

LIMITNÍ VĚTY

DALŠÍ SPOJITÁ ROZDĚLENÍ PR.



8. cvičení

Způsoby statistického šetření

Vyčerpávající šetření

prošetření všech jednotek statistického souboru (populace)

Výhody: přesnost a detailnost zjištěných informací

Nevýhody: personální, finanční a časová náročnost

Výběrové šetření

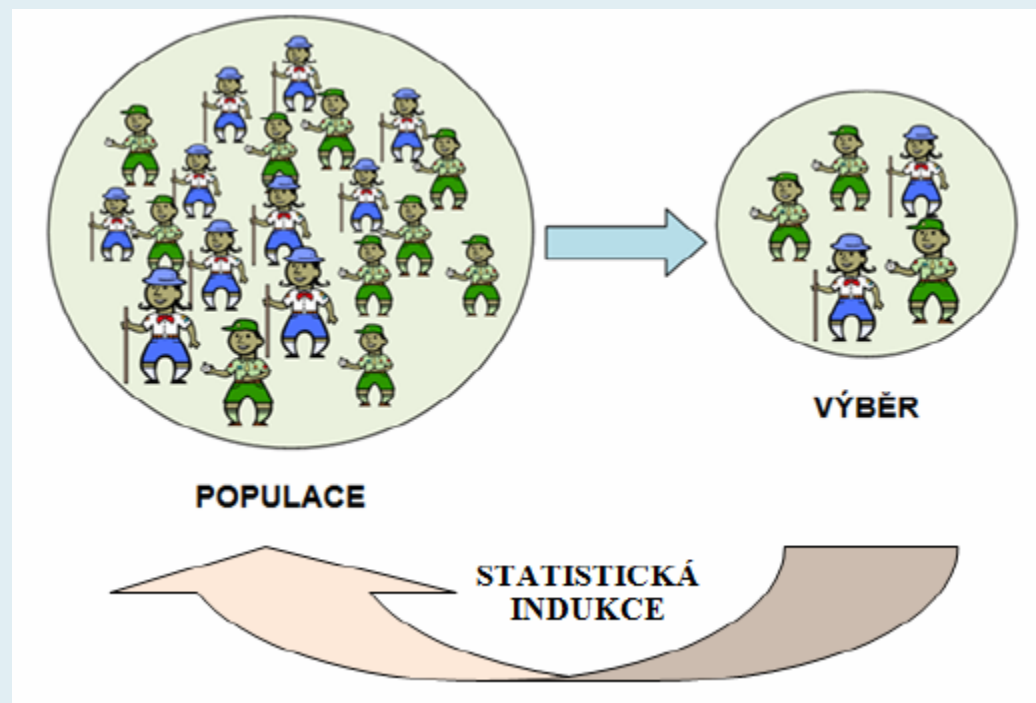
ze základního souboru (populace) o rozsahu N vybereme jeho část, tzv. **výběrový soubor**, zkráceně **výběr**, o rozsahu n . Ze zpracovaných výsledků pak usuzujeme na vlastnosti celé populace

Výhody: menší personální, finanční a časová náročnost

Nevýhody: mírou objektivnosti získaných informací je kvalita provedení výběrového šetření

Náhodné výběry

- ❖ Používáme k modelování a zkoumání **populace**.
- ❖ Výběr musí být **reprezentativní** (všechny jednotky, z nichž se skládá populace, musí mít stejnou šanci na zařazení do zkoumaného výběru)
- ❖ **Statistická indukce** na základě zkoumání omezeného počtu statistických jednotek (výběru) usuzuje na vlastnosti celé populace.



Parametry populace vs. výběrové charakteristiky

- ❖ **Parametry základního souboru** (populace) – střední hodnota μ , rozptyl σ^2 , směrodatná odchylka σ , pravděpodobnost $\pi, \dots$.

Parametry populace jsou **konstantní hodnoty** (pro určitou náhodnou veličinu, v pevném čase). Jestliže neznáme rozdělení náhodné veličiny, nedokážeme parametry většinou přesně určit.

- ❖ **Parametry výběrového souboru** jsou příslušné protějšky parametrů populace, ozn. **výběrové charakteristiky**.

Výběrů ze základního souboru může být mnoho
=> výběrové charakteristiky jsou proměnlivé (variabilní), proto mají charakter náhodných veličin, lze je popsat nějakým rozdělením, mají také svou střední hodnotu, rozptyl a všechny ostatní charakteristiky.

Základní parametry populace a příslušné výběrové charakteristiky

Základní soubor (populace)	střední hodnota $\mu (E(X))$	medián $x_{0,5}$	rozptyl σ^2	směrodatná odchylka σ	pravděpodobnost π
Výběrový soubor (výběr)	(výběrový) průměr $\bar{X}$	výběrový medián $\tilde{X}_{0,5}$	výběrový rozptyl S^2	výběrová směrodatná odchylka S	relativní četnost p

výběrové charakteristiky jsou náhodné veličiny, jejichž jednotlivé realizace lze získat výpočtem pozorovaných hodnot těchto charakteristik pro jednotlivé výběry o rozsahu n . (Např. Průměrný plat 20 občanů ČR je náhodná veličina. Výpočtem průměrného platu konkrétních 20 občanů získáme jednu realizaci tohoto průměru, výpočtem průměrného platu jiného vzorku 20 občanů ČR získáme jinou realizaci průměru.)

výběrová rozdělení = rozdělení pravděpodobností výběrových charakteristik.

