LSM

Bojana Milošević

10/14/2019

Pretpostavke

$$Y = X\beta + \varepsilon,$$

1. $E(\varepsilon) = 0$
2. $Cov(\varepsilon) = \sigma^2 I$
3. X i ε su nezavisni slučajni vektori
4. $\varepsilon \sim \mathcal{N}_n(0, \sigma^2 I)$

MNK:

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

- $E(\hat{\beta}) = \beta$ i $Cov(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$

Osobine ocene uz pretpostavku 4.

- $Y \sim \mathcal{N}_n(X\beta, \sigma^2 I)$
- $\hat{\beta}$ je linearna transformacija Y pa ima $\mathcal{N}_n(\beta, \sigma^2(X^TX)^{-1})$
- Označimo $\Sigma = \sigma^2(X^TX)^{-1}$. Tada $\hat{\beta}_k \sim \mathcal{N}(\beta_k, \Sigma_{kk})$

Važne višedimenzione raspodele

- Slučajni vektor X ima n-dimenzionu normalnu raspodelu $N_n(\mu, \Sigma)$ ukoliko je njegova funckija gustine

$$f(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{(x-\mu)^T\Sigma^{-1}(x-\mu)}{2}}, \ \ x \in R^n$$

gde je Σ simetrična, pozitivno definitna kovarijaciona matrica a sa $|\Sigma|$ je označena njena determinanta.

Osobine normalne raspodele

- Ukoliko X ima $N(\mu, \Sigma)$ raspdelu onda $AX + b$ ima $N(A\mu + b, A\Sigma A^T)$
- X se može predstaviti u obliku $X = AZ + \mu$ gde je $AA^T = \Sigma$ a Z ima standardnu višedimenzionalnu normalnu raspodelu
- Ako slučajni vektor $Z = (X^T, Y^T)^T$ ima $N\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_y \end{pmatrix}\right)$ raspodelu onda su marginalne raspodele za X i Y redom $N(\mu_x, \Sigma_x)$ i $N(\mu_y, \Sigma_y)$, a uslovne raspodele $Y|x \sim N(\mu_y + \Sigma_{xy}^T \Sigma_{xx}^{-1}(x - \mu_x), \Sigma_y - \Sigma_{xy}^T \Sigma_x^{-1} \Sigma_{xy})$
- X ima višedimenzionu normalnu raspodelu akko za svaki vektor a (različit od nule) $a^T X$ ima jednodimenzionalnu normalnu raspodelu

χ^2 raspodela

Neka su $X_1, ..., X_k$ nezavisne slučajne veličine sa $N(\theta_1, 1)$,...,$N(\theta_k, 1)$ raspodelma, redom. Tada slučajna veličina

$$Y = \sum_{j=1}^{k} X_j^2$$

ima $\chi_k^2(\mu)$ raspodelu, gde je parametar položaja $\mu = \sum_{j=1}^{k} \theta_j^2$. Ukoliko je $\mu = 0$ parametar položaja ćemo izostaviti u notaciji.

Fišerova F raspodela

Neka $X \sim \chi^2_{n_1}(\mu)$ i $Y \sim \chi^2_{n_2}$ i nezavisne su. Tada

$$\frac{\frac{X}{n_1}}{\frac{Y}{n_2}}$$

ima Fišerovu $F_{n_1,n_2}(\mu)$ raspodelu.

Studentova raspodela

Neka X ima normalnu $N(\theta, 1)$ raspodelu i Y ima χ^2_m raspodelu i nezavisne su. Tada

$$\frac{X}{\sqrt{\frac{Y}{m}}}$$

ima Studentovu $t_m(\theta)$ raspodelu, gde je θ parametar položaja.

