

Harmonogram předmětu MV011 Statistika I pro JS 2014

1. 17.2. Tabulkové a grafické zpracování datových souborů.
2. 24.2. Funkcionální a číselné charakteristiky znaků.
3. 3.3. Statistický software, základy práce v SAS.
4. 10.3. Popisná statistika v MS Excel a SAS.
5. 17.3. Regresní analýza v MS Excelu a SAS.
6. 24.3. Úvod do teorie pravděpodobnosti.
7. 31.3. Náhodné veličiny (NV).
(v rámci cvičení (31.3., 2.4. a 3.4. zápočtový test).
8. 7.4. Diskrétní a spojité náhodné veličiny, vybraná rozložení NV.
9. 14.4. Stochasticky nezávislé náhodné veličiny, transformace NV, generování realizací NV.
10. 21.4. Velikonoce
25.4. deadline pro odevzdání projektu.
11. 28.4. Číselné charakteristiky NV.
1.5. svátek ...cvičení odpadá (samostudium)
12. 5.5. Slabý zákon velkých čísel a centrální limitní věta, úvod do testování hypotéz.
8.5. svátek ...cvičení odpadá (samostudium)
9.5. deadline pro odevzdání opraveného neuznaného projektu
13. 12.5. Testování hypotéz v MS Excel a SAS.
12. 14. a 15. 5. oprava zápočtového testu (v rámci cvičení)
14. 19.5. Zkouška - předtermín
Cvičení 19., 21. a 22. 5. - konzultace

Základní literatura:

- Řezáč, M. *Statistika I*, 2014. <http://www.math.muni.cz/~mrezac/vyuka/StatistikaI.pdf>

- Budíková, Marie - Králová, Maria - Maroš, Bohumil. *Průvodce základními statistickými metodami*. vydání první. Praha : Grada Publishing, a.s., 2010. 272 s. edice Expert. ISBN 978-80-247-3243-5.
- Budíková, Marie - Mikoláš, Štěpán - Osecký, Pavel. *Popisná statistika*. 3., doplněné vyd. Brno : Masarykova univerzita, 1998. 52 s. ISBN 80-210-1831-3.
- Budíková, Marie - Mikoláš, Štěpán - Osecký, Pavel. *Teorie pravděpodobnosti a matematická statistika. Sbírká příkladů*. 3. vyd. Brno : Masarykova univerzita, 2004. 127 s. ISBN 80-210-3313-4.

Podmínky pro absolvování předmětu MV011 Statistika I

Účast na cvičeních je povinná. Případnou neúčast je třeba omluvit oficiální cestou.

Vnitrosemestrální test („zápočtový test“): Výsledek se započítává do zkoušky. Úspěšné zvládnutí písemky předpokládá zisk aspoň 10 bodů. V případě neúspěchu je možnost jedné opravy.

Písemka potrvá 20 minut, jsou povoleny učební materiály (tj. skripta Popisná statistika, Teorie pravděpodobnosti a matematická statistika – sbírka příkladů, záznamy z přednášek a cvičení). Sestává ze dvou příkladů, celkem 20 bodů.

Možnost skládat zkoušku či získat zápočet je podmíněna dostatečnou účastí na cvičeních, úspěšným zvládnutím vnitrosemestrální písemky a uznáním semestrálního projektu.

Zkouška potrvá 90 minut, je výhradně písemná, jsou povoleny učební materiály (tj. skripta Popisná statistika, Teorie pravděpodobnosti a matematická statistika – sbírka příkladů, záznamy z přednášek a cvičení). Je tvořena čtyřmi příklady, každý za 20 bodů, celkem 80 bodů.

Hodnocení zkoušky:

(90, 100] ... A, (80, 90] ... B, (70, 80] ... C, (60, 70] ... D, (50, 60] ... E, [0, 50] ... F

Hodnocení zápočtu:

(45, 100] ... Z, [0, 45] ... N

Pokyny k semestrálnímu projektu

Datový soubor je uložen v ISu ve Studijních materiálech. Soubor german1.xlsx jsou vlastní data, soubor german.doc obsahuje popis dat.

Projekt zpracujte pomocí software SAS. Výstupy včetně tabulek a grafů vložte do textového editoru, kde k nim napíšete komentář.

Projekt odevzdejte v elektronické podobě (formát **.pdf**) do **Odevzdávárny** nejpozději do **25.4.**, každý student jeden soubor. Později odevzdané projekty nebudou uznány. **Uznání projektu je nutnou podmínkou** připuštění ke zkoušce. Hodnocení projektu najdete v **Poznámkovém bloku**. Nebude-li projekt uznán, máte jednu možnost opravy, a to do **9.5.**

Nepoužívejte řešení někoho jiného, existují prostředky na odhalení podvodu. Podvodem se nic nenaučíte.

Nevytvářejte příliš rozsáhlá řešení, 10 stránek určitě stačí.

Snažte se, aby Vaše řešení bylo pěkně upravené, přehledné, se správným pravopisem a celkově kultivované.

3 nejlépe zpracované projekty v každé seminární skupině obdrží **bonus ve výši 10 bodů**.

Zadání semestrálního projektu

K dispozici je 1000 záznamů o žadatelích o úvěr.

Úkol 1.: Zjistěte absolutní a relativní četnosti proměnných „class“ (klasifikátor dobrého klienta) a „at7“ (délka současného zaměstnání), přičemž pro proměnnou „at7“ zjistěte též kumulativní absolutní a relativní četnosti. Pro proměnnou „class“ vytvořte sloupcový diagram, pro proměnnou „at7“ polygon četností.

Úkol 2.: Pro proměnné „at9“ (pohlaví a rodinný status) a „at4“ (účel úvěru) sestavte kontingenční tabulky absolutních a relativních četností, sloupcově a řádkově podmíněných relativních četností. Kolik procent žen financuje úvěrem vzdělání? Kolik procent klientů, kteří financovali úvěrem koupi nového auta, jsou muži?

Úkol 3.: Podle Sturgersova pravidla stanovte optimální počet třídících intervalů pro proměnné „at5“ (výše úvěru) a „at13“ (věk) a nakreslete jejich histogramy, a to

- pro celý soubor
- pro dobré klienty (class=1)
- pro špatné klienty (class=2).

Úkol 4.: Vypočtěte minimum, maximum, medián, průměr, směrodatnou odchylku, šikmost a špičatost proměnných „at5“ a „at13“

- pro celý soubor
- pro dobré klienty (class=1)
- pro špatné klienty (class=2).

Úkol 5.: Vypočtěte a interpretujte Pearsonův koeficient korelace proměnných „at5“ a „at13“

- pro celý soubor
- pro dobré klienty (class=1)
- pro špatné klienty (class=2).

Úkol 6.: Najděte rovnici regresní přímky vyjadřující závislost proměnné „at5“ na proměnné „at13“. Jaký je index determinace a co vyjadřuje? Jaká je predikovaná hodnota výše úvěru pro věk 40 let? Nalezenou regresní přímku zakreslete do dvourozměrného tečkového diagramu.

Úkol 7.: Otestujte následující hypotézy:

- průměrná výše úvěru (at5) je rovna hodnotě 3200
- průměrná výše úvěru je shodná pro muže a ženy (at9)
- průměrná výše úvěru je shodná pro dobré a špatné klienty
- účel úvěru (at4) a typ bydlení (at15) jsou nezávislé