadované výpočty musíme mít k dispozici statistické tabulky (viz příloha).

Příklad 1

Zjistěte pravděpodobnost náhodné veličiny U , která má normované normální rozdělení.

$$P(U < 2)$$

$$P(U < -2)$$

$$P(U = 1)$$

$$F(1)$$

$$F(-1)$$

$$P(1 < U < 2)$$

$$P(-1 < U < 2)$$

Výsledek

- $P(U < 2) = F(2) = \mathbf{0,97725}$
- $P(U < -2) = 1 - F(2) = 1 - 0,97725 = \mathbf{0,02275}$
- $P(U = 1) = \mathbf{0}$
- $F(1) = \mathbf{0,84134}$
- $F(-1) = 1 - F(1) = 1 - 0,84134 = \mathbf{0,15866}$
- $P(1 < U < 2) = F(2) - F(1) = 0,97725 - 0,84134 = \mathbf{0,13591}$
- $P(-1 < U < 2) = F(2) - F(-1) = F(2) - (1 - F(1)) = 0,97725 - (1 - 0,84134) = \mathbf{0,81859}$

Příklad 2

Zjistěte kvantil příslušného normovaného normálního rozdělení.

$u_{0,95}$

$u_{0,75}$

$u_{0,15}$

$u_{0,35}$

Výsledek

- $u_{0,95} = \mathbf{1,65}$
- $u_{0,75} = \mathbf{0,68}$
- $u_{0,15} = \mathbf{-1,036}$
- $u_{0,35} = \mathbf{-0,385}$

Příklad 3

Zjistěte pravděpodobnost náhodné veličiny X , která má normální rozdělení se střední hodnotou 9 a rozptylem 25?

$$P(X < 10)$$

$$P(X < -1)$$

$$P(10 < X < 19)$$

$$P(-1 < X < -0,5)$$

Výsledek

- $P(X < 10) = P\left(U < \frac{10-9}{5}\right) = P(U < 0,2) = \mathbf{0,57926}$
- $P(X < -1) = P\left(U < \frac{-1-9}{5}\right) = P(U < -2) = 1 - 0,97725 = \mathbf{0,02275}$
- $P(10 < X < 19) = P\left(\frac{10-9}{5} < U < \frac{19-9}{5}\right) = P(0,2 < U < 2) = F(2) - F(0,2) = 0,97725 - 0,57926 = \mathbf{0,39799}$
- $P(-1 < X < -0,5) = P\left(\frac{-1-9}{5} < U < \frac{-0,5-9}{5}\right) = P(-2 < U < -1,9) = F(-1,9) - F(-2) = 1 - F(1,9) - (1 - F(2)) = 0,02872 - 0,02275 = \mathbf{0,00597}$

Příklad 4

Zjistěte kvantil příslušného normálního rozdělení.

$$x_{0,95}$$

$$x_{0,75}$$

$$x_{0,15}$$

$$x_{0,35}$$

Výsledek

- $x_{0,95} = 17,25$
- $x_{0,75} = 12,4$
- $x_{0,15} = 3,82$
- $x_{0,35} = 7,075$

Příklad 5

Zjistěte kvantil příslušného studentova rozdělení.

$t_{0,9} [10]$

$t_{0,95} [10]$

$t_{0,975} [10]$

$t_{0,95} [14]$

$t_{0,975} [21]$

Výsledek

- $t_{0,9} [10] = 1,372$
- $t_{0,95} [10] = 1,812$
- $t_{0,975} [10] = 2,228$
- $t_{0,95} [14] = 1,761$
- $t_{0,975} [21] = 2,080$

Kapitola 6: Základy teorie odhadu (intervaly spolehlivosti)

Populační průměr, podíl či rozptyl lze stanovit na základě předem zvoleného výběrového souboru. Platí, že vlastnosti a charakteristiky jak výběrového, tak i populačního průměru by měly být přibližně identické. Teorie odhadu slouží jako další fáze v rámci analýzy a aplikací statistických testů – především jednovýběrových a dvouvýběrových parametrických a neparametrických testů. Tato kapitola shrnuje veškeré statistické ukazatele popisné statistiky a slouží jako podklad pro budoucí detailnější analýzy.

