

Zápočtová úloha MAI059 2006/2007

11. prosince 2006

V bance je k dispozici 5 nezávislých obslužných míst (přepážek), doba obsluhy jednoho klienta u přepážky se řídí exponenciálním rozdělením se střední hodnotou 6 minut. Klienti přicházejí do banky nezávisle, doba mezi příchody se také řídí exponenciálním rozdělením, ale průměrný počet příchozích klientů (a tedy také střední doba mezi příchody) se v průběhu otevírací doby banky mění následujícím způsobem:

9:00 – 10:30 32 klientů za hodinu;
10:30 – 15:30 26 klientů za hodinu;
15:30 – 17:00 50 klientů za hodinu.

Dále předpokládejme, že žádný klient neodejde z banky kvůli čekání ve frontě a že čekárna je dostatečně velká (neexistuje tedy žádné omezení pro příchody). Konečně předpokládejme, že banka obslouží všechny klienty i po zavírací době, tj. všichni klienti, kteří přišli před 17:00 a v 17:00 stále čekají, budou obslouženi.

Přiložený zdrojový kód v programu R (funkce `banka.jedenden`) simuluje výše popsané příchody a obsluhu klientů během jednoho dne.

Úkoly:

- Proveďte simulační studii pro $D = 250$ dnů (for cyklus). Podrobně okomentujte, jak funguje funkce `banka.jedenden` a které její výstupy si ukládáte pro následující úkoly. Zajímají nás statistické principy, ne syntaktický rozbor.
- Graficky znázorněte průměrnou dobu čekání ve frontě v závislosti na denní době, (i) jsou-li celý den otevřené všechny přepážky, resp. (ii) jsou-li otevřené 4 přepážky do 10:30, 3 přepážky 10:30 – 15:30 a všechny od 15:30. Výsledky okomentujte.
- Navrhněte a interpretujte přibližný 90% interval spolehlivosti pro střední hodnotu času ukončení obsluhy posledního klienta. Je použití CLV teoreticky v pořádku? Je použitá aproximace dostatečně přesná? Proč? (Porovnejte teoretické normální kvantily s empirickými.)
- Otestujte na 5% hladině významnosti hypotézu, že průměrný počet klientů během jednoho dne je roven 253. Úlohu řádně statisticky zformulujte. (Bonus: proč právě 253?)

Nápověda: Pro (c) si uvědomte, že obsluha posledního klienta v jednom dni je ukončena v čase $\max_i T_i$, kde T_i je okamžik ukončení obsluhy i -tého klienta. Nakreslete si histogram $\max_i T_i$ a vizuálně posuďte shodu s normálním rozdělením. Pro (d) použijte náhodné počty klientů v jednotlivých

dnech N_1, \dots, N_D , které tvoří výběr z rozdělení, jehož střední hodnota odpovídá střednímu počtu klientů za den. Pro sestavení intervalu spolehlivosti i testování využijte CLV.

Z eRka se Vám mohou hodit funkce `boxplot`, `for`, `hist`, `length`, `plot`, `qnorm`, `source`, `summary`, `t.test`. Náповědu ke všem funkcím získáte příkazem `help(jmenofunkce)`.

Pokyny k vypracování: Vaše řešení musí obsahovat především **komentář** Vašich statistických úvah s relevantními výstupy simulací a jejich interpretací (**maximálně 3** stránky textu + obrázky). Dále nám zašlete **Váš kód** použitý k analýzám a malování obrázků. Abychom si mohli Vaše výsledky ověřit, nastavte `set.seed(ddmmyyyy)` podle svého data narození před prvním generováním náhodných čísel. Tento použitý seed nezapomeňte uvést ve svém řešení.

Po loňských zkušenostech vyžadujeme komentář v souvislých větách se slušně vysázenou matematikou (nejlépe \TeX em, Word a podobné nevadí, preferujeme PDF). Nebojte se obrázků (vyzkoušejte například `histogram`, `boxplot`).

Jedná se o samostatnou práci. Pokud usoudíme, že se některá řešení navzájem příliš podobají, odměníme autory bonusem minus 30 bodů.

Práci posílejte e-mailem svému cvičícímu. Deadline pro odeslání finální verze práce je **31.3.2007, 23:59**. Za práci je možné získat maximálně 40 bodů, přičemž 25 bodů bude uděleno za statistiku, 8 bodů za komentář k eRkovému kódu a 7 bodů za celkový dojem z práce.

V Praze 11. prosince 2006
JS & KH & LP & MB & MS & OV & PK