Час 1

- Популација, обележје
- узорак, репрезентативност узорка
- типови обележја:
  - квалитативно (категоричко):
    * номинално: крвна група, пол, сешуално опредељење, вериска припадност...
    * ординално: разред, интензитет бола, статус студената (будзет, самофинансирајући)
  - (нумеричко):
    * дискретно: број деце, оцена на испиту, број искори71ених дана одмора...
    * непрекидно: тежина, висина, време чекања у реду у банци....
- дескриптивне статистике: мода,медијана,узорачка средина,распон узорка, интерквартилно растојање,узорачка дисперзија
- графичко представљање података: полигони фрквенција, кумулативни полигони фреквенција, хистограми (апсолутних, релативних фреквенција, густине), боксплот, барплот....
- идентификација аутлајера

Час 2

- емпиријска функција расподеле је дефинисана са

  $$F_n(x) = \frac{\sum_{i=1}^{n} I\{X_i \leq x\}}{n}$$

  Неке њене особине су:
  - $E(F_n(x)) = F(x)$, $D(F_n(x)) = \frac{F(x)(1-F(x))}{n}$
  - $nF_n(x)$ има $\mathcal{B}(n, F(x))$ расподелу
  - уз одговарајуће скалирање, $F(x)$ се може апроксимирати нормалном расподелом

теор. м.	узор. м.	теор. цент. м.	узор. цент. м.
EX	$\bar{X}_n$	$--$	$--$
EX^2	$\frac{\sum X_i^2}{n}$	DX	$\bar{S}_n^2$
EX^3	$\frac{\sum X_i^3}{n}$	$E(X-EX)^3$	$\frac{\sum (X_i-\bar{X}_n)^3}{n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
EX^k	$\frac{\sum X_i^k}{n}$	$E(X-EX)^k$	$\frac{\sum (X_i-\bar{X}_n)^k}{n}$

Табела 1: Теоријски и одговарајући узорачки моменти

  - неопадајућа је функција
  - функција за одређивање у R-у је ecdf(x)

- Оцењивање параметара EX са $\bar{X}_n$ и DX са поправљеном узорачком дисперзијом $\tilde{S}_n^2$;
- $E\bar{X}_n = EX_1$ и $E\tilde{S}_n^2 = DX_1$
- оцењивање параметара расподеле, опис статистичког модела и поставка проблема
- метод момената за оцењивање θ, при чему θ може бити вишедимензионалан параметар. Оцене непознатих параметара се добијају као решење система једначина који се добије кад се изједначе теоријски моменти са одговарајућим узорачким моментима. Илустровали смо метод на оцењивању параметара униформнр $\mathcal{U}[a,b]$ расподеле, нормалне $\mathcal{N}(m,\sigma^2)$, експоненцијалне $\mathcal{E}(\lambda)$, Пуасонове $\mathcal{P}(\lambda)$ и биномне $\mathcal{B}(N,p)$ расподеле.
- особине оцена: непристрасност, постојаност, асимптотска непристрасност
- поређење оцена: оцена $\hat{\theta}_n^{(1)}$ је боља од $\hat{\theta}_n^{(2)}$ (за параметар θ) у средњеквадратном смислу ако је $E(\hat{\theta}_n^{(1)} - \theta)^2 < E(\hat{\theta}_n^{(2)} - \theta)^2$

Час 3

Метод максималне веродостојности

Основни принцип овог метода је да је оцена непознатог параметра (који може бити вишедимензионални) вредност која максимизира функцију веродостојности.

У случају дискретног обележја функција веродостојности је

$$L(\theta) = P_\theta\{X_1 = x_1, ..., X_n = x_n\}.$$

У случају простог случајног узорка

$$L(\theta) = \prod_{i=1}^{n} P_\theta\{X_i = x_i\}.$$

У случају апсолутно непрекидног обележја функција веродостојности је

$$L(\theta) = f_\theta(X_1, ..., X_n).$$

У случају простог случајног узорка

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i).$$

Веома често је лакше маскимизирати неку монотону трансформацију функције веродостојности. Најчешће се максимизира $\log L(\theta)$.

Оцена добијена овомо методом не мора бити јединствена (радили смо пример $X \sim U[\theta-1, \theta+1]$ расподелу.)

Оцене добијене овом методом имају следеће лепо својство: Нека је g нека функција.Уколико је $\hat{\theta}_n$ оцена методом максималне веродостојности за θ онда је $g(\hat{\theta}_n)$ оцена методом максималне веродостојности за $g(\theta)$.

На часу смо илустровали метод прмерима обележја из Пуасонове $\mathcal{P}(\lambda)$, нормалне $\mathcal{N}(m, \sigma^2)$, биномне $\mathcal{B}(N, p)$, униформне $\mathcal{U}[0, \theta]$ и униформне $\mathcal{U}[\theta-1, \theta+1]$ расподеле и дискретне расподеле за коју важи $P\{X=-1\} = P\{X=1\} = \theta$ и $P\{X=0\} = 1-2\theta$.

Час 4

На часу смо објаснили основне принципе Бајесовске статистике и приказали сличности и разлике са класичним (фреквенионистичким) приступом. Главна разлика је што је у Бајесовском свету непознат параметар случајна величина са неком (априорном расподелом). Та расподела се мења након сазнања о реализованом узорку. Управо та (апостериорна) расподела је основ за Бајесовско оцењивање. Најчешће се за оцену парамтера узима математичко очекивање у односу на апостериорну расподелу.

На часу смо урадили примере за оцењивање параметра p код Биномне расподеле у случају да је априорна расподела униформна, затим

Бета и нека дискретна расподела. Поред тога, Бајесовско оцењивање смо илустровали на примеру оцене параметра средње вредности Пуасонове расподеле у случају да је априорна расподела експоненцијална расподела.