Limitní věty

- ❖ Tvzení, která jsou důležitá pro popis pravděpodobnostních modelů v případě rostoucího počtu náhodných pokusů.
 - ❖ **Zákon velkých čísel** – Zabývá se konvergencí výběrového průměru ke střední hodnotě v posloupnosti nezávislých veličin.
 - ❖ **Centrální limitní věta** - zabývá se konvergencí rozdělení k normálnímu rozdělení, dvě dílčí formulace:
 - ❖ Lindebergova-Lévyho věta
 - ❖ Moivreova-Laplaceova věta

Zákon velkých čísel

- ❖ Říká, že průměr nezávislých náhodných veličin se stejnými středními hodnotami konverguje podle pravděpodobnosti k jejich střední hodnotě.
- ❖ Důsledkem je **Bernoulliho věta: Relativní četnost** sledovaného jevu **konverguje** podle pravděpodobnosti k jeho **pravděpodobnosti**.
- ❖ Umožňuje experimentálně odhadovat neznámou pravděpodobnost pomocí pozorované relativní četnosti (viz. pravděpodobnost toho, že padne šestka na „cinknuté“ kostce, cv.3).

Lindebergova-Lévyho věta

- ❖ **Aproximace rozdělení součtů náhodných veličin**
- ❖ Jestliže $X_1, X_2, \dots, X_n$ jsou nezávislé náhodné veličiny se stejným (libovolným) rozdělením, stejnými středními hodnotami a se stejnými (konečnými) rozptyly, pak jejich **součet konverguje v distribuci k normálnímu rozdělení** o parametrech: $n\mu; n\sigma^2$.

$\{X_n\}$:

- nezávislé NV
- NV se stejným typem rozdělení
- $EX_1 = EX_2 = \dots = EX_n$
- $DX_1 = DX_2 = \dots = DX_n < \infty$,

Pak:

$$X = \sum_{i=1}^n X_i \rightarrow N(n\mu; n\sigma^2)$$

Důsledek Linderbergovy-Lévyho věty

- ❖ **Aproximace rozdělení průměru náhodných veličin**
- ❖ Jestliže $X_1, X_2, \dots, X_n$ jsou nezávislé náhodné veličiny se stejným (libovolným) rozdělením, stejnými středními hodnotami a se stejnými (konečnými) rozptyly, pak jejich **průměr konverguje v distribuci k normálnímu rozdělení** o parametrech: $\mu; \sigma^2/n$.

$\{X_n\}$:

- nezávislé NV
- NV se stejným typem rozdělení
- $EX_1 = EX_2 = \dots = EX_n$
- $DX_1 = DX_2 = \dots = DX_n < \infty$,

Pak:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \rightarrow N\left(\mu; \frac{\sigma^2}{n}\right)$$

1. Zatížení letadla s 64 místy nemá překročit 6 000 kg. Jaká je pravděpodobnost, že při plném obsazení bude tato hodnota překročena, jestliže má hmotnost cestujícího střední hodnotu 90 kg a směrodatnou odchylku 10 kg?

Řešení:

X_i ... hmotnost i -tého cestujícího

Rozdělení NV neznáme, ale víme:

$$EX_i = 90 \text{ kg}, \quad DX_i = 10^2 \text{ kg}^2,$$

X ... celková hmotnost všech cestujících podle Lindebergovy-Lévyho věty:

$$X = \sum_{i=1}^n X_i \rightarrow N(n \cdot 90; n \cdot 10^2)$$

$$X = \sum_{i=1}^{64} X_i \rightarrow N(64 \cdot 90 = 5760; 64 \cdot 10^2 = 6400)$$

$$P(X > 6000) = \underline{\underline{0,0014}}$$

Pravděpodobnost, že při plném obsazení bude překročena hodnota 6000 kg je 0,14 %.

2. Denní procento přijatých kalorií z tuku můžeme u Američanů modelovat normálním rozdělením se střední hodnotou 36 % a směrodatnou odchylkou 10 %. Předpokládejte, že jste náhodně vybrali 45 Američanů, určete:
- a) rozdělení průměrného procenta přijatých kalorií z tuků,
 - b) pravděpodobnost, že 45 Američanů bude přijímat denně průměrně více než 40 % kalorií z tuku,

Řešení:

X_i ... denní procento přijatých kalorií z tuku u i -tého Američana

$$X_i \rightarrow N(36; 100)$$

a)
 $\bar{X}$... průměrné denní procento kalorií přijatých z tuku u 45 Američanů podle Lindebergovy-Lévyho věty:

$$\bar{X} = \frac{\sum_{i=1}^{45} X_i}{45} \rightarrow N\left(36; \frac{100}{45}\right)$$

$$\underline{\underline{\bar{X} \rightarrow N(36; 2,2)}}$$

Průměrné procento přijatých kalorií z tuků můžeme u 45 Američanů modelovat normálním rozdělením se střední hodnotou 36 % a směrodatnou odchylkou 1,5 %.

Řešení:

b)

$\bar{X}$... průměrné denní procento kalorií přijatých z tuku u 45 Američanů

$$\bar{X} \rightarrow N(36;2,2)$$

$$P(\bar{X} > 40) = \underline{\underline{0,004}}$$

Pravděpodobnost, že průměrné procento denně přijatých kalorií z tuku je u 45 Američanů větší než 40 % je 0,4%.

