

Математичка статистика

Задатак 1. Нека је \mathbf{X}_n узорак из популације чије обележје X има нормалну $\mathcal{N}(m, \sigma^2)$ расподелу. Одредити расподелу случајне величине

- а) $\frac{\bar{X}_n - m}{\sigma} \sqrt{n}$;
 б) $\frac{n\bar{S}_*^2}{\sigma^2}$, где је $\bar{S}_*^2 = \frac{1}{n} \sum_{k=1}^n (X_k - m)^2$;
 в) $\frac{n\bar{S}_n^2}{\sigma^2}$;
 г) $\frac{\bar{X}_n - m}{\bar{S}_n} \sqrt{n-1}$.

Решење. а) Случајна величина \bar{X}_n има нормалну $\mathcal{N}(m, \frac{\sigma^2}{n})$ расподелу. Ово можемо лако показати коришћењем карактеристичних функција:

$$\varphi_{\bar{X}_n}(t) = \varphi_{\frac{1}{n} \sum_{k=1}^n X_k}(t) = \varphi_{\sum_{k=1}^n X_k}\left(\frac{t}{n}\right) = \prod_{k=1}^n \varphi_{X_k}\left(\frac{t}{n}\right) = \prod_{k=1}^n e^{im\frac{t}{n} - \frac{\sigma^2 t^2}{2n^2}} = e^{itmn - \frac{\sigma^2 t^2}{2n}}$$

Добијени израз представља карактеристичну функцију случајне величине која има $\mathcal{N}(m, \frac{\sigma^2}{n})$ расподелу. Према томе,

$$\frac{\bar{X}_n - m}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{X}_n - m}{\sigma} \sqrt{n} \in \mathcal{N}(0, 1).$$

б) Приметимо да

$$\frac{n\bar{S}_*^2}{\sigma^2} = n \frac{1}{n} \frac{1}{\sigma^2} \sum_{k=1}^n (X_k - m)^2 = \sum_{k=1}^n \left(\frac{X_k - m}{\sigma} \right)^2,$$

при чему $\frac{X_k - m}{\sigma} \in \mathcal{N}(0, 1)$. Дакле, $\frac{n\bar{S}_*^2}{\sigma^2}$ је збир квадрата n независних случајних величина са нормалном $\mathcal{N}(0, 1)$ расподелом. Самим тим, према дефиницији χ_n^2 расподеле, $\frac{n\bar{S}_*^2}{\sigma^2} \in \chi_n^2$.

в) Слично као у претходном задатку

$$\begin{aligned} \frac{n\bar{S}_n^2}{\sigma^2} &= n \frac{1}{n} \frac{1}{\sigma^2} \sum_{k=1}^n (X_k - \bar{X}_n)^2 = \frac{1}{\sigma^2} \sum_{k=1}^n (X_k - m + m - \bar{X}_n)^2 \\ &= \frac{1}{\sigma^2} \sum_{k=1}^n (X_k - m)^2 + \frac{2}{\sigma^2} \sum_{k=1}^n (X_k - m)(m - \bar{X}_n) + \frac{1}{\sigma^2} \sum_{k=1}^n (m - \bar{X}_n)^2 \\ &= \frac{n\bar{S}_*^2}{\sigma^2} - \frac{2}{\sigma^2} (n\bar{X}_n - mn)(\bar{X}_n - m) + \frac{1}{\sigma^2} n(\bar{X}_n - m)^2 \\ &= \frac{n\bar{S}_*^2}{\sigma^2} - \frac{1}{\sigma^2} n(\bar{X}_n - m)^2 = \frac{n\bar{S}_*^2}{\sigma^2} - \left(\frac{\bar{X}_n - m}{\sigma} \sqrt{n} \right)^2 \end{aligned}$$

Случајна величина $\frac{n\bar{S}_*^2}{\sigma^2} \in \chi_n^2$, а $\left(\frac{\bar{X}_n - m}{\sigma} \sqrt{n} \right)^2 \in \chi_1^2$ расподелу, као квадрат случајне величине која има $\mathcal{N}(0, 1)$ расподелу.

Да бисмо добили расподелу случајне величине $\frac{n\bar{S}_n^2}{\sigma^2}$ користићемо следећу теорему:

Теорема 1. Ако је X_1, \dots, X_n упрости случајан узорак из $\mathcal{N}(0, 1)$ расподеле, онда су \bar{X}_n и \bar{S}_n^2 независне случајне величине. *

Дакле, пошто је $\frac{n\bar{S}_n^2}{\sigma^2} = \frac{n\bar{S}_n^2}{\sigma^2} + \left(\frac{\bar{X}_n - m}{\sigma} \sqrt{n}\right)^2$, из наведене теореме следи да је карактеристична функција случајне величине $U_1 = \frac{n\bar{S}_n^2}{\sigma^2}$ једнака производу карактеристичних функција случајних величина $U_2 = \frac{n\bar{S}_n^2}{\sigma^2}$ и $U_3 = \left(\frac{\bar{X}_n - m}{\sigma} \sqrt{n}\right)^2$. Према томе, користећи задатак 1. д) и везу између χ^2 и гама расподеле, добија се

$$\begin{aligned} \varphi_{U_1}(t) &= \varphi_{U_2}(t)\varphi_{U_3}(t) \\ \frac{1}{(1-2it)^{\frac{n}{2}}} &= \varphi_{U_2}(t) \frac{1}{(1-2it)^{\frac{1}{2}}} \\ \implies \varphi_{U_2}(t) &= \frac{1}{(1-2it)^{\frac{n-1}{2}}}. \end{aligned}$$

Наведене функција је карактеристична функција случајне величине која има гама $\gamma\left(\frac{n-1}{2}, \frac{1}{2}\right)$ расподелу. Према томе, $\frac{n\bar{S}_n^2}{\sigma^2} \in \gamma\left(\frac{n-1}{2}, \frac{1}{2}\right)$, односно, користећи везу између гама расподеле и χ^2 расподеле, $\frac{n\bar{S}_n^2}{\sigma^2} \in \chi_{n-1}^2$.

