1. p : $(1-p)\cdot(1-p)\,p = (1-p)^2\cdot p = L(p)$

Ennek kell a maximuma p-ben $p\in[0,1]$

$L(p) = (p^2-2p+1)\,p = p^3-2p^2+p$

$L'(p) = 3p^2-4p+1 = (p-1)(3p-1)$

$L'(p)=0 \iff p=1$ vagy $\underbrace{p=\frac{1}{3}}$.

$\Downarrow$

$\hat{p}_{ML} = \frac{1}{3}$. //

2. $X\sim \mathrm{Exp}(\lambda)$ $\quad E(X) = 3 \Rightarrow \lambda = \frac{1}{3}$

$E(X) = \frac{1}{\lambda}$

$P(X\leq 2) = 1-e^{-\lambda\cdot 2} = 1-e^{-\frac{2}{3}}$.

$X, Y\sim \mathrm{Exp}(\frac{1}{3})$ ftl.

$P(X>1,\ Y>1) \underset{\text{ftl}}{=} P(X>1)\cdot P(Y>1) = e^{-\frac{1}{3}}\cdot e^{-\frac{1}{3}} = e^{-\frac{2}{3}}$.

3. 500 vendég $\quad S$: A menüt választók

$S\sim \mathrm{Binomiális}(500, \frac{5}{6})$

$S\leq 420$ és $500-S\leq 100$

$\Rightarrow \quad 400\leq S\leq 420$

$$P(400 \leq S \leq 420) = P\left(\frac{400-\frac{5}{6}\cdot 500}{\sqrt{\frac{5}{6}\cdot\frac{1}{6}\cdot 500}} \leq \ldots \leq \frac{420-}{\ldots}\right)$$

np

$np(1-p)$

$$\approx \Phi(\ldots) - \Phi(\ldots).$$

A sztochasztika alapjai

14. feladatsor: konfidenciaintervallum, próbák

1. Egy játék fizikai terhelhetőségére elvégzett tesztek kg-ban a következő eredményeket adták: 40, 45, 40, 42, 36. Tegyük fel, hogy a háttérváltozó normális eloszlást követ 2 szórással. (a) Adjunk 95%-os konfidenciaintervallumot a várható értékre! (b) Oldjuk meg a feladatot ismeretlen szórás esetén is!

(a) Láttuk

$$N(0,1) \sim \frac{\sqrt{n}}{\sigma}(\widehat{\mu} - \mu) \qquad \widehat{\mu} = \overline{X}_n = \frac{1}{n} \cdot \Sigma X_i$$

standard normális véletlen változó. Tehát

$$\begin{aligned} 0{,}95 = \mathbf{P}(|\widehat{\mu} - \mu| \le \delta) &= \mathbf{P}\left(\frac{\sqrt{n}}{\sigma}(\widehat{\mu} - \mu) \in [-\delta\sqrt{n}/\sigma, \delta\sqrt{n}/\sigma]\right) \\ &= \Phi(\delta\sqrt{n}/\sigma) - \Phi(-\delta\sqrt{n}/\sigma) \\ &= 2\Phi(\delta\sqrt{n}/\sigma) - 1. \end{aligned}$$

$$\Phi(-x) = 1 - \Phi(x)$$

Ez utóbbi kell 0,95 legyen, azaz

$$\Phi(\delta\sqrt{n}/\sigma) = 0,975.$$

A normális eloszlás táblázatából kapjuk, hogy

$$\frac{\delta\sqrt{n}}{\sigma} = \Phi^{-1}(0,975) = 1,96$$

vagyis

$$\delta = \frac{\sigma}{\sqrt{n}} 1,96 = 1,75.$$

Most használjuk csak a konkrét adatainkat. Eszerint $n = 5$ és $\overline{x}_n = 40,6$, ezért a 95%-os konfidenciaintervallum [40,6 − 1,75, 40,6 + 1,75] = [38,85, 42,35].