3. Předpokládejme, že délku odevzdávaných seminárních projektů ze statistiky můžeme modelovat rovnoměrným rozdělením od 13 do 25 stran. Mějme cvičícího se 47 studenty, určete pravděpodobnost, že cvičící musí za semestr přečíst více než 1000 stran semestrálních projektů.

Řešení:

X_i ... délka i -tého odevzdaného semestrálního projektu

$$X_i \rightarrow R(13;25)$$

Střední hodnotu a rozptyl dopočítáme podle vzorečků pro rovnoměrné rozdělení:

$$E(X_i) = \frac{13 + 25}{2} = 19 \qquad D(X_i) = \frac{(13 - 25)^2}{12} = 12$$

X ... celková délka všech odevzdaných projektů podle Lindebergovy-Lévyho věty:

$$X = \sum_{i=1}^{47} X_i \rightarrow N(47 \cdot 19 = \underline{\underline{893}}; 47 \cdot 12^2 = \underline{\underline{6768}})$$

$$P(X > 1000) = \underline{\underline{0,097}}$$

Pravděpodobnost, že cvičící musí za semestr přečíst více než 1000 stran semestrálních projektů je 9,7 %.

4. Předpokládejme, že světová třída vytrvalostních běžců uběhne maratón průměrně za 135 minut se směrodatnou odchylkou 14 minut.

Nalezněte:

- a) rozdělení průměrného času na maratónu u 49 závodníků,
- b) pravděpodobnost, že průměrná doba na dokončení maratónu je u těchto 49 závodníků mezi 132 a 136 minutami.

Řešení:

a)

X_i ... doba běhu i -tého závodníka

Rozdělení NV neznáme, ale víme:

$$EX_i = 135 \text{ min}, DX_i = 14 \text{ min}^2,$$

$\bar{X}$... průměrná doba běhu 49 závodníků podle Lindebergovy-Lévyho věty:

$$\bar{X} = \frac{\sum_{i=1}^{49} X_i}{49} \rightarrow N\left(135; \frac{196}{49} = \left(\frac{14}{7}\right)^2 = 4\right)$$

$$\bar{X} \rightarrow N(135;4)$$

Průměrný čas běhu maratónu u 49 závodníků můžeme modelovat pomocí normálního rozdělení se střední hodnotou 135 minut a směrodatnou odchylkou 2 minuty.

Řešení:

b)

$\bar{X}$... průměrná doba běhu 49 závodníků

$$\bar{X} \rightarrow N(135;4)$$

$$P(132 < \bar{X} < 136) = \underline{\underline{0,625}}$$

Pravděpodobnost, že průměrná doba na dokončení maratónu je u 49 závodníků mezi 132 a 136 minutami je 62,5 %.

Moivreova-Laplaceova věta

- ❖ Vyjadřuje konvergenci binomického rozdělení k normálnímu rozdělení

$$X \rightarrow Bi(n;p); \quad EX = np; DX = np(1-p) \quad ,$$

pak pro dostatečně velká n : $X \rightarrow N(np; np(1-p))$

Aproximace dává dobré výsledky, když:

$$np(1-p) > 9 \quad \text{nebo} \quad \min\{np; np(1-p)\} > 5$$

Další aplikace CLV

- ❖ Aproximace výběrové relativní četnosti normálním rozdělením:

$$p \rightarrow N\left(p; \frac{n \cdot (1-p)}{n}\right)$$

- ❖ Aproximace Poissonova rozdělení normálním rozdělením:

$$X \rightarrow Po(\lambda t), \quad EX = \lambda t, \quad DX = \lambda t$$

pak pro dostatečně velké t platí, že X můžeme aproximovat norm. rozdělením

$$X \rightarrow N(\lambda t, \lambda t)$$

- ❖ Aproximace průměrného počtu výskytu události za časovou jednotku normálním rozdělením:

Nechť Y je průměrný počet výskytu události za časovou jednotku, pak:

$$Y \rightarrow N\left(\lambda, \frac{\lambda}{t}\right)$$

Oprava na spojitost

Oprava na spojitost (u pravděpodobnostní funkce):

Jestliže aproximujeme diskrétní náhodnou veličinu spojitou, pak:

$$P(X = a) = P(a - 0,5 \leq X < a + 0,5)$$

Obecně – oprava na spojitost:

Posouzení vhodnosti použití opravy na spojitost provádíme vždy při řešení konkrétního příkladu důsledným převodem pravděpodobnosti výskytu náhodné veličiny na nějakém intervalu na vztah mezi distribučními funkcemi v příslušných bodech.

5. Na telefonní ústřednu je napojeno 3000 účastníků. Každý z nich bude volat telefonní ústřednu během hodiny s pravděpodobností 10%. Jaká je pravděpodobnost, že během následující hodiny zavolá ústřednu:
- a) právě 300 účastníků?
 - b) více než 310 účastníků?
 - c) mezi 200 a 450 účastníky(včetně)?

