2024. december 4-i gyakorlat

Két várható érték különbségének becslése

- Független minták esete: $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$

  – σ_Y^2 és σ_X^2 sokasági varianciák ismertek.
  Az $1-\alpha$ megbízhatósági szinthez tartozó konfidencia intervallum:

$$P\left(\overline{y}-\overline{x}-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_X^2}{n_X}} \;<\; \mu_Y-\mu_X \;<\; \overline{y}-\overline{x}+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_X^2}{n_X}}\right)=1-\alpha$$

  azaz

$$\mathrm{Int}_{1-\alpha}(\mu_Y-\mu_X)=\overline{y}-\overline{x}\pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_X^2}{n_X}}$$

  – σ_Y^2 és σ_X^2 sokasági varianciákat a mintákból kell becsülni, és feltételezhető, hogy $\sigma_Y^2=\sigma_X^2$.
  Legyen $\overline{d}=\overline{y}-\overline{x}$. Ekkor

$$s_{\overline{d}}=s_c\cdot\sqrt{\frac{1}{n_Y}+\frac{1}{n_X}}=\sqrt{\frac{(n_Y-1)s_Y^2+(n_X-1)s_X^2}{n_Y+n_X-2}}\cdot\sqrt{\frac{1}{n_Y}+\frac{1}{n_X}}$$

  Az $1-\alpha$ megbízhatósági szinthez tartozó konfidencia intervallum:

$$P\Big(\overline{d}-t_{1-\frac{\alpha}{2}}(\nu)\cdot s_{\overline{d}} \;<\; \mu_Y-\mu_X \;<\; \overline{d}+t_{1-\frac{\alpha}{2}}(\nu)\cdot s_{\overline{d}}\Big)=1-\alpha$$

  azaz

$$\mathrm{Int}_{1-\alpha}(\mu_Y-\mu_X)=\overline{d}\pm t_{1-\frac{\alpha}{2}}(\nu)\cdot s_{\overline{d}}$$

- Páros minták esete
  $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ és $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ között sztochasztikus kapcsolat van
  $n := n_X = n_Y$, továbbá σ_Y^2 és σ_X^2 ismeretlenek.
  Elemenként kivonjuk egymásból a két mintát: $d_i = y_i - x_i$, $i=1,\dots,n$. Ekkor

$$\overline{d}=\overline{y}-\overline{x}=\frac{\sum_{i=1}^n d_i}{n} \qquad s_d^2=\frac{\sum_{i=1}^n (d_i-\overline{d})^2}{n-1}$$

  Az $1-\alpha$ megbízhatósági szinthez tartozó konfidencia intervallum:

$$P\left(\overline{d}-t_{1-\frac{\alpha}{2}}(n-1)\frac{s_d}{\sqrt{n}} \;<\; \mu_Y-\mu_X \;<\; \overline{d}+t_{1-\frac{\alpha}{2}}(n-1)\frac{s_d}{\sqrt{n}}\right)=1-\alpha$$

  azaz

$$\mathrm{Int}_{1-\alpha}(\;\mu_Y-\mu_X\;)=\overline{d}\pm t_{1-\frac{\alpha}{2}}(n-1)\frac{s_d}{\sqrt{n}}$$

1. Egy autóklub megvizsgálta a Lada Borscs és a Skoda Sztrapacska gépkocsik fogyasztását. A vizsgálat néhány eredményét az alábbi táblázat tartalmazza. A fogyasztás szórásának gyári értékei mindkét típusú gépkocsi esetén 0,95 liter/100km. Az autók fogyasztásáról feltételezzük, hogy az normális eloszlást követ.

Group Statistics

	Autótípus	N	Mean	Std. Deviation	Std. Error Mean
fogyasztás(l/100km)	Lada Borscs	12	8,4806	1,0703	0,3090
	Skoda Sztrapacska	15	7,3799	0,8967	0,2315

(a) Adjon 98%-os konfidencia intervallumot a Lada Borscs átlagfogyasztására! Adjon becslést abban az esetben is, ha a szórás gyári értékét ismeretlennek tekintjük!

