Stat1: Lukács tétel és következménye

Lukács tétel. Legyen $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ fae. minta. Akkor

1. $\bar{X} \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$;
2. $nS_n^2/\sigma^2 \sim \chi^2(n-1)$, vagy ekvivalens módon $(n-1){S_n^*}^2/\sigma^2 \sim \chi^2(n-1)$;
3. $\bar{X}$ és S_n^2 függetlenek egymástól, vagy ekvivalens módon $\bar{X}$ és ${S_n^*}^2$ is függetlenek.

Bizonyítás. Az (1) állítás nyilvánvaló. A (2) és (3) állítás bizonyításához transzformáljuk X_i-ket egy olyan $n \times n$-es $\mathbf{V}$ ortogonális mátrixszal, amelynek utolsó sorában végig az $1/\sqrt{n}$ értékek állnak, a többi sora pedig tetszőleges (persze ortogonális az utolsó sorra, így az első $n-1$ sorban a sorösszegek 0-k lesznek). A transzformáció tehát:

$$Y_i = \sum_{k=1}^{n} v_{ik} X_k, \quad (i = 1, \ldots, n),$$

ahol v_{ik}-k a $\mathbf{V}$ mátrix elemei. Az így definiált $Y_1, Y_2, \ldots, Y_n$ val. változók nyilván normális eloszlásúak lesznek és függetlenek egymástól, ugyanis

$$\begin{aligned}
\mathrm{Cov}(Y_i, Y_j) &= \mathbb{E}\left[(Y_i - \mathbb{E}Y_i)(Y_j - \mathbb{E}Y_j)\right] = \\
&= \mathbb{E}\left[\sum_{k=1}^{n} v_{ik}(X_k - \mathbb{E}X_k) \sum_{l=1}^{n} v_{jl}(X_l - \mathbb{E}X_l)\right] = \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n} v_{ik}v_{jl}\mathbb{E}\left[(X_k - \mathbb{E}X_k)(X_l - \mathbb{E}X_l)\right] = \sum_{k=1}^{n}\sum_{l=1}^{n} v_{ik}v_{jl}\delta_{kl}\sigma^2 = \\
&= \sum_{k=1}^{n} v_{ik}v_{jk}\sigma^2 = \sigma^2 \sum_{k=1}^{n} v_{ik}v_{jk} = \sigma^2\delta_{ij} = \begin{cases} 0, & i \neq j \\ \sigma^2, & i = j. \end{cases}
\end{aligned}$$

Ebből látható, hogy Y_i-k korrelálatlanok, és a normalitás miatt függetlenek is, továbbá varianciájuk σ^2. Várható értékük:

$$\mathbb{E}Y_i = \sum_{j=1}^{n} v_{ij}\mathbb{E}X_j = \begin{cases} 0, & i = 1, \ldots, n-1 \\ \sqrt{n}\mu, & i = n. \end{cases}$$

mivel a $\mathbf{V}$ mátrix első $n-1$ sorában a sorösszeg 0. Ezért a független $Y_1, \ldots, Y_{n-1}$ változók eloszlása $\mathcal{N}(\mathbf{0}, \sigma^2)$ lesz, míg tőlük függetlenül $Y_n \sim \mathcal{N}(\sqrt{n}\mu, \sigma^2)$, sőt $Y_n = \sqrt{n}\bar{X}$. Ugyanakkor

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n}\Big(\sum_{k=1}^{n} v_{ik}X_k\Big)\Big(\sum_{l=1}^{n} v_{il}X_l\Big) = \sum_{k=1}^{n}\sum_{l=1}^{n} X_k X_l \sum_{i=1}^{n} v_{ik}v_{il} = \sum_{k=1}^{n}\sum_{l=1}^{n} X_k X_l \delta_{kl} = \sum_{k=1}^{n} X_k^2.$$

Végül a Steiner formulával:

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

Innen

$$\frac{nS_n^2}{\sigma^2} = \sum_{i=1}^{n-1} \left(\frac{Y_i}{\sigma} \right)^2,$$

ami $\chi^2(n-1)$-eloszlású (hisz $\frac{Y_i}{\sigma} \sim \mathcal{N}(0,1)$ fae, $i = 1, \ldots, n-1$) és független Y_n-től (igy $\bar{X}$-tól is).

Következmény: A χ^2 eo. paramétereit nézve,

$$\mathbb{E}\left((n-1) S_n^{*2} / \sigma^2 \right) = n - 1,$$

ezért

$$\mathbb{E}\left(S_n^{*2} \right) = \sigma^2.$$

Továbbá

$$\mathbb{D}^2\left((n-1) S_n^{*2} / \sigma^2 \right) = 2(n-1),$$

ezért

$$\mathbb{D}^2(S_n^{*2}) = \frac{2(n-1)}{(n-1)^2}\sigma^4 = \frac{2\sigma^4}{(n-1)} \to 0$$

as $n \to \infty$. Így S_n^{*2} négyzetes középben konzisztens becslése a valódi szórásnégyzetnek.