Řešení:

X ... počet účastníků volajících ústřednu (z 3000)

$$X \rightarrow Bi(3000;0,1)$$

a)

$$P(X = 300) = \frac{3000!}{2700!300!} \cdot (0,1)^{300} \cdot (0,9)^{2700}$$

Na kalkulačce nelze vypočítat => aproximujeme pomocí Moivreovy-Laplaceovy věty:

$$EX=n \cdot p$$

$$DX=n \cdot p \cdot (1-p)$$

$$X \rightarrow Bi(3000;0,1) \approx X \rightarrow N(3000 \cdot 0,1; 3000 \cdot 0,1 \cdot 0,9)$$

$$X \rightarrow N(300;270)$$

$P(X = 300) = 0 \Rightarrow$ oprava na spojitost:

$$P(299,5 \leq X < 300,5) = P(X < 300,5) - P(X < 299,5) = \underline{\underline{0,024}}$$

Pravděpodobnost, že během následující hodiny zavolá ústřednu právě 300 účastníků je 2,4 %.

Řešení:

b) Aproximujeme pomocí Moivreovy-Laplaceovy věty:

$$X \rightarrow Bi(3000;0,1) \approx X \rightarrow N(300;270)$$

$$P(X > 310) = 1 - P(X \leq 310) \Rightarrow \text{oprava na spojitost:}$$

$$P(X > 310) = 1 - P(X < 310,5) = \underline{\underline{0,261}}$$

Pravděpodobnost, že během následující hodiny zavolá ústřednu více než 310 účastníků je 26,1 %.

Řešení:

c) Aproximujeme pomocí Moivreovy-Laplaceovy věty:

$$X \rightarrow Bi(3000;0,1) \approx X \rightarrow N(300;270)$$

$$P(200 \leq X \leq 450) = P(X \leq 450) - P(X < 200) = \underline{\underline{1}}$$

Pravděpodobnost, že během následující hodiny zavolá na ústřednu mezi 200 a 450 účastníky (včetně) je 100 %.

Použitá aproximace dává velmi dobré výsledky (velmi blízké skutečným), protože je splněna podmínka, že: $np(1-p) > 9$, $(270 > 9)$.

6. Sekretářka Petra píše na stroji rychlosti 250 úhozů / min. Při této rychlosti udělá průměrně 3 chyby za 10 minut. Jaká je pravděpodobnost, že při 30 minutovém diktátu udělá více než 10 chyb?

Řešení:

X ... počet chyb v diktátu (30 minut)

$$X \rightarrow Po(9) \leftarrow \text{průměrný počet chyb za 30 min}$$

$$P(X > 10) = 1 - P(X \leq 10) = 1 - \sum_{k=0}^{10} \frac{9^k}{k!} \cdot e^{-9} = 1 - e^{-9} \cdot \left(\frac{9^0}{0!} + \frac{9^1}{1!} + \dots + \frac{9^{10}}{10!} \right)$$

$$= \underline{\underline{0,294}}$$

Srovnání s výpočtem pomocí centrální limitní věty:

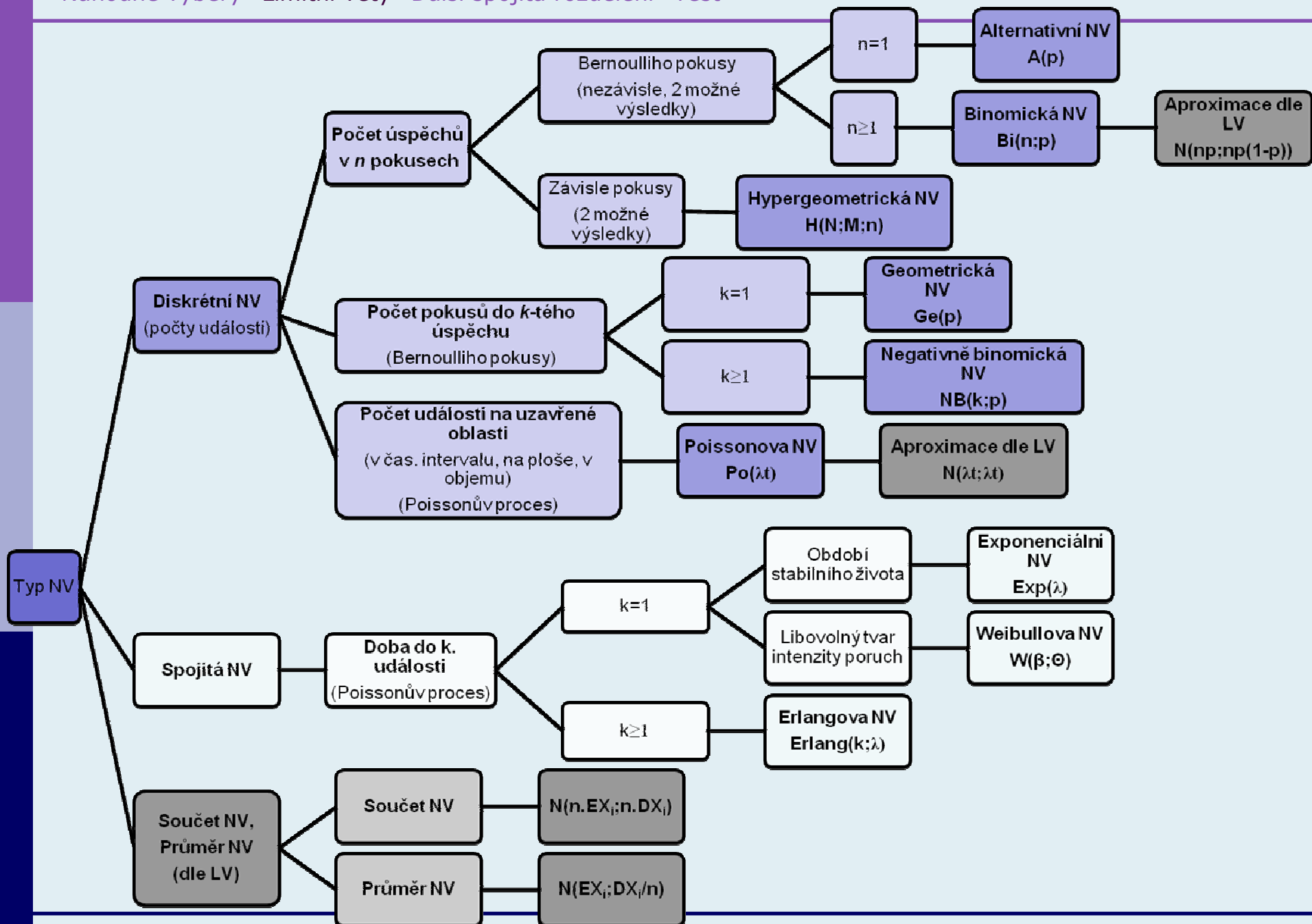
Poissonovu náhodnou veličinu s parametrem λt můžeme aproximovat pro dostatečně velká λt normálním rozdělením s parametry: $\mu = \lambda t$, $\sigma^2 = \lambda t$.

