STA2 - zk (60 min) ....

jméno a příjmení

Varianta S1

Teoretické otázky:

1) Je pravdivé tvrzení, že 70 %-ní kvantil normovaného normálního rozdělení je záporné číslo? Vysvětlete svůj názor.
2) Jakou metodu byste použili k posouzení, jestli je prodej určitého zboží v průběhu dne rovnoměrně rozložen?
3) Co je multikolinearita? Kdy začíná být škodlivá?
4) Charakterizujte individuální jednoduché indexy. Uveďte příklad jejich použití.
5) Podle čeho se rozhodnete, jakou funkci zvolit k vyrovnání časové řady a jak ověříte, zda vaše volba byla správná?

Příklady:

1) Náhodná veličina má rozdělení s hustotou pravděpodobnosti

$$f(x) = cx^2 \qquad \text{pro } 0 < x < 2\,,$$
$$= 0 \qquad \text{pro ostatní x.}$$

Stanovte:
   a) konstantu c ,
   b) rozptyl X ,
   c) distribuční funkci F (x). 3 BODY

2) Máme odhadnout podíl pracovníků, kteří mají zájem o zvyšování kvalifikace. Předpokládáme, že tento podíl nepřesáhne 20 %. Jak velký máme zvolit rozsah výběru, žádáme-li, aby při 95 %-ní spolehlivosti nepřekročila přípustná chyba 2 %? 1 BOD

3) O produkci podniku, který vyrábí dva druhy výrobků, máte v následující tabulce tyto údaje za rok 2000 a 2001:

Výrobek	**Hodnota výroby v tis. Kč**		**Individuální jednoduchý index množství roku 2001/2000**
	2000	**2001**	
pastelky	60	88	1,6
fixy	72	57,6	0,9

Určete, jak se celkově změnila úroveň prodejních cen ve firmě, jak se změnilo množství vyrobeného a prodaného zboží ve sledovaném období a jak se změnila hodnota výroby. Dále zhodnoťte, jak se na změně hodnoty výroby podílel faktor cen a jak faktor prodaného množství výrobků. 3 BODY

Celkový počet bodů: 12
Minimální počet bodů pro úspěšné složení zkoušky: 7,25
Minimální počet bodů z teorie: 1,75
Minimální počet bodů z příkladů: 2,5
Každá teoretická otázka je za 1 bod.

Teoretické otázky - řešení:

1) Není to pravda. Normované normální rozdělení je symetrické podle nuly, proto medián je roven právě hodnotě 0. Tím pádem bude 70% kvantil tohoto rozdělení kladný, protože je větší než medián. Pozn.: Lze ověřit ve statistických tabulkách, které jsou ke zkoušce povolené, a na základě toho odvodit jednoduše vysvětlení.
2) Lze použít chí-kvadrát test dobré shody nebo Kolmogorovův-Smirnovův test.
3) Jedná se o lineární závislost mezi vysvětlujícími proměnnými ve vícerozměrném regresním modelu. Škodlivá začíná být tehdy, když párový korelační koeficient mezi dvojicí některých vysvětlujících proměnných je větší než 0,8 (někdy se uvádí i 0,75).
4) Slouží ke srovnávání dvou hodnot téhož ukazatele, který není složen z dílčích částí. Příklad: srovnání prodaného množství jednoho produktu u jedné firmy v aktuálním roce s množstvím téhož produktu (téže firmy) v roce předchozím.
5) Rozhodnutí – podle bodového diagramu vyberu funkci nebo funkce, které vypadají, že nejlépe přiléhají skutečným hodnotám. Ověřím pomocí celkového F-testu o vhodnosti vybrané trendové funkce a pomocí individuálních t-testů, kterými testuji přínos parametrů zvolené trendové funkce.

Příklady – řešení:

1) a) Aby funkce byla hustotou pravděpodobnosti, musí platit obecně:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Konkrétně zde:

$$\int_{0}^{2} cx^2dx = 1$$

Vhodné bude nejprve vyřešit integrál a pak výsledek položit rovno 1.

$$c\int_{0}^{2} x^2dx = c\left[\frac{x^3}{3}\right]_0^2 = c\cdot\frac{8}{3}$$

Nyní výsledek položím rovno 1:

$$\frac{8}{3}c = 1$$

$$8c = 3$$

$$c = \frac{3}{8}$$

Pro konstantu 3/8 je tato funkce hustotou pravděpodobnosti, tj.:

$f(x) = \frac{3}{8}x^2$ pro $0 < x < 2$

$= 0$ pro ost. x

b) Rozptyl spojité náhodné veličiny lze vypočítat obecně jako:

$$D(X) = \int_{-\infty}^{\infty} x^2f(x)dx - \left[\int_{-\infty}^{\infty} xf(x)dx\right]^2$$

$$D(X)=\int_0^2 x^2\frac{3}{8}x^2dx-\left[\int_0^2 x\frac{3}{8}x^2dx\right]^2=\frac{3}{8}\int_0^2 x^4\,dx-\left[\frac{3}{8}\int_0^2 x^3\,dx\right]^2=\frac{3}{8}\left[\frac{x^5}{5}\right]_0^2-\left[\frac{3}{8}\left[\frac{x^4}{4}\right]_0^2\right]^2=$$

$$=\frac{3}{8}\cdot\frac{32}{5}-\left[\frac{3}{8}\cdot\frac{16}{4}\right]^2=\frac{12}{5}-\frac{9}{4}=\frac{3}{20}=0{,}15$$

