

## Zadání zápočtového příkladu ze STATISTIKY II

Každý student připraví soubor ekonomických dat podle vlastního výběru a dodrží při tom podmínky na tvar a velikost souboru uvedené dále. Výběr dat lze provést podle vlastního uvážení. Lze využít statistické ročenky, vlastní zdroje nebo dostupné databáze, např.:

Databázi německé firmy DSI Data Service, která obsahuje následující databáze:

World Statistics;  
OECD Statistics;  
Intra- Extra-OECD Trade;  
Eurostat Statistics;  
Intra- and extra-EU Trade Statistics a  
German Statistical Office, German FED.

Všechny DB obsahují téměř nepřehledné množství statistických dat i v retrospektivě. Např. vyhledání indexu spotřebitelských cen třeba v Bangladéši za roky 1986 - 2001 tak nepředstavuje žádný vážnější problém.

Do DB můžete vstoupit na následující adrese:  
<http://195.145.59.167/ISAPI/LogIn.dll/login?lg=e>

Soubor dat, které si pro zpracování připravíte, bude obsahovat realizace 3 náhodných výběrů z rozdělení náhodných veličin X, Y a Z, rozsah každého výběru bude  $n=30$ . Datový soubor každý student připraví na disketu se jménem s12345.txt, kde místo 12345 zapíše své identifikační číslo. Při přípravě datového souboru je nutné dodržet následující pokyny:

### Popis datového souboru:

Vzorový soubor s názvem "s12345.txt" patří studentovi, jehož uco je 12345 (viz také první sloupec v uvedeném souboru).

KAZDY soubor má 30 radku, za posledním radkem MUSÍ BYT 1 ENTER (ne víc !!!).

Máte-li více dat, použijte jen napr. posledních 30 údajů

Jednotlivé sloupce jsou od sebe odděleny tabulátorem (nikoli hromadou mezer!!!).

První sloupec obsahuje 30x pod sebou studentovo UCO.

Sloupce 2, 3 a 4 obsahují data (tj. hodnoty tří veličin X, Y a Z, které si vyberete).

Při zápisu desetinného čísla použijte DESETINNOU TECKU, nikoli čárku.

Jiný formát dat není přípustný a bude považován za nevyhovující, protože by při zpracování výsledku působil značné komplikace.

Takto zpracovaný soubor odevzdate na disketu, na jejíž popisce bude vaše jméno, UCO a název souboru (pro kontrolu). Disketa vám bude samozřejmě vrácena.

Pokud necemu nerozumíte, NEBOJTE SE A ZAVCAS SE ZEPTEJTE - kamaradu nebo cvicicich.

**Vzorový příklad datového souboru studenta, jehož UCO je 12345:**

12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4
12345	1.0	4.0	7.0
12345	2.1	5.1	8.1
12345	3.2	6.3	9.4

**Zpracování datového souboru musí obsahovat:**

1. Ekonomický popis všech tří proměnných  $X$ ,  $Y$  a  $Z$  a důvody, proč je žádoucí a vhodné právě tyto proměnné statisticky zpracovat.

2. Výpočet průměru, rozptylu, směrodatné odchylky, 1.-4. obecného i centrálního momentu, šikmosti, špičatosti a mediánu.

3. Graf výběrové distribuční funkce a histogram rozdělení četností pro každou proměnnou.

4. Intervaly spolehlivosti pro střední hodnotu a pro směrodatnou odchylku každé z proměnných  $X$ ,  $Y$  a  $Z$ . Je třeba dbát, aby výběry byly vybrány z normálního, alternativního nebo Poissonova rozdělení, aby byly splněny předpoklady pro konstrukci intervalů spolehlivosti.

5. Intervaly spolehlivosti pro rozdíl středních hodnot veličin  $X$  a  $Y$  nebo  $X$  a  $Z$  nebo  $Y$  a  $Z$ . Opět je potřeba dbát na splnění

matematických předpokladů. Tyto intervaly spolehlivosti stačí konstruovat pro jednu vybranou dvojici proměnných.

6. Závěr statistické analýzy a ekonomická interpretace získaných výsledků.

**Zpracování** doporučuji provést na počítači a použít software, který byl probírán v dřívějších kurzech, např. Excel nebo MATLAB. Lze použít i jiný software, který studentovi vyhovuje.

**Termín pro odevzdání práce** cvičícímu je pátek 14.5. Po tomto termínu již nebudou žádné práce akceptovány s výjimkou prací studentů, kteří předloží potvrzení o pracovní neschopnosti.

Každý student odevzdá disketu s popisem a zdaným textovým souborem podle výše uvedených pokynů a dále počítačový výtisk zpracování úlohy. Zpracování musí obsahovat všech 6 položek uvedených v odstavci **Zpracování datového souboru** včetně zvoleného datového souboru.

**Práce bude součástí vašeho závěrečného hodnocení.**

**Ve vašem úsilí přeji hodně zdaru, šťastnou ruku při volbě dat a hodně dobrých nápadů při statistické analýze.**

Na vaše výsledky se těší

Jaroslav Michálek spolu se všemi cvičícími