

Varianta S4

Teoretické otázky:

- 1) K čemu slouží histogram rozdělení četností?
- 2) Stručně popište princip rozkladu souhrnného indexu hodnoty metodou postupných změn. Jaký má tento rozklad význam?
- 3) Časová řada měsíčních údajů je rostoucí a sezónní výkyvy rostou úměrně trendu. Jaký byste použili model časové řady a jak byste posoudili sezónnost takové časové řady?
- 4) Jakým způsobem lze snížit pravděpodobnost chyby I. druhu a jak pravděpodobnost chyby II. druhu (při dané chybě I. druhu)?
- 5) Co to znamená, když mezi 2 alternativními proměnnými existuje úplná záporná asociace? Vysvětlete!

Příklady:

- 1) Mezi studenty vysoké školy proběhla anketa, která měla za cíl zjistit, z jakého důvodu se rozhodli na této škole studovat. Jejich odpovědi jsou zaznamenány v následující tabulce. Vypočítejte obě charakteristiky variability, které znáte. Interpretujte jejich hodnoty. Potom jejich hodnoty porovnejte – uveďte, proč se liší, a které byste dali při výpočtu přednost a proč. 2 BODY

Důvod studia x_i	Počet studentů n_i
další sebevzdělávání	10
odklad vojny	2
prodloužení mládí	8
perspektiva při hledání zaměstnání	9
ostatní	3
Součet	32

- 2) V následující tabulce jsou uvedeny údaje o počtu dětí v domácnosti a výdajích na školní potřeby v tis. Kč za rok. Pomocí analýzy rozptylu posuďte na hladině významnosti 5 %, zda existuje závislost výdajů na školní potřeby na počtu dětí v domácnosti. Případně změřte sílu závislosti vhodnou charakteristikou (Není třeba ověřovat podmínky pro použití analýzy rozptylu). 2 BODY

Počet dětí v domácnosti x_i	Výdaje na školní potřeby v tis. Kč za rok y_{ij}	n_i
1	12 13,5 7 9,3 8,1 10,2 9 9,8	8
2	18 19,5 24 20,1 22,7 17,5 14,8 16,2 18,4 21,5	10
3	24,3 28,7 27,1 32 30,4 23 26,5	7
4	35 42,5 38,6 39,2 46	5
Součet		30

- 3) V níže uvedené tabulce jsou údaje o prodeji 15-litrových plastových odpadkových košů v několika pobočkách určité firmy v lednu a srpnu 2003.
- a) Určete, jak se změnila tržba, počet prodaných kusů a průměrná cena v srpnu oproti lednu.
 b) Určete vliv změny cen v jednotlivých pobočkách a vliv změny struktury prodaných výrobků na změnu průměrné ceny.

3 BODY

Pobočka	Cena za jednotku v Kč		Tržby v Kč	
	leden	srpen	leden	srpen
Brno	120	145	26 400	24 070
Jihlava	109	120	20 056	21 600
Plzeň	112	131	21 952	26 462
Liberec	105	126	22 260	28 980

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) Slouží ke grafickému znázornění intervalového rozdělení četností.
- 2) Rozklad I_Q může mít dvojí podobu:

$$I_Q = {}^p I_p \cdot {}^L I_q$$

NEBO

$$I_Q = {}^L I_p \cdot {}^p I_q$$

Význam: Je možné určit, jak se na změně hodnoty podílel faktor cen a jak faktor množství.

- 3) Použili bychom model proporcionální sezónnosti a velikost sezónních výkyvů bychom vyčíslili prostřednictvím indexních sezónních faktorů.
- 4) Snížit pravděpodobnost chyby I. druhu můžeme sami – tím, že si koeficient α zvolíme menší, např. z 0,05 snížím na 0,01. Musíme ale přitom vzít v úvahu, že se při této aktivitě zároveň zvýší pravděpodobnost chyby II. druhu. Pravděpodobnost chyby II. druhu (při dané chybě I. druhu) snížíme pomocí zvětšení rozsahu výběru.
- 5) Znamená to, že pokud nastává jeden z jevů, druhý jev nenastává. Je to v případě, kdy koeficient asociace je roven -1. Použití v analýze zkoumání závislostí (alternativních znaků).

Příklady – řešení:

- 1) Tabulka rozdělení četností obsahuje údaje o nominální proměnné, proto budeme počítat míru mutability a nominální varianci.

Důvod studia x_i	Počet studentů n_i	n_i^2	p_i	p_i^2
další sebevzdělávání	10	100	0,3125	0,0977
odklad vojny	2	4	0,0625	0,0039
prodloužení mládí	8	64	0,2500	0,0625
perspektiva při hledání zaměstnání	9	81	0,2813	0,0791
ostatní	3	9	0,0938	0,0088
Součet	32	258	1,0000	0,252

Míru mutability vypočítáme podle:

$$M = \frac{n^2 - \sum_{i=1}^k n_i^2}{n(n-1)} = \frac{32^2 - 258}{32 \cdot 31} = \frac{766}{992} = 0,772$$

Pozn.: Pomocný výpočet v tabulce výše v modré barvě.

77,2 % dvojic studentů se odlišuje z pohledu důvodu studia na VŠ. Variabilita důvodu studia je poměrně vysoká.

Nominální varianci vypočítáme podle:

$$NOMVAR = 1 - \sum_{i=1}^k p_i^2$$

$$NOMVAR = 1 - 0,252 = 0,748$$

Pozn.: Pomocný výpočet v tabulce výše v hnědé barvě.

Variabilita důvodu studia je poměrně vysoká.

K charakterizování variability důvodu studia je zde vhodnější použít míru mutability, protože NOMVAR je charakteristika, která skutečnou variabilitu hodnot nominální proměnné podhodnocuje.