$$X \rightarrow N(9;9)$$

$$P(X > 10) = 1 - P(X \leq 10) = 1 - P(X < 10,5) = 1 - F(10,5) = 1 - \Phi\left(\frac{10,5 - 9}{\sqrt{9}}\right) =$$

$$= 1 - \Phi(0,5) = 1 - 0,691 = \underline{\underline{0,309}}$$

Výsledky obou postupů se liší o 0,015, to je asi 5%ní chyba (0,015/0,294·100%).



χ^2 - rozdělení (Pearsonovo)

Mějme nezávislé náhodné veličiny $Z_1, Z_2, \dots, Z_n$.

$\forall i = 1, \dots, n : Z_i \rightarrow N(0;1)$

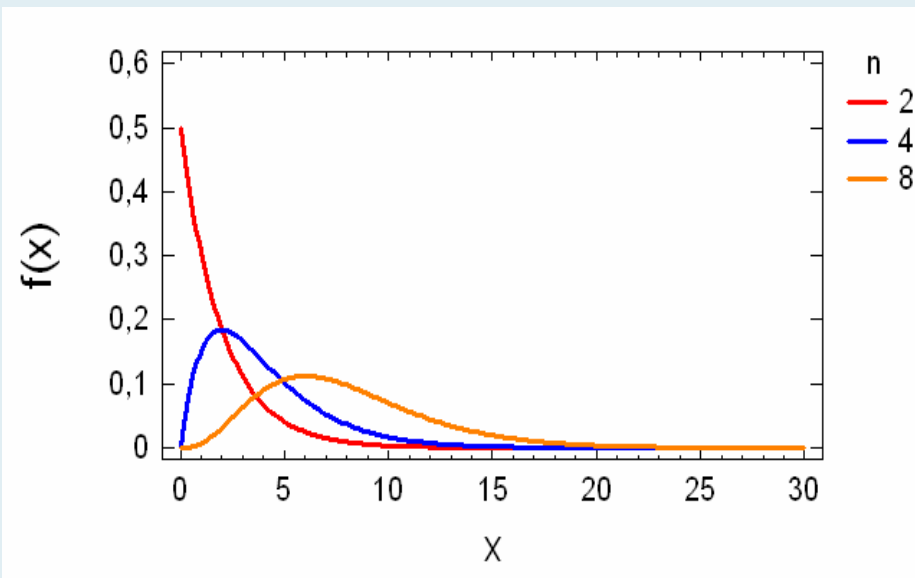
Pak

$$X = \sum_{i=1}^n Z_i^2 \quad X \rightarrow \chi_n^2$$

n ... počet sčítaných nezávislých veličin = počet stupňů volnosti
= jediný parametr rozdělení

$$E(\chi_n^2) = n \quad D(\chi_n^2) = 2n$$

x_p ... tabelovány



Hustota pravděpodobnosti χ^2 - rozdělení

χ^2 - rozdělení - vlastnosti

- ❖ Součet nezávislých náhodných veličin s χ^2 rozdělením má opět χ^2 rozdělení a počet stupňů volnosti je roven součtu stupňů volnosti jednotlivých veličin v součtu.

$$X_i \rightarrow \chi_{n_i}^2, \quad X = \sum_{(i)} \chi_{n_i}^2, \quad \text{pak} \quad X \rightarrow \chi_{\sum_{(i)} n_i}^2$$

- ❖ Jestliže provedeme náhodný pokus spočívající v náhodném výběru o rozsahu n z populace **podléhající normálnímu rozdělení** s rozptylem σ^2 . Pro uvedený výběr určíme výběrovou směrodatnou odchylku s . Lze ukázat, že náhodná veličina
$$\frac{(n-1) s^2}{\sigma^2}$$

má χ^2 rozdělení s $n-1$ stupni volnosti.

$$X_i \rightarrow N(\mu; \sigma^2) \Rightarrow \frac{(n-1) s^2}{\sigma^2} \rightarrow \chi_{n-1}^2$$

χ^2 - rozdělení - použití

- ❖ $\frac{(n-1)s^2}{\sigma^2}$ je testová statistika pro testy o rozptylu (je-li výběr z $N(0;1)$)
- ❖ χ^2 se používá při testování nezávislosti kategoriálních veličin (testy nezávislosti v kontingenční tabulce)
- ❖ Test dobré shody (testování toho, zda data pocházejí z určitého rozdělení).
- ❖ Se vzrůstajícím počtem stupňů volnosti se χ^2 rozdělení blíží normálnímu rozdělení $N(n;2n)$.