Příklad 1

K dispozici máte výběrový soubor 10 automobilů v rámci jejich spotřeby pohonných hmot v litrech na 100 kilometrů: 5,5; 5,6; 6,3; 7,2; 5,7; 4,9; 5,1; 6,9; 5,8; 7,1. Uvažujte hladinu významnosti 5 %. Směrodatnou odchylku populačního souboru neznáme. Stanovte odhad populačního průměru.

Výsledek

- průměr výběrového souboru = 6,01
- směrodatná odchylka výběrového souboru = 0,82388
- $5,420672699 < \mu < 6,599327301$

Příklad 2

K dispozici máte výběrový soubor obsahující 35 studentů. Statistický znak vyjadřuje počet bodů, který studenti obdrželi z testu. Průměr výběrového souboru je 40. Rozptyl populačního souboru je 324. Uvažujte hladinu významnosti 1 %. Stanovte pravostranný odhad populačního průměru.

Výsledek

- $\mu < 47,07698367$

Příklad 3

K dispozici máte výběrový soubor obsahující 20 studentů. Statistický znak vyjadřuje počet bodů, který studenti obdrželi z testu. Body jsou uvedeny v tabulce. Rozptyl populačního souboru neznáme. Uvažujte hladinu významnosti 10 %. Stanovte odhad populačního průměru podle následující tabulky.

48	63	22	39	70	21	53	12	34	15
22	70	30	33	19	32	55	12	34	41

Dále:

- určete, jakou hodnotu nepřekročí průměr populačního souboru s pravděpodobností 99 %,
- stanovte levostranný odhad populačního průměru s hladinou významnosti 1 %.

Výsledek

- průměr výběrového souboru = 36,25
- směrodatná odchylka výběrového souboru = 18,40445
- $29,13454234 < \mu < 43,36545766$
- $24,47595464\mu < 48,02404533$
- $25,80110063 < \mu$

Přílohy

Číselné charakteristiky polohy (úrovně) a variability

Aritmetický průměr

$$\bar{x}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

V aplikaci Excel funkce: PRŮMĚR.

Modus

Hodnota, která se ve statistickém souboru vyskytuje nejčastěji. (Hodnota s největší absolutní četností.)

Kvantily

Medián

zn. $X_{0,5}$ je hodnota, jež dělí řadu podle velikosti seřazených výsledků na dvě stejně početné poloviny. V aplikaci excel funkce MEDIAN(množina dat).

Kvartily

První kvartil zn. $X_{0,25}$ je hodnota, jež dělí řadu podle velikosti seřazených výsledků na dvě části tak, že čtvrtina dat se nachází před a tři čtvrtiny za touto hodnotou. V aplikaci excel funkce QUARTIL(množina dat; 1) 3.2.2.

Druhý kvartil = medián

Třetí kvartil zn. $X_{0,75}$ je hodnota, jež dělí řadu podle velikosti seřazených výsledků na dvě části tak, že tři čtvrtiny dat se nachází před a jedna čtvrtina za touto hodnotou. V aplikaci excel funkce QUARTIL(množina dat; 3).

Percentily

Percentil dělí statistický soubor na setiny. V aplikaci excel příkaz PERCENTIL(množina dat; číslo percentilu).

Variační rozpětí

Je to rozdíl mezi největší a nejmenší hodnotou kvantitativního znaku neboli $R = x_{\max} - x_{\min}$.

Kvartilové rozpětí

Je rozdíl třetího a prvního kvartilu. $R_q = x_{0,75} - x_{0,25}$.

Rozptyl

(Též střední kvadratická odchylka, střední kvadratická fluktuace, variance nebo také disperze). Rozptyl je definován jako střední hodnota kvadrátů odchylek od střední hodnoty.