г) Наведену статистику можемо написати на следећи начин

$$\frac{\bar{X}_n - m}{\bar{S}_n} \sqrt{n-1} = \frac{\bar{X}_n - m}{\frac{\bar{S}_n}{\sqrt{n-1}}} \frac{\sqrt{n}}{\sigma} = \frac{\bar{X}_n - m}{\frac{\bar{S}_n}{\sigma} \sqrt{\frac{1}{n-1}}}.$$

Из наведене теореме, закључујемо да су у бројиоцу и имениоцу независне случајне величине. Дакле, имамо $\frac{U_1}{\sqrt{U_2 \frac{1}{n-1}}}$, где $U_1 \in \mathcal{N}(0, 1)$, а $U_2 \in \chi_{n-1}^2$ и U_1 и U_2 су независне, па $\frac{U_1}{\sqrt{U_2 \frac{1}{n-1}}}$ има t_{n-1} расподелу. ✓

Задатак 2. Дати су независни узорци \mathbf{X}_n из популације чије обележје X има нормалну $\mathcal{N}(m_1, \sigma^2)$ расподелу и \mathbf{Y}_k из популације чије обележје Y има нормалну $\mathcal{N}(m_2, \sigma^2)$ расподелу. Одредити расподелу случајне величине

$$\frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sqrt{n\bar{S}_n^2(X) + k\bar{S}_k^2(Y)}} \sqrt{\frac{nk}{n+k}} (n+k-2).$$

Решење. Из претходног задатка следи да

$$\begin{aligned} \frac{\bar{X}_n - m_1}{\sigma} \sqrt{n} \in \mathcal{N}(0, 1) &\implies \frac{\bar{X}_n - m_1}{\sigma} \in \mathcal{N}\left(0, \frac{1}{n}\right) \\ \frac{\bar{Y}_k - m_2}{\sigma} \sqrt{k} \in \mathcal{N}(0, 1) &\implies \frac{\bar{Y}_k - m_2}{\sigma} \in \mathcal{N}\left(0, \frac{1}{k}\right) \end{aligned}$$

при чему су \bar{X}_n и \bar{Y}_k независне случајне величине, јер су узорци независни. Дакле

$$\frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sigma} \in \mathcal{N}\left(0, \frac{1}{n} + \frac{1}{k}\right).$$

Одавде,

$$U_1 = \frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}} = \frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sigma} \sqrt{\frac{nk}{n+k}} \in \mathcal{N}(0, 1)$$

Са друге стране, $\frac{n\bar{S}_n^2(X)}{\sigma^2} \in \chi_{n-1}^2$ и $\frac{k\bar{S}_k^2(Y)}{\sigma^2} \in \chi_{k-1}^2$ и независне су па

$$U_2 = \frac{n\bar{S}_n^2(X) + k\bar{S}_k^2(Y)}{\sigma^2} \in \chi_{n+k-2}^2.$$

Знамо да је \bar{X}_n независна од $\bar{S}_n^2(X)$ по теореме, а од $\bar{S}_k^2(Y)$ јер су узорци независни, па је независна и од њихове линеарне комбинације. Аналогно важи за \bar{Y}_k , па закључујемо да су U_1 и U_2 независне случајне величине. Дакле,

$$\begin{aligned} \frac{U_1}{\sqrt{U_2 \frac{1}{n+k-2}}} &= \frac{\frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sigma} \sqrt{\frac{nk}{n+k}}}{\sqrt{\frac{n\bar{S}_n^2(X) + k\bar{S}_k^2(Y)}{\sigma^2} \frac{1}{n+k-2}}} \\ &= \frac{(\bar{X}_n - m_1) - (\bar{Y}_k - m_2)}{\sqrt{n\bar{S}_n^2(X) + k\bar{S}_k^2(Y)}} \sqrt{\frac{nk(n+k-2)}{n+k}} \in t_{n+k-2}. \end{aligned}$$

✓

Задатак 3. Дати су независни узорци \mathbf{X}_n из популације чије обележје X има нормалну $\mathcal{N}(m_1, \sigma_1^2)$ расподелу и \mathbf{Y}_k из популације чије обележје Y има нормалну $\mathcal{N}(m_2, \sigma_2^2)$ расподелу. Одредити расподелу случајне величине $\frac{n(k-1)\sigma_2^2\bar{S}_n^2(X)}{k(n-1)\sigma_1^2\bar{S}_k^2(Y)}$.

Решење. Знамо да $U_1 = \frac{n\bar{S}_n^2(X)}{\sigma_1^2} \in \chi_{n-1}^2$ и $U_2 = \frac{k\bar{S}_k^2(Y)}{\sigma_2^2} \in \chi_{k-1}^2$ и да су случајне величине U_1 и U_2 независне јер су узорци независни. Стога:

$$\frac{\frac{U_1}{n-1}}{\frac{U_2}{k-1}} = \frac{\frac{n\bar{S}_n^2(X)}{\sigma_1^2(n-1)}}{\frac{k\bar{S}_k^2(Y)}{\sigma_2^2(k-1)}} = \frac{n(k-1)\sigma_2^2\bar{S}_n^2(X)}{k(n-1)\sigma_1^2\bar{S}_k^2(Y)} \in \mathcal{F}_{n-1, k-1}.$$

✓