2. tétel

Nagy Dániel

2019. június 8.

Kivonat

Bootstrap módszerek. A maximum likelihood módszer. Hipotézis tesztelés. Extrém statisztikák. Post hoc analízis. Regresszió. Függetlenségvizsgálat. Egzakt tesztek.

1. Alapfogalmak

- Eseménytér (ez egy abstrakt fogalom): $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ pl. kockadobás esetén $\Omega = \{\omega_1 = \text{"1est dobok"}, \omega_2 = \text{"2est dobok"}, \omega_3 = \text{"párosat dobok"}...\}$
- Valószínűségi változó: $X : \Omega \to \mathbb{R}$ pl. kockadobás esetén $X(\omega_1) = 1, X(\omega_2) = 2, ...$
- Valószínűség: P egy mérték, amely Ω részhalmazaihoz számot rendel:
  - $P : \mathcal{P}(\Omega) \to \mathbb{R}$
  - $P(\Omega) = 1$ és $P(\varnothing) = 0$
  - $0 \leq P(A) \leq 1 \ \forall A \in \Omega$
  - Ha $A_1, A_2, ...$ diszjunkt részhalmazai Ω-nak, akkor
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$
- Hasznos összefüggések:
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - Két esemény független $\iff P(A \cap B) = P(A)P(B)$
  - $P(A|B) = \frac{P(A \cap B)}{P(B)}$
  - Teljes valószínűség: Ha $A_1, A_2, ...$ az Ω egy felosztása, akkor
$$P(B) = \sum_k P(B|A_k)P(A_k)$$
  - Bayes-tétel: Ha $A_1, A_2, ...$ az Ω egy felosztása, akkor
$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_j P(B|A_j)P(A_j)}$$
- Eloszlásfüggvény (CDF - cumulative distribution function):
$$F_X(x) = P(X < x) = P(\{\omega \in \Omega | X(\omega) < x\})$$
diszkrét esetben
$$F_X(x) = P(X = x) = P(\{\omega \in \Omega | X(\omega) = x\})$$
- Sűrűségfüggvény (PDF - Probability density function):
Ha az X változó eloszlásfüggvénye $F_X(x)$, akkor a sűrűségfüggvény definíciója
$$F_X(x) = \int_{-\infty}^{x} \rho_X(\xi)\mathrm{d}\xi \iff P(a \leq X(\omega) \leq b) = \int_a^b \rho_X(x)\mathrm{d}x$$
Megjegyzés: sűrűségfüggvénye csak folytonos eloszlású valószínűségi változónak van.

- Várható érték

$$\text{folytonos eset } E(X) = \langle X \rangle = \int\limits_{-\infty}^{\infty} x\rho(x)\mathrm{d}x$$

$$\text{diszkrét eset } E(X) = \langle X \rangle = \sum_k x_k P(X = x_k)$$

- Várható értékre vonatkozó azonosságok:
  - Ha $Y = g(X) \Rightarrow E(Y) = E(g(X)) = \int\limits_{-\infty}^{\infty} g(x)\rho(x)\mathrm{d}x$
  - $E\left(\sum\limits_k a_k X_k\right) = \sum\limits_k a_k E(X_k)$
  - Ha $X_1, X_2, ...$ független változók, akkor $E\left(\prod\limits_k X_k\right) = \prod\limits_k E(X_k)$

2. Bootstrap módszerek

2.1. Jackknife módszer

3. Maximum likelihood

4. Extrém statisztikák

5. Post-hoc analízis

6. Regresszió

7. Hipotézis tesztelés

8. z-teszt, t-test

8.1. Konfidenciaintervallumok

8.2. Függetlenségvizsgálat, χ^2-próba

Hivatkozások