Zn $D(X)$, $\text{var}(x)$, σ^2 . V aplikaci excel funkce VAR.VÝBĚR(množina dat).

Směrodatná odchylka

Zn. σ , je odchylka od střední hodnoty, která má rozměr stejný jako daná veličina.

V aplikaci excel funkce SMODCH.VÝBĚR(množina dat).

Variační koeficient

Je definován jako podíl směrodatné odchylky a aritmetického průměru. Je-li absolutní hodnota variačního koeficientu menší než 0,5, hovoříme o homogenních datech, pokud je roven 0,5 nebo větší, hovoříme o nehomogenních datech.

Vybrané statistické funkce v Excelu

Základní statistické charakteristiky	
průměr	průměr
výběrová směrodatná odchylka	smodch.výběr
výběrový rozptyl	var.výběr
medián	median
kvartil	quartil
maximum	max
minimum	min
populační rozptyl	var
populační směrodatná odchylka	smodch
modus	mode
max	large
min	small
četnosti	četnosti
Základní pravděpodobnostní modely	
binomické rozdělení	binom.dist
Poissonovo rozdělení	poisson.dist
normální rozdělení	norm.dist
hypergeometrické rozdělení	hypgeom.dist
Kvantilové funkce	
kvantil standardního normálního rozdělení	norm.inv
určení kvantilu studentova rozdělení	t.inv
určení kvantilu chí kvadrát rozdělení	chisq.inv

Přehled intervalů spolehlivosti

1	$\bar{X} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ $\mu \leq \bar{X} + u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$ $\mu \geq \bar{X} - u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	<p>u-interval spolehlivosti pro populační průměr</p> <p>podmínky: normální rozdělení, σ známe</p> <p>excel: = norm.inv() R: <i>distributions > continuous > normal > normal quantiles</i></p>
2	$\bar{X} - t_{1-\frac{\alpha}{2}}(n-1) \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{1-\frac{\alpha}{2}}(n-1) \cdot \frac{s}{\sqrt{n}}$ $\mu \leq \bar{X} + t_{1-\alpha}(n-1) \cdot \frac{s}{\sqrt{n}}$ $\mu \geq \bar{X} - t_{1-\alpha}(n-1) \cdot \frac{s}{\sqrt{n}}$	<p>t-interval spolehlivosti pro populační průměr</p> <p>podmínky: normální rozdělení, σ neznáme</p> <p>excel: = t.inv() R: <i>distributions > continuous > t > t quantiles</i></p>
3	$\bar{X} - u_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + u_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ $\mu \leq \bar{X} + u_{1-\alpha} \cdot \frac{s}{\sqrt{n}}$ $\mu \geq \bar{X} - u_{1-\alpha} \cdot \frac{s}{\sqrt{n}}$	<p>asymptotický u-interval spolehlivosti pro populační průměr</p> <p>podmínky: nezáleží na typu rozdělení, $n \geq 30$</p> <p>excel: = norm.inv() R: <i>distributions > continuous > normal > normal quantiles</i></p>
4	$P - u_{1-\frac{\alpha}{2}} \cdot \frac{\sqrt{P \cdot (1-P)}}{\sqrt{n}} \leq \pi \leq P + u_{1-\frac{\alpha}{2}} \cdot \frac{\sqrt{P \cdot (1-P)}}{\sqrt{n}}$ $\pi \leq P + u_{1-\alpha} \cdot \frac{\sqrt{P \cdot (1-P)}}{\sqrt{n}}$ $\pi \geq P - u_{1-\alpha} \cdot \frac{\sqrt{P \cdot (1-P)}}{\sqrt{n}}$	<p>asymptotický u-interval spolehlivosti pro populační poměr (podíl v populaci)</p> <p>podmínky: $n \cdot P \cdot (1-P) > 9$</p> <p>excel: = norm.inv() R: <i>distributions > continuous > normal > normal quantiles</i></p>
5	$\frac{(n-1) \cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \leq \sigma^2 \leq \frac{(n-1) \cdot s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$ $\sigma^2 \leq \frac{(n-1) \cdot s^2}{\chi^2_{\alpha}(n-1)}$ $\sigma^2 \geq \frac{(n-1) \cdot s^2}{\chi^2_{1-\alpha}(n-1)}$	<p>interval spolehlivosti pro populační rozptyl</p> <p>podmínky: normální rozdělení</p> <p>excel: = chisq.inv() R: <i>distributions > continuous > chi squared > chi squared quantiles</i></p>