7. Odvodte distribuční funkci a hustotu pravděpodobnosti náhodné veličiny X , která má rozdělení χ_n^2 s jedním stupněm volnosti.

Řešení:

$$X = Z^2$$

$$Z \rightarrow N(0;1) \Rightarrow X \rightarrow \chi_1^2$$

Náhodná veličina X je funkcí náhodné veličiny Z (náhodná veličina s rozdělením χ_1^2 nabývá pouze kladných hodnot):
pro $x > 0$:

$$\begin{aligned} \underline{\underline{F(x)}} &= P(X < x) = P(Z^2 < x) = P(-\sqrt{x} < Z < \sqrt{x}) = \\ &= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}) = \Phi(\sqrt{x}) - [1 - \Phi(\sqrt{x})] = 2\Phi(\sqrt{x}) - 1 = \\ &= \frac{2}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\sqrt{x}} e^{-\frac{t^2}{2}} dt - 1 = \underline{\underline{\sqrt{\frac{2}{\pi}} \cdot \int_{-\infty}^{\sqrt{x}} e^{-\frac{t^2}{2}} dt - 1}} \end{aligned}$$

pro $x \leq 0$ $F(x) = 0$

Řešení:

Hustotu pravděpodobnosti určíme jako derivaci distribuční funkce:

pro $x > 0$:

$$\underline{\underline{f(x) = \frac{dF(x)}{dx} = \frac{d(2\Phi(\sqrt{x}) - 1)}{dx} = 2 \cdot \frac{1}{2\sqrt{x}} \cdot \varphi(\sqrt{x}) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}}}$$

pro $x \leq 0$:

$$\underline{\underline{f(x) = \frac{dF(x)}{dx} = 0}}$$

Studentovo rozdělení (t rozdělení)

Z a X jsou nezávislé náhodné veličiny, kde

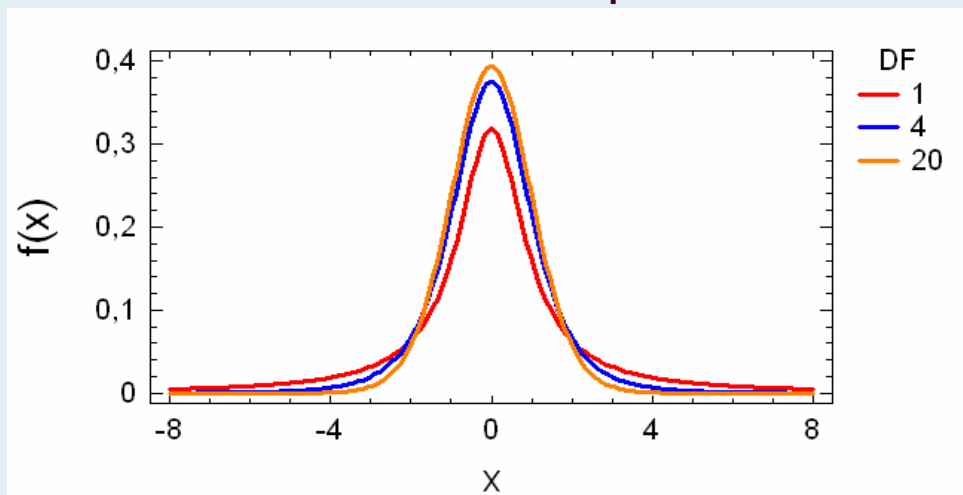
$$Z \rightarrow N(0;1), \quad X \rightarrow \chi_n^2, \quad \text{pak}$$

náhodná veličina T má studentovo rozdělení s n stupni volnosti

$$T_n = \frac{Z}{\sqrt{\frac{X}{n}}} \rightarrow t_n$$

$$E(T_n) = 0 \quad D(T_n) = \frac{n}{n-2}$$

t_p tabelovány



Hustota pravděpodobnosti Studentova t rozdělení

pro $n \rightarrow \infty$ (v praxi $n > 30$) se Studentovo t rozdělení blíží normovanému normálnímu rozdělení.

Studentovo rozdělení - vlastnosti

Pokud náhodné veličiny $X_1, X_2, \dots, X_n$ mají normální rozdělení $N(\mu; \sigma^2)$ a jsou navzájem nezávislé, pak náhodná veličina definována jako

$$\frac{\bar{X} - \mu}{s} \cdot \sqrt{n}$$

má Studentovo t rozdělení s $(n-1)$ stupni volnosti:

$$\frac{\bar{X} - \mu}{s} \cdot \sqrt{n} \rightarrow t_{n-1}$$

Studentovo rozdělení - použití

Studentovo t rozdělení má uplatnění zejména při modelování založeném na analýze malých výběrů.

