

Cvičení č.8

Název:

**Testování statistických hypotéz
Závislost náhodných veličin**

Termín odevzdání: 3.5.2009 (do IS)

Cíl:

Naučit se s pomocí MS EXCEL testovat statistické hypotézy, vypočítat korelační koeficient a zkonstruovat regresní přímku

Úkol A: Testování statistických hypotéz

1. **zapište obecný postup testování** (viz. přednáška nebo skripta BRÁZDIL A kol.)
2. Pro soubory dat Procentuelní zastoupení narození dětí v ČR a Procentuelní zastoupení narození dětí v ČR v Brno - venkov zjistěte, zda oba soubory lze považovat za shodné podle vybraných testů :
 - a. stanovte hladinu významnosti pro porovnání souborů na 95 %.
 - b. formulujte nulovou hypotézu pro shodu souborů.
 - c. porovnejte shodnost rozložení četností souboru pro Brno – venkov se souborem pro ČR pomocí testu CHÍ – kvadrát (CHITEST)
 - D. Testujte významnost rozdílu rozptylů souborů pomocí F –testu (FTEST)
 - e. Testujte významnost rozdílu pomocí T - testu.(TTEST)
 - f. Přijměte nebo odmítněte nulovou hypotézu.
3. Výsledky **přehledně zpracujte** do protokolu

Pozn.: Pracujte s podporou statist. funkcí MS Excel. Použijte výše jmenované funkce.

Tab. 1 : Rozdělení počtu narozených dětí v ČR okrese Brno – venkov v roce 2005, cvičná data

měsíc	% dětí narozených v ČR	% dětí narozených v okrese Brno - venkov
1	8,39	8,52
2	7,91	8,81
3	9,02	9,01
4	9,03	8,72
5	9,15	9,12
6	8,64	7,94
7	8,45	7,84
8	8,04	7,93
9	8,28	7,74
10	7,93	8,13
11	7,41	7,44
12	7,75	8,81

- Postup:
1. připravíme tabulku dat - zde % dětí v jednotlivých měsících
 2. zajímá nás, nakolik jsou tyto soubory shodné
 3. k tomuto zjištění použijeme testy shody
 4. hladina významnosti je pro nás 95 %, tj soubory považujeme za shodné,

pokud výsledky testování ukáží shodu min. 95 %

5. hypotéza : Soubory jsou shodné na hladině významnosti 95%

6. použití testů - Chi - test, T - test, F - test, pozorně pracovat s nápovědou F test

7. porovnat pravděpodobnosti z testů s námi zvolenou hladinou významnosti.

8. Pokud splňují výsledky naši podmínku, pak můžeme přijmout naši hypotézu.

9. Pokud nespĺňují, musíme ji odmítnout s tím, že na zvolené hladině významnosti 95 % nevyhověly podmínce při použití vybraných testů a tedy se neshodují.

10. zformulujeme odpověď

Úkol B: Výpočet korelačního koeficientu, konstrukce regresní přímky

1. **Vypočítejte koeficienty korelace** mezi soubory s pomocí MS Excel a statistické funkce Correl. Zpracujte do tabulky.
2. **Vyhodnoťte podle koeficientu korelace závislost klimatických charakteristik.**
3. Sestrojte **bodový graf** pro dvojici **nejvíce závislých st. souborů**, sestrojte **regresní přímku (v Excelu)**
4. **Zpracujte protokol s obvyklými náležitostmi**

Tab. 2 Klimatické charakteristiky naměřené na meteorologické stanici Brno Tuřany v roce 2002

měsíc	1	2	3	4	5	6	7	8	9	10	11	12
prům. teplota, ; °C	-1,2	3,6	5,8	9,4	17,1	19,1	20,9	20,4	14,0	7,6	6,0	3,1
prům úhrn srážek; mm	8,1	21,3	21,0	28,6	45,8	81,7	58,0	91,2	39,2	71,9	48,2	21,5
délka slun. svitu; h	55,6	82,7	183,4	169,5	238,3	291,4	288,0	221,2	174,5	89,4	44,7	41,3

Pozn. Ukázka zpracování grafu, bodový, upravit dle potřeby protínání osy y, vložení regresní přímky pomocí nástroje "přidat spojnici trendu".