- 2) Doporučeno řešit ve statistickém programu.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

(NEBO H_0 : Výdaje na školní potřeby nejsou závislé na počtu dětí v domácnosti)

$$H_1: \text{non } H_0$$

Výstup z programu STATGRAPHICS Centurion 18:

ANOVA Table for výdaje skol potr by Pocet deti

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	3125,7	3	1041,9	114,66	0,0000
Within groups	236,266	26	9,08716		
Total (Corr.)	3361,97	29			

$$F = 114,66$$

$$P\text{-Value} = 0,0000$$

$P\text{-Value} < \alpha$, proto zamítáme H_0 a přijímáme H_1 .

Na hladině významnosti 5 % jsme prokázali, že výše výdajů na školní potřeby závisí na počtu dětí v domácnosti.

Sílu závislosti lze změřit pomocí poměru determinace, tj.

$$p^2 = \frac{S_{ym}}{S_y}$$

$$p^2 = \frac{S_{ym}}{S_y} = \frac{3125,7}{3361,97} = 0,930$$

Pozn.: S_{ym} Hodnotu najdeme v tabulce výše (modrá barva)

S_y Hodnotu najdeme v tabulce výše (růžová barva)

93 % z celkové variability výdajů na školní potřeby můžeme vysvětlit pomocí počtu dětí v domácnosti, tj. závislost výše výdajů na školní potřeby na počtu dětí v domácnosti je velmi silná.

- 3) Pracujeme se stejnorodým ukazatelem, jehož dílčí části lze shrnovat buď pomocí součtu (množství, tržby) nebo pomocí průměru (cena), proto k charakterizování změn indikátorů použijeme individuální složené indexy.

Pobočka	Cena za jednotku v Kč		Tržby v Kč		Počet prodaných kusů		p_0q_1	p_1q_0
	leden	srpen	leden	srpen	leden	srpen		
Brno	120	145	26 400	24 070	220	166	19920	31900
Jihlava	109	120	20 056	21 600	184	180	19620	22080
Plzeň	112	131	21 952	26 462	196	202	22624	25676
Liberec	105	126	22 260	28 980	212	230	24150	26712
Celkem	x	x	90668	101112	812	778	86314	106368

Pozn.: Pomocné výpočty jsou v tabulce uvedeny modře.

a)

$$I_Q = \frac{\sum Q_1}{\sum Q_0} = \frac{101112}{90668} = 1,115$$

$$\Delta_Q = \sum Q_1 - \sum Q_0 = 101112 - 90668 = 10444$$

Tržba se v srpnu zvýšila oproti lednu o 11,5%, tedy o 10 444 Kč.

$$I_q = \frac{\sum q_1}{\sum q_0} = \frac{778}{812} = 0,958$$

$$\Delta_q = \sum q_1 - \sum q_0 = 778 - 812 = -34$$

Počet prodaných kusů se snížil v srpnu oproti lednu o 4,2%, tedy o 34 ks.

$$I_{\bar{p}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\frac{\sum Q_1}{\sum q_1}}{\frac{\sum Q_0}{\sum q_0}} = \frac{\frac{101112}{778}}{\frac{90668}{812}} = \frac{129,964}{111,66} = 1,164$$

$$\Delta_{\bar{p}} = \bar{p}_1 - \bar{p}_0 = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0} = 129,964 - 111,66 = 18,3$$

Průměrná cena kusů se zvýšila v srpnu oproti lednu o 16,4%, tedy o 18,30 Kč.

b)

$$I_{\bar{p}} = I_{SS}(q_1) \cdot I_{STR}(p_0)$$

$$I_{SS}(q_1) = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_1}{\sum q_1}} = \frac{\frac{101112}{778}}{\frac{86314}{778}} = \frac{129,964}{110,943} = 1,171$$

$$I_{STR}(p_0) = \frac{\frac{\sum p_0 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{86314}{778}}{\frac{90668}{812}} = \frac{110,943}{111,66} = 0,994$$

$$I_{\bar{p}} = I_{SS}(q_1) \cdot I_{STR}(p_0) = 1,171 \cdot 0,994 = 1,164$$

$$\Delta_{\bar{p}} = (129,964 - 110,943) + (110,943 - 111,66) = 19,021 + (-0,717) = 18,3$$

V důsledku změn cen na jednotlivých pobočkách (při zachování struktury prodeje z běžného období) vzrostla průměrná cena o 17,1%, tj. o 19 Kč a změna struktury prodaného zboží (při cenové hladině základního období) vyvolala pokles průměrné ceny o 0,6%, tj. o 0,7 Kč.

NEBO

$$I_{\bar{p}} = I_{SS}(q_0) \cdot I_{STR}(p_1)$$

$$I_{SS}(q_0) = \frac{\frac{\sum p_1 q_0}{\sum q_0}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{106368}{812}}{\frac{90668}{812}} = \frac{130,995}{111,66} = 1,173$$

$$I_{STR}(p_1) = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_1 q_0}{\sum q_0}} = \frac{\frac{101112}{778}}{\frac{106368}{812}} = \frac{129,964}{130,995} = 0,992$$

$$I_{\bar{p}} = I_{SS}(q_0) \cdot I_{STR}(p_1) = 1,173 \cdot 0,992 = 1,164$$

$$\Delta_{\bar{p}} = (130,995 - 111,66) + (129,964 - 130,995) = 19,335 + (-1,031) = 18,3$$

V důsledku změn cen na jednotlivých pobočkách (při zachování struktury prodeje ze základního období) vzrostla průměrná cena o 17,3%, tj. o 19,3 Kč a změna struktury prodaného zboží (při cenové hladině běžného období) vyvolala pokles průměrné ceny o 0,8%, tj. o 1 Kč.