- ❖ $\frac{\bar{X} - \mu}{s} \cdot \sqrt{n}$ je testovou statistikou využívanou při testování střední hodnoty (neznáme-li σ a jde-li o výběr z $N(0;1)$).
- ❖ Užívá se k testování hypotéz o shodě středních hodnot dvou náhodných výběrů, se stejnými předpoklady jako v předcházejícím případě - navíc musí být tyto výběry nezávislé.
- ❖ Rozdělení je vhodným prostředkem pro analýzu výsledků regresní analýzy.

Fisher-Snedecorovo rozdělení

V a W jsou nezávislé náhodné veličiny, obě s rozdělením chí-kvadrát, s různým počtem stupňů volnosti:

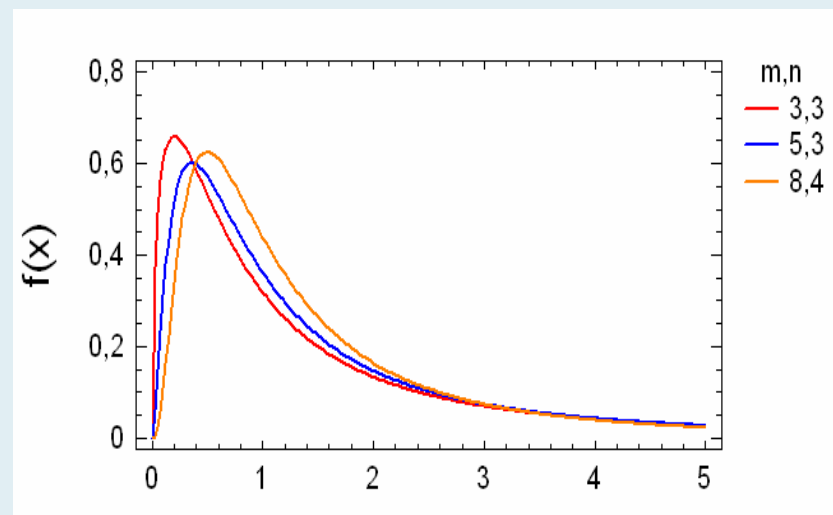
$$V \rightarrow \chi_m^2, \quad W \rightarrow \chi_n^2, \quad \text{pak}$$

$$F = \frac{\frac{V}{m}}{\frac{W}{n}} \rightarrow F_{m,n}$$

$$E(F) = \frac{n}{n-2} \quad \text{pro } n > 2$$

$$D(F) = \frac{2n^2 \left(1 + \frac{n-2}{m}\right)}{(n-2)^2 (n-4)} \quad \text{pro } n > 4$$

$$F_p = \frac{1}{F_{1-p}(m,n)}$$



Hustota pravděpodobnosti Fisher-Snedecorova rozdělení

Fisher-Snedecorovo rozdělení - vlastnosti

Mějme dva výběry z normálního rozdělení.

$\forall i = 1, 2, \dots, n_1$, kde n_1 je rozsah prvního výběru: $X_{1i} \rightarrow N(\mu_1; \sigma_1^2)$

$\forall j = 1, 2, \dots, n_2$, kde n_2 je rozsah druhého výběru: $X_{2j} \rightarrow N(\mu_2; \sigma_2^2)$

Nechť výběrové rozptyly S_1^2 , S_2^2 a jsou náhodné veličiny definované jako

$$S_1^2 = \frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2}{n_1 - 1}, S_2^2 = \frac{\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2}{n_2 - 1}$$

pak

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \rightarrow F_{n_1-1, n_2-1}$$

Fisher-Snedecorovo rozdělení - použití

Široké uplatnění, zejména při hodnocení výsledků statistických analýz. Používá se především

- ❖ testy o shodnosti rozptylů dvou náhodných výběrů
- ❖ testy o shodě středních hodnot pro více náhodných výběrů, v analýze rozptylu
- ❖ testy v regresní analýze

8. Firma Edison vyrábí žárovky Ed. Životnost těchto žárovek je průměrně 5 let se směrodatnou odchylkou 6 měsíců. Žárovky jsou vyráběny na dvou linkách. Předpokládejme, že obě linky mají srovnatelné parametry, tj. že průměrná životnost a variabilita životnosti žárovek Ed vyrobených ve firmě Edison nezávisí na tom, na jaké lince byly vyrobeny. Pro ověření kvality výroby bude testována životnost 20 žárovek z linky 1 a 30 žárovek z linky 2. Jaká je pravděpodobnost, že u vzorku z linky 1 bude zjištěn více než dvojnásobný rozptyl oproti rozptylu zjištěnému u vzorku z linky 2?

Řešení:

S_1^2 ... rozptyl životnosti zjištěný u vzorku z linky 1

S_2^2 ... rozptyl životnosti zjištěný u vzorku z linky 2

Hledáme pravděpodobnost, že $S_1^2 > 2S_2^2$, tzn. $\frac{S_1^2}{S_2^2} > 2$.

$$P\left(\frac{S_1^2}{S_2^2} > 2\right) = ?$$

Za předpokladu, že oba vzorky jsou výběrem z normálního rozdělení, platí:

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \rightarrow F_{n_1-1, n_2-1}$$

Řešení:

Ze zadání předpokládáme, že rozptyl životnosti žárovek vyrobených na jednotlivých linkách je stejný:

$$\sigma_1^2 = \sigma_2^2$$

Pak

$$\frac{S_1^2}{S_2^2} \rightarrow F_{n_1-1, n_2-1}$$

Testujeme: 20 žárovek z linky 1 $\Rightarrow n_1 = 20$ a 30 žárovek z linky 2 $\Rightarrow n_2 = 30$. $\Rightarrow \frac{S_1^2}{S_2^2} \rightarrow F_{19,29}$

$$P\left(\frac{S_1^2}{S_2^2} > 2\right) = 1 - F_{19,29}(2) \cong \underline{\underline{0,045}}$$

Pravděpodobnost, že u vzorku z linky 1 bude zjištěn více než dvojnásobný rozptyl oproti rozptylu zjištěnému u vzorku z linky 2 je přibližně 4,5 %.