Vybraná pravděpodobnostní rozdělení

Statistické tabulky

Distribuční funkce normálního rozdělení

$$\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

<i>U</i>	$\Phi(u)$	<i>u</i>	$\Phi(u)$	<i>u</i>	$\Phi(u)$	<i>u</i>	$\Phi(u)$
0,00	0,50000	0,40	0,65542	0,80	0,78814	1,20	0,88493
0,01	,50399	0,41	,65910	0,81	,79103	1,21	,88686
0,02	,50798	0,42	,66276	0,82	,79389	1,22	,88877
0,03	,51197	0,43	,66640	0,83	,79673	1,23	,89065
0,04	,51595	0,44	,67003	0,84	,79955	1,24	,89251
0,05	,51994	0,45	,67364	0,85	,80234	1,25	,89435
0,06	,52392	0,46	,67724	0,86	,80511	1,26	,89617
0,07	,52790	0,47	,68082	0,87	,80785	1,27	,89796
0,08	,53188	0,48	,68439	0,88	,81057	1,28	,89973
0,09	,53586	0,49	,68793	0,89	,81327	1,29	,90147
0,10	,53983	0,50	,69146	0,90	,81594	1,30	,90320
0,11	,54380	0,51	,69497	0,91	,81859	1,31	,90490
0,12	,54776	0,52	,69847	0,92	,82121	1,32	,90658
0,13	,55172	0,53	,70194	0,93	,82381	1,33	,90824
0,14	,55567	0,54	,70540	0,94	,82639	1,34	,90988
0,15	,55962	0,55	,70884	0,95	,82894	1,35	,91149
0,16	,56356	0,56	,71226	0,96	,83147	1,36	,91309
0,17	,56749	0,57	,71566	0,97	,83398	1,37	,91466
0,18	,57142	0,58	,71904	0,98	,83646	1,38	,91621
0,19	,57535	0,59	,72240	0,99	,83891	1,39	,91774
0,20	,57926	0,60	,72575	1,00	,84134	1,40	,91924
0,21	,58318	0,61	,72907	1,01	,84375	1,41	,92073
0,22	,58706	0,62	,73237	1,02	,84614	1,42	,92220
0,23	,59095	0,63	,73565	1,03	,84850	1,43	,92364
0,24	,59483	0,64	,73891	1,04	,85083	1,44	,92507
0,25	,59871	0,65	,74215	1,05	,85314	1,45	,92647
0,26	,60257	0,66	,74537	1,06	,85543	1,46	,92786
0,27	,60642	0,67	,74857	1,07	,85769	1,47	,92922
0,28	,61026	0,68	,75175	1,08	,85993	1,48	,93056
0,29	,61409	0,69	,75490	1,09	,86214	1,49	,93189
0,30	,61791	0,70	,75804	1,10	,86433	1,50	,93319
0,31	,62172	0,71	,76115	1,11	,86650	1,51	,93448
0,32	,62552	0,72	,76424	1,12	,86864	1,52	,93574
0,33	,62930	0,73	,76730	1,13	,87076	1,53	,93699
0,34	,63307	0,74	,77035	1,14	,87286	1,54	,93822
0,35	,63683	0,75	,77337	1,15	,87493	1,55	,93943
0,36	,64058	0,76	,77637	1,16	,87698	1,56	,94062
0,37	,64431	0,77	,77935	1,17	,87900	1,57	,94179
0,38	,64803	0,78	,78230	1,18	,88100	1,58	,94295
0,39	,65173	0,79	,78524	1,19	,88298	1,59	,94408