Rozptyl této náhodné veličiny je 0,15.

c) Vztah distribuční funkce a hustoty pravděpodobnosti je obecně popsán vztahem:

$$F(x)=\int_{-\infty}^{x} f(t)dt$$

Zde konkrétně:

$$F(x)=\int_0^x \frac{3}{8}t^2\,dt=\frac{3}{8}\int_0^x t^2\,dt=\frac{3}{8}\left[\frac{t^3}{3}\right]_0^x=\frac{1}{8}x^3$$

Distribuční funkce této náhodné velečiny je:

$$F(x)=\begin{cases}=0 & \text{pro } x\le 0\\ \dfrac{x^3}{8} & pro\ 0<x<2\\ =1 & \text{pro } x\ge 2\end{cases}$$

2) Výstupem, který hledáme, je minimální rozsah výběru pro dané šetření. Známé skutečnosti:

p = 0,2 (odhadovaný podíl zájemců o zvyšování kvalifikace)
1-α = 0,95 (požadovaná spolehlivost odhadu)
Δ = 0,02 (přípustná chyba odhadu)

Odhadujeme relativní četnost, proto vycházíme ze vzorce přípustné chyby odhadu relativní četnosti, jehož úpravou dostaneme:

$$n\ge\frac{u^2_{1-\frac{\alpha}{2}}\cdot p(1-p)}{\Delta^2}$$

Ještě je potřeba zjistit v tabulkách či statistickém programu hodnotu kvantilu normovaného normálního rozdělení $u_{1-\frac{\alpha}{2}}$. Jde o kvantil $u_{0,975}=1{,}96$.

$$n\ge\frac{1{,}96^2\cdot 0{,}2\cdot 0{,}8}{0{,}02^2}$$
$$n\ge 1537$$

Výběr by mělo tvořit alespoň 1537 pracovníků.

3) Popis známých údajů:

Výrobek	**Hodnota výroby v tis. Kč**		**Individuální jednoduchý index množství roku 2001/2000 ($i_q=q_1/q_0$)**
	2000 (Q_0)	**2001 (Q_1)**	
pastelky	60	88	1,6
fixy	72	57,6	0,9

Máme dva různé produkty, jejichž množství ani cenu tudíž nemůžeme shrnovat, pracujeme tedy s nestejnorodým ukazatelem. Pro charakterizování změn úrovně prodejních cen, hodnoty výroby a množství vyrobených výrobků je vhodné použít souhrnné indexy.

- Změna vyrobeného množství – můžeme použít Laspeyresův nebo Paascheho objemový index, ovšem v upraveném tvaru:

  $${}_{P}I_q = \frac{\sum Q_1}{\sum \frac{Q_1}{i_q}}$$

  $${}_{P}I_q = \frac{\sum Q_1}{\sum \frac{Q_1}{i_q}} = \frac{145{,}6}{119} = 1{,}224$$

  Ve sledovaném období došlo k nárůstu vyrobeného množství výrobků o 22,4 %, bereme-li v úvahu ceny běžného období (rok 2001).
  NEBO

  $${}_{L}I_q = \frac{\sum i_q \cdot Q_0}{\sum Q_0} = \frac{160{,}8}{132} = 1{,}218$$

  Ve sledovaném období došlo k nárůstu vyrobeného množství výrobků o 21,8 %, bereme-li v úvahu ceny základního období (rok 2000).

- Změna hodnoty výroby – použijeme souhrnný hodnotový index:

  $$I_Q = \frac{\sum Q_1}{\sum Q_0} = \frac{145{,}6}{132} = 1{,}103$$

  Hodnota výroby ve sledovaném období vzrostla o 10,3 %.

- Při hodnocení změny cenové úrovně využijeme toho, co víme o rozkladu souhrnného indexu hodnoty:

  $$I_Q = {}_{P}I_p \cdot {}_{L}I_q$$

  $$1{,}103 = {}_{P}I_p \cdot 1{,}218$$
  $${}_{P}I_p = 0{,}906$$

  Ve sledovaném období došlo k poklesu cen výrobků o 9,4 %, bereme-li v úvahu vyrobené množství výrobků jako v běžném období, zde tedy v roce 2001.

  NEBO

  $$I_Q = {}_{L}I_p \cdot {}_{P}I_q$$
  $$1{,}103 = {}_{L}I_p \cdot 1{,}224$$
  $${}_{L}I_p = 0{,}901$$

  Ve sledovaném období došlo k poklesu cen výrobků o 9,9 %, bereme-li v úvahu vyrobené množství výrobků jako v základním období, zde tedy v roce 2000.

- Pro zhodnocení, jak se na změně hodnoty výroby podílel faktor cen a jak faktor prodaného množství výrobků využijeme předchozí výpočty, tj. rozklad souhrnného hodnotového indexu metodou postupných změn:

  $$1{,}103 = 0{,}906 \cdot 1{,}218$$

  Změna cen způsobila pokles hodnoty výroby o 9,4 %, změna vyrobeného množství výrobků způsobila nárůst hodnoty výroby o 21,8 %.

  Pozn.: Lze komentovat i druhou podobu rozkladu.