Osobine ocene parametara modela

- $$\frac{(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)}{\sigma^2} \sim \chi^2_{p+1}$$
- $$\frac{(n-p-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p-1}$$

  i nezavisno je sa $\hat{\beta}$
- $$\frac{(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)}{(p+1)\hat{\sigma}^2} \sim F_{p+1,n-p-1}.$$

Osobine ocene parametara modela

- $V_k = [(X^T X)^{-1}]_{kk}$
- Kako su $Y - \hat{Y}$ i $\hat{\beta}$ nekorelisani onda je i $\hat{\sigma}^2$ nekorelisano sa $\hat{\beta}$ a samim tim i nezavisno (zbog normalnosti)

$$D(\hat{\beta}_k) = \sigma^2 V_k$$

- $$\frac{\hat{\beta}_k - \beta_k}{\sigma\sqrt{V_k}} \sim N(0,1)$$

$$\frac{\hat{\beta}_k - \beta_k}{\hat{\sigma}\sqrt{V_k}} \sim t_{n-p-1}$$

Raspodela kvadratne forme - Kohranova teorema

Neka su $X_1, ..., X_n$ nezavisne $\mathcal{N}(0, \sigma^2)$ slučajne velićine i neka je

$$\sum_{i=1}^{n} X_i^2 = \sum_{j=1}^{k} Q_j,$$

gde je Q_j kvadratna forma definisana sa $Q_j = X^T A_j X$, za $j = 1, 2, ..., k$, pri čemu je $R(A_j) = r_j$. Tada je $\sum_{j=1}^{k} r_j = n$ ako i samo ako

1. $Q_1, ..., Q_k$ su nezavisne slučajne veličine i
2. Q_j/σ^2 ima $\chi^2_{r_j}$ raspodelu.

Lema koja je ključna u dokazu

Lemma

Neka su $x_1, ...x_n$ *realni brojevi. Pretpostavimo da se suma* $\sum_{i=1}^{n} x_i^2$ *mozhe predstaviti kao zbir* k *kvadratnih formi* $\sum_{j=1}^{k} Q_j$ *gde je* $Q_i = x^T A_i x$ *i* $R(A_i) = r_i$*. Ako je* $\sum_{i=1}^{k} r_i = n$ *tada postoji ortogonalna matrica* M*, takva da za* $x = My$ *važi*

$$\begin{aligned} Q_1 &= y_1^2 + \cdots y_{r_1}^2 \\ Q_2 &= y_{r_1+1}^2 + \cdots y_{r_1+r_2}^2 \\ &\cdots \\ Q_k &= y_{n-r_k+1}^2 + \cdots y_n^2 \end{aligned}$$

Primena teoreme

- Uzoračka sredina $\bar{X}$ i popravljena uzoračka disperzija S^2 su nezavine slučajne veličine, i $(n-1)S^2/\sigma^2$ ima χ^2_{n-1} raspodelu

$$(n-1)S^2 = \sum_{i=1}^{n} X_i^2 - n(\bar{X})^2$$

$$X^T X = X^T \left(I - \frac{J}{n}\right) X + X^T \frac{J}{n} X$$

Primena teoreme

- $$\frac{SSTO}{\sigma^2} \sim \chi^2_{n-1}$$
  $$\frac{SSE}{\sigma^2} \sim \chi^2_{n-p-1}$$
  $$\frac{SSR}{\sigma^2} \sim \chi^2_{p}$$

  i nezavisne su

Uopštenje Kohranove teoreme

Neka slučajni vektor X ima $\mathcal{N}(\theta, \sigma^2 I)$ raspodelu i neka je

$$\sum_{i=1}^{n} X_i^2 = \sum_{j=1}^{k} Q_j,$$

gde je Q_j kvadratna forma definisana sa $Q_j = X^T A_j X$, za $j = 1, 2, ..., k$, pri čemu je $R(A_j) = r_j$. Tada je $\sum_{j=1}^{k} r_j = n$ i $\sum_{j=1}^{k} \mu_j = \theta^T \theta$ ako i samo ako

1. $Q_1, ..., Q_k$ su nezavisne slučajne veličine i
2. Q_j/σ^2 ima $\chi^2_{r_j}(\mu_j)$ raspodelu, pri čemu je $\mu_j = \theta^T A_j \theta$

Važno tvrđenje o nezavisnosti kvadratnih formi

Neka slučajni vektor X ima $N(\theta, \sigma^2 I)$ raspodelu i neka je $Q_1 = X^T A_1 X$ i $Q_2 = X^T A_2 X$, gde su A_1 i A_2 dve simetrične matrice. Tada su Q_1 i Q_2 nezavisne ako i samo ako je $A_1 A_2 = 0$.