Distribuční funkce normálního rozdělení

$$\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
1,60	0,94520	2,00	0,97725	2,40	0,99180	3,10	0,99903
1,61	94630	2,01	97778	2,41	99202	3,12	99910
1,62	94738	2,02	97831	2,42	99224	3,14	99916
1,63	94845	2,03	97882	2,43	99245	3,16	99921
1,64	94950	2,04	97932	2,44	99266	3,18	99926
1,65	95053	2,05	97982	2,45	99286	3,20	99931
1,66	95154	2,06	98030	2,46	99305	3,22	99936
1,67	95254	2,07	98077	2,47	99324	3,24	99940
1,68	95352	2,08	98124	2,48	99343	3,26	99944
1,69	95449	2,09	98169	2,49	99361	3,28	99948
1,70	95543	2,10	98214	2,50	99379	3,30	99952
1,71	95637	2,11	98257	2,52	99413	3,32	99955
1,72	95728	2,12	98300	2,54	99446	3,34	99958
1,73	95818	2,13	98341	2,56	99477	3,36	99961
1,74	95907	2,14	98382	2,58	99506	3,38	99964
1,75	95994	2,15	98422	2,60	99534	3,40	99966
1,76	96080	2,16	98461	2,62	99560	3,42	99969
1,77	96164	2,17	98500	2,64	99585	3,44	99971
1,78	96246	2,18	98537	2,66	99609	3,46	99973
1,79	96327	2,19	98574	2,68	99632	3,48	99975
1,80	96407	2,20	98610	2,70	99653	3,50	99977
1,81	96485	2,21	98645	2,72	99674	3,55	99981
1,82	96562	2,22	98679	2,74	99693	3,60	99984
1,83	96638	2,23	98713	2,76	99711	3,65	99987
1,84	96712	2,24	98745	2,78	99728	3,70	99989
1,85	96784	2,25	98778	2,80	99744	3,75	99991
1,86	96856	2,26	98809	2,82	99760	3,80	99993
1,87	96926	2,27	98840	2,84	99774	3,85	99994
1,88	96995	2,28	98870	2,86	99788	3,90	99995
1,89	97062	2,29	98899	2,88	99801	3,95	99996
1,90	97128	2,30	98928	2,90	99813	4,00	99997
1,91	97193	2,31	98956	2,92	99825	4,05	99997
1,92	97257	2,32	98983	2,94	99836	4,10	99998
1,93	97320	2,33	99010	2,96	99846	4,15	99998
1,94	97381	2,34	99036	2,98	99856	4,20	99999
1,95	97441	2,35	99061	3,00	99865	4,25	99999
1,96	97500	2,36	99086	3,02	99874	4,30	99999
1,97	97558	2,37	99111	3,04	99882	4,35	99999
1,98	97615	2,38	99134	3,06	99890	4,40	99999
1,99	97670	2,39	99158	3,08	99897	4,45	1,00000