(b) Adjon 95%-os megbízhatóságú intervallumbecslést a Lada Borscs fogyasztásának szórására!

(c) Adjon 98%-os megbízhatóságú intervallumbecslést arra, hogy a Lada Borscs átlagosan mennyivel fogyaszt többet, mint a Skoda Sztrapacska! Adjon becslést abban az esetben is, ha a szórások gyári értékeit ismeretlennek tekintjük, de feltesszük, hogy azonosak!

(d) Egy másik autóklub adatai szerint 400-ból 12 Skoda Sztrapacska hűtővize elfolyik. Adjon 95%-os konfidencia intervallumot a hibás hűtőrendszerű Skoda Sztrapacskák arányára!

(e) Mekkora minta kellene ahhoz, hogy a hibás hűtőrendszerű Skoda Sztrapacskák arányára 90%-os biztonsággal adjunk az előzőnél kétszer pontosabb becslést (fele akkora hiba)?

2. A pincérek láthatatlan jövedelmének becslése céljából 10 kiválasztott pincér bevallott havi borravalójának ismeretében a vendégkör véleménye alapján megbecsülték a tényleges borravaló nagyságát is. A minta adatai a következők:

	Bevallott	Tényleges
Sorszám	borravaló, Ft/hó	
1.	4 000	9 000
2.	2 000	5 300
3.	3 500	6 000
4.	5 000	9 800
5.	1 800	4 300
6.	6 000	10 100
7.	2 800	5 900
8.	1 500	4 200
9.	3 900	9 400
10.	4 400	10 500

A tényleges és bevallott borravaló különbsége normális eloszlású változónak tekinthető. Becsülje meg 90% biztonsággal, hogy átlagosan mekkora összegű borravalót nem vallanak be a pincérek!

3. A Hörömpő Cirkusz (világszám!) bolha szekciójának vezető (és egyben egyedüli) artistája, Lajoska messze földön híres távolugró tudományáról. Minden egyes előadásnak van egy olyan pontja, amikor Lajoska helyből távolugrik. A mellékelt táblázat 8 délutáni és 5 délelőtti előadás ugrásának centiméterben mért adatait adja meg (az ugrások hosszát normális eloszlásúnak tekinthetjük).

Group Statistics

	előadás	N	Mean	Std. Deviation	Std. Error Mean
ugráshossz	délután	8	100,0000	15,58387	5,50973
	délelőtt	5	85,0000	11,18034	5,00000

(a) Adjon 95%-os megbízhatóságú intervallumbecslést a délelőtti ugrások szórására!

(b) Adjon 90%-os megbízhatóságú intervallumbecslést arra, hogy Lajoska átlagosan hány centiméterrel ugrik nagyobbat délután, mint délelőtt! Lajoska délelőtti és délutáni teljesítményét függetlennek tekintjük.

(c) Több éves közös munkájuk során Lajoska segédje minden előadáson feljegyezte a mért ugráshosszt és ebből meghatározta, hogy a délelőtti ugrások szórása 10 cm. Az adatok alapján adjon 95 %-os megbízhatóságú intervallumbecslést a délelőtti ugrások átlagos hosszára. Hány délelőtti ugrás adataira lenne szükség egy fele ilyen hosszú, de 98%-os megbízhatóságú konfidenciaintervallum előállításához?

(d) Lajoska fő attrakciója a háromszoros *Salto mortale*, azaz a védőháló nélküli halálugrás három átfordulással. Ez azonban nem mindig sikerül, ilyenkor Lajoska csak 2 bukfencet csinál a levegőben. A legutóbbi Kökörcsin utcai turnén Lajoska 50 ugrásból 30-at rontott. Adjon 90%-os megbízhatóságú intervallumbecslést a sikeres ugrások arányára!