Test

1. Střední hodnota pevně zvolené náhodné veličiny je
 - a) náhodná veličina,
 - b) konstanta,
 - c) náhodný jev,
 - d) výběrová charakteristika.

1. Střední hodnota pevně zvolené náhodné veličiny je

~~a)~~ náhodná veličina,

☒ b) konstanta,

~~c)~~ náhodný jev,

~~d)~~ výběrová charakteristika.

2. Výběrový průměr je

- a) náhodná veličina,
- b) konstanta,
- c) náhodný jev,
- d) populační charakteristika.

2. Výběrový průměr je

- ☒ a) náhodná veličina,
- ☐ b) konstanta,
- ☐ c) náhodný jev,
- ☐ d) populační charakteristika.

3. S rostoucím rozsahem výběru se obvykle rozptyl průměru

- a) snižuje,
- b) zvyšuje,
- c) nemění.

3. S rostoucím rozsahem výběru se obvykle rozptyl průměru

☒ a) snižuje,

☐ b) zvyšuje,

☐ c) nemění.

4. Statistická indukce je

- a) experiment,
- b) metoda, která umožňuje odhadnout vlastnosti výběru na základě znalostí vlastností populace,
- c) zobecnění statistických výsledků získaných zpracováním výběru na celou populaci,
- d) metoda sběru dat.

4. Statistická indukce je

- ☒ a) experiment,
- ☒ b) metoda, která umožňuje odhadnout vlastnosti výběru na základě znalostí vlastností populace,
- ☐ c) zobecnění statistických výsledků získaných zpracováním výběru na celou populaci,
- ☒ d) metoda sběru dat.

5. Zákon velkých čísel v důsledku říká, že při dostatečném rozsahu výběru
- a) má průměr normální rozdělení,
 - b) má průměr Studentovo rozdělení,
 - c) se střední hodnota přibližuje teoretické hodnotě průměru,
 - d) se relativní četnost přibližuje teoretické hodnotě pravděpodobnosti.

5. Zákon velkých čísel v důsledku říká, že při dostatečném rozsahu výběru

~~a)~~ má průměr normální rozdělení,

~~b)~~ má průměr Studentovo rozdělení,

~~c)~~ se střední hodnota přibližuje teoretické hodnotě průměru,

☒ d) se relativní četnost přibližuje teoretické hodnotě pravděpodobnosti.

6. Pro modelování průměru výběru dostatečně velkého rozsahu je vhodné použít rozdělení

- a) normální,
- b) Pearsonovo (χ^2),
- c) Studentovo,
- d) Fisher-Snedecorovo.

6. Pro modelování průměru výběru dostatečně velkého rozsahu je vhodné použít rozdělení

- ☒ a) normální,
- ☐ b) Pearsonovo (χ^2),
- ☐ c) Studentovo,
- ☐ d) Fisher-Snedecorovo.

7. Pro modelování průměru výběru malého rozsahu je vhodné použít rozdělení
- a) normální,
 - b) Pearsonovo (χ^2),
 - c) Studentovo,
 - d) Fisher-Snedecorovo.

7. Pro modelování průměru výběru malého rozsahu je vhodné použít rozdělení

~~a)~~ normální,

~~b)~~ Pearsonovo (χ^2),

☒ c) Studentovo,

~~d)~~ Fisher-Snedecorovo.

8. Pro modelování relativní četnosti ve výběru o dostatečném rozsahu je vhodné použít rozdělení
- a) normální,
 - b) Pearsonovo (χ^2),
 - c) Studentovo,
 - d) Fisher-Snedecorovo.

8. Pro modelování relativní četnosti ve výběru o dostatečném rozsahu je vhodné použít rozdělení

☒ a) normální,

☐ b) Pearsonovo (χ^2),

☐ c) Studentovo,

☐ d) Fisher-Snedecorovo.

9. Pro modelování rozptylu výběru z normálního rozdělení je vhodné použít rozdělení
- a) normální,
 - b) Pearsonovo (χ^2),
 - c) Studentovo,
 - d) Fisher-Snedecorovo.

9. Pro modelování rozptylu výběru z normálního rozdělení je vhodné použít rozdělení

~~a)~~ normální,

☒ b) Pearsonovo (χ^2),

~~c)~~ Studentovo,

~~d)~~ Fisher-Snedecorovo.

10. Pro modelování poměru rozptylů dvou výběrů z normálního rozdělení je vhodné použít rozdělení

- a) normální,
- b) Pearsonovo (χ^2),
- c) Studentovo,
- d) Fisher-Snedecorovo.

10. Pro modelování poměru rozptylů dvou výběrů z normálního rozdělení je vhodné použít rozdělení

~~a)~~ normální,

~~b)~~ Pearsonovo (χ^2),

~~c)~~ Studentovo,

☒ d) Fisher-Snedecorovo.