Kvantily normovaného normálního rozdělení

P	u_P	P	u_P	P	u_P	P	u_P
0,50	0,000	0,75	0,674	0,950	1,645	0,975	1,960
0,51	0,025	0,76	0,706	0,951	1,655	0,976	1,970
0,52	0,050	0,77	0,739	0,952	1,665	0,977	1,995
0,53	0,075	0,78	0,772	0,953	1,675	0,978	2,014
0,54	0,100	0,79	0,806	0,954	1,685	0,979	2,034
0,55	0,126	0,80	0,842	0,955	1,695	0,980	2,054
0,56	0,151	0,81	0,878	0,956	1,706	0,981	2,075
0,57	0,176	0,82	0,915	0,957	1,717	0,982	2,097
0,58	0,202	0,83	0,954	0,958	1,728	0,983	2,120
0,59	0,228	0,84	0,994	0,959	1,739	0,984	2,144
0,60	0,253	0,85	1,036	0,960	1,751	0,985	2,170
0,61	0,279	0,86	1,080	0,961	1,762	0,986	2,197
0,62	0,305	0,87	1,126	0,962	1,774	0,987	2,226
0,63	0,332	0,88	1,175	0,963	1,787	0,988	2,257
0,64	0,358	0,89	1,227	0,964	1,799	0,989	2,290
0,65	0,385	0,90	1,282	0,965	1,812	0,990	2,326
0,66	0,412	0,905	1,311	0,966	1,825	0,991	2,366
0,67	0,440	0,910	1,341	0,967	1,838	0,992	2,409
0,68	0,468	0,915	1,372	0,968	1,852	0,993	2,457
0,69	0,496	0,920	1,405	0,969	1,866	0,994	2,512
0,70	0,524	0,925	1,440	0,970	1,881	0,995	2,576
0,71	0,553	0,930	1,476	0,971	1,896	0,996	2,652
0,72	0,583	0,935	1,514	0,972	1,911	0,997	2,748
0,73	0,613	0,940	1,555	0,973	1,927	0,998	2,878
0,74	0,643	0,945	1,598	0,974	1,943	0,999	3,090

Kvantily Studentova rozdělení

ν	P				
	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,553	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750

Základní vzorce popisné statistiky a pravděpodobnosti

Sturgesovo pravidlo

$$1 + 3,3 * \log (n)$$

populační aritmetický průměr

$$mí = \frac{1}{N} \sum_{i=1}^N x_i$$

výběrový aritmetický průměr

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

vážený průměr pro absolutní četnosti

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

vážený průměr pro relativní četnosti

$$\bar{x} = \sum_{i=1}^k x_i p_i$$

obecně

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i}$$

harmonický průměr dle povahy zadaných hodnot

$$\bar{x}_H = \frac{\sum_{i=1}^k n_i}{\sum_{i=1}^k \frac{n_i}{x_i}}$$

$$\bar{x}_H = \frac{1}{\sum_{i=1}^k \frac{p_i}{x_i}}$$

$$\bar{x}_H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

geometrický průměr

$$\bar{x}_G = \sqrt[n]{\prod_{i=1}^n x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

medián

lichý počet dat

$$x_{\left(\frac{n+1}{2}\right)}$$

sudý počet dat

$$\frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)}}{2}$$

rozptyl pro výběrový soubor

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

rozptyl pro populační soubor

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

směrodatná odchylka pro výběrový soubor

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

směrodatná odchylka pro populační soubor

$$s = \sqrt{\frac{1}{N} \sum_{i=1}^N (\bar{x} - x_i)^2}$$

variační koeficient

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

klasická pravděpodobnost

$$P(X) = m/n$$

statistická pravděpodobnost

$$P(X) \approx m/n$$

podmíněná pravděpodobnost

$$P(A \cap B) = P(A) * P(B|A) = P(B) * P(A|B)$$

Úplná pravděpodobnost

$$P(A) = P(B) \times P(A|B) + P(\bar{B}) \times P(A|\bar{B})$$

$$P(A) = \sum_{i=1}^n P(B_i) \times P(A | B_i)$$

Bayesův vzorec

$$P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B) \times P(A | B)}{P(A)}$$

$$P(B_i | A) = \frac{P(B_i) \times P(A | B_i)}{\sum_{j=1}^n P(B_j) \times P(A | B_j)}$$

Binomické rozdělení

$$P(X = k) = \binom{n}{k} * p^k * (1 - p)^{n - k}$$

Poissonovo rozdělení

$$P(X = x) = \frac{(t * \lambda)^x}{x!} * e^{-t * \lambda}$$

Hypergeometrické rozdělení

$$P(x) = \frac{\binom{M}{x} * \binom{N-M}{n-x}}{\binom{N}{n}}$$

Rovnoměrné rozdělení

$$F(X) = \frac{x-a}{b-a}$$