(b) Minden ugyanaz mint az előbb, csak nem ismerjük a σ szórást. Ezért becsüljük

$$\sigma \text{ helyett} \qquad Z_n = \left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2\right)^{1/2}.$$

A

$$\frac{1}{Z_n(X)\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu). \sim \text{Student}(n-1)$$

hányados Student-eloszlású $n-1$ szabadsági fokkal. Tehát

$$\begin{aligned} \mathbf{P}(|\widehat{\mu} - \mu| \le \delta) &= \mathbf{P}\left(\frac{\sqrt{n}}{Z_n}(\widehat{\mu} - \mu) \in [-\delta\sqrt{n}/Z_n, \delta\sqrt{n}/Z_n]\right) \\ &= \Phi_{n-1}(\delta\sqrt{n}/Z_n) - \Phi_{n-1}(-\delta\sqrt{n}/Z_n) \\ &= \underbrace{2\Phi_{n-1}(\delta\sqrt{n}/Z_n) - 1.} = 0{,}95 \end{aligned}$$

$n=5$

Itt most Φ_{n-1} az $n-1$ szabadsági fokú Student-eloszlás eloszlásfüggvénye. Ez utóbbi kell 0,95 legyen, azaz

$$\Phi_{n-1}(\delta\sqrt{n}/Z_n) = 0,975.$$

A táblázatából kapjuk, hogy

$$\frac{\delta\sqrt{n}}{Z_n} = \Phi_{n-1}^{-1}(0,975) = 2,776$$

vagyis

$$\delta = \frac{Z_n}{\sqrt{n}}\Phi_{n-1}^{-1}(0,975),$$

és a konfidenciaintervallum:

$$\left[\overline{X}_n - \delta, \overline{X}_n + \delta\right].$$

Beírva az adatainkat, $n = 5$, $\overline{X}_5 = 40,6$, $Z_5 = 3,29$, és $\Phi_4^{-1}(0,975) = 2,776$ (itt a 4. sort figyeltük), azaz $\delta = 3,29 \cdot 2,776/\sqrt{5} = 4,08$. Tehát a konfindenciaintervallum: [36,52, 44,68].

$P(|Z| > 1{,}96) = 0{,}05$

2. Egy véletlen változó értékeit megfigyelve a következő statisztikai mintát kapjuk: $6,5,\ 7,3,\ 5,4,\ 6,5,\ 2,1$. (a) Tegyük fel, hogy a háttéreloszlás normális, ismert $\sigma = 2$ szórással és ismeretlen μ várható értékkel. Teszteljük 5%-os szignifikanciaszinten azt a nullhipotézist, hogy az elméleti várható érték 8.

(b) Oldjuk meg a feladatot ismeretlen szórás esetén is.

A mintaátlag $\overline{X}_5 = 5,56$.

$$u = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma} = \underline{-2,72}.$$

← Ha igaz H_0, akkor ez standard normális

Választunk egy α szignifikanciaszintet, ez most 0,05. Keressük azt az u_α küszöbértéket, amelyre teljesül, hogy $\mathbf{P}(|Z| > u_\alpha) = \alpha$, ahol Z standard normális. A normális eloszlás szimmetriája miatt

$$\mathbf{P}(|Z| > u_\alpha) = \mathbf{P}(Z > u_\alpha) + \mathbf{P}(Z < -u_\alpha) = 2(1 - \Phi(u_\alpha)),$$

tehát $u_\alpha = \Phi^{-1}(1 - \alpha/2)$. Esetünkben $u_{0,05} = \Phi^{-1}(0,975) = 1,96$. Mivel $|u| = 2,72 > 1,96$, ezért ha igaz lenne a nullhipotézis, akkor a bekövetkezett esemény valószínűsége 0,05-nél kisebb, vagyis egy valószínűtlen esemény következett be. Ekkor 5%-os szignifikanciaszinten *elvetjük* a nullhipotézist. Azt mondjuk, hogy az elfogadási tartomány a $[-1,96,\ 1,,96]$ intervallum. Ha a próbastatisztika értéke ide esik, elfogadjuk a nullhipotézist, különben elvetjük.

(b) Most a próbastatisztika

$$t = \sqrt{n}\frac{\overline{X}_n - \mu_0}{Z_n} = \underline{-2,67}.$$

Itt persze Z_n a korrigált empirikus szórás. Ha a nullhipotézis igaz, akkor t véletlen változó Student-eloszlást követ $n-1$ szabadsági fokkal. Az ehhez tartozó kritikus érték az a t_α, melyre $\mathbf{P}(|T| > u_\alpha) = \alpha$, ahol T Student-eloszlású $n-1$ szabadsági fokkal. A Student-eloszlás szimmetriája miatt

$$\mathbf{P}(|T| > t_\alpha) = \mathbf{P}(T > t_\alpha) + \mathbf{P}(T < -t_\alpha) = 2(1 - \Phi_{n-1}(t_\alpha)),$$

tehát $t_\alpha = \Phi_{n-1}^{-1}(1 - \alpha/2)$. Esetünkben $t_{0,05} = \Phi_4^{-1}(0,975) = 2,776$. Most $|t| \le 2,776$, azaz 5%-os szignifikanciaszinten elfogadjuk a nullhipotézist.