Matematikai statisztika

7. előadás második része

2022. október 17.

Konfidenciaintervallum

- Olyan intervallum, mely legalább 1-α valószínűséggel tartalmazza a keresett paramétert:

$$P_\theta\left(T_1(X) < \theta < T_2(X)\right) \geq 1 - \alpha$$

Példa (normális eloszlás)

- A Gyorskenyér Kft automata kenyérsütő készülékei egyszerre 100 kenyeret sütnek ki. Ezek tömegei grammban mérve $N(m,10^2)$ eloszlással közelíthetőek, ahol m a kezelő beállításától függ. Egy ellenőrzésnél megmérték mind a 100 kenyér tömegét. Az átlag 990 g volt. Készítsünk 95%-os megbízhatóságú konfidencia intervallumot m-re!

Konfidencia intervallum normális eloszlás várható értékére (ismert szórás esetén)

$$\xi_1,\dots,\xi_n \sim N(m,\sigma^2),\ \sigma \text{ ismert},\ \Phi(u_y)=y \Rightarrow$$

$$P\left(\bar{\xi}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}<m<\bar{\xi}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha,$$

$$P\left(m>\bar{\xi}-u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha,$$

$$P\left(m<\bar{\xi}+u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

Konfidencia intervallum várható értékre (ismert szórás esetén)

$$\xi_1,\ldots,\xi_n,\ E\xi_i = m,\ D^2\xi_i = \sigma^2,\ \sigma \text{ ismert} \Rightarrow$$

$$P\left(\bar{\xi} - \sqrt{\frac{1}{\alpha}}\frac{\sigma}{\sqrt{n}} < m < \bar{\xi} + \sqrt{\frac{1}{\alpha}}\frac{\sigma}{\sqrt{n}}\right) \geq 1-\alpha.$$

α	$\boxed{u_{1-\alpha/2}}$	$\sqrt{\frac{1}{\alpha}}$
10%	1,64	3,16
5%	1,96	4,47
2,50%	2,24	6,32
1%	2,58	10,00

Konfidencia intervallum ”sok” megfigyelés esetén

$$\xi_1, \ldots, \xi_n,\ D^2\xi_i = \sigma^2 \text{ ismert} \Rightarrow$$

$$P\left(\bar{\xi} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < m < \bar{\xi} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \sim 1-\alpha.$$

Példák (milyen valószínűséggel születik fiúgyermek?)

- Svájcban 1871 és 1900 között a 2.644.757 megszületett gyermekből 1.359.671 fiú és 1.285.086 lány volt.
- Fiúk relatív gyakorisága így 0,5141.

$$p(1-p) \leq \frac{1}{4} \Rightarrow$$

$$P\left(\bar{\xi} - \frac{u}{2\sqrt{n}} < p < \bar{\xi} + \frac{u}{2\sqrt{n}}\right) \sim 2\Phi(u) - 1$$

Esetünkben 0,9973 valószínűséggel $0{,}5132 \leq p \leq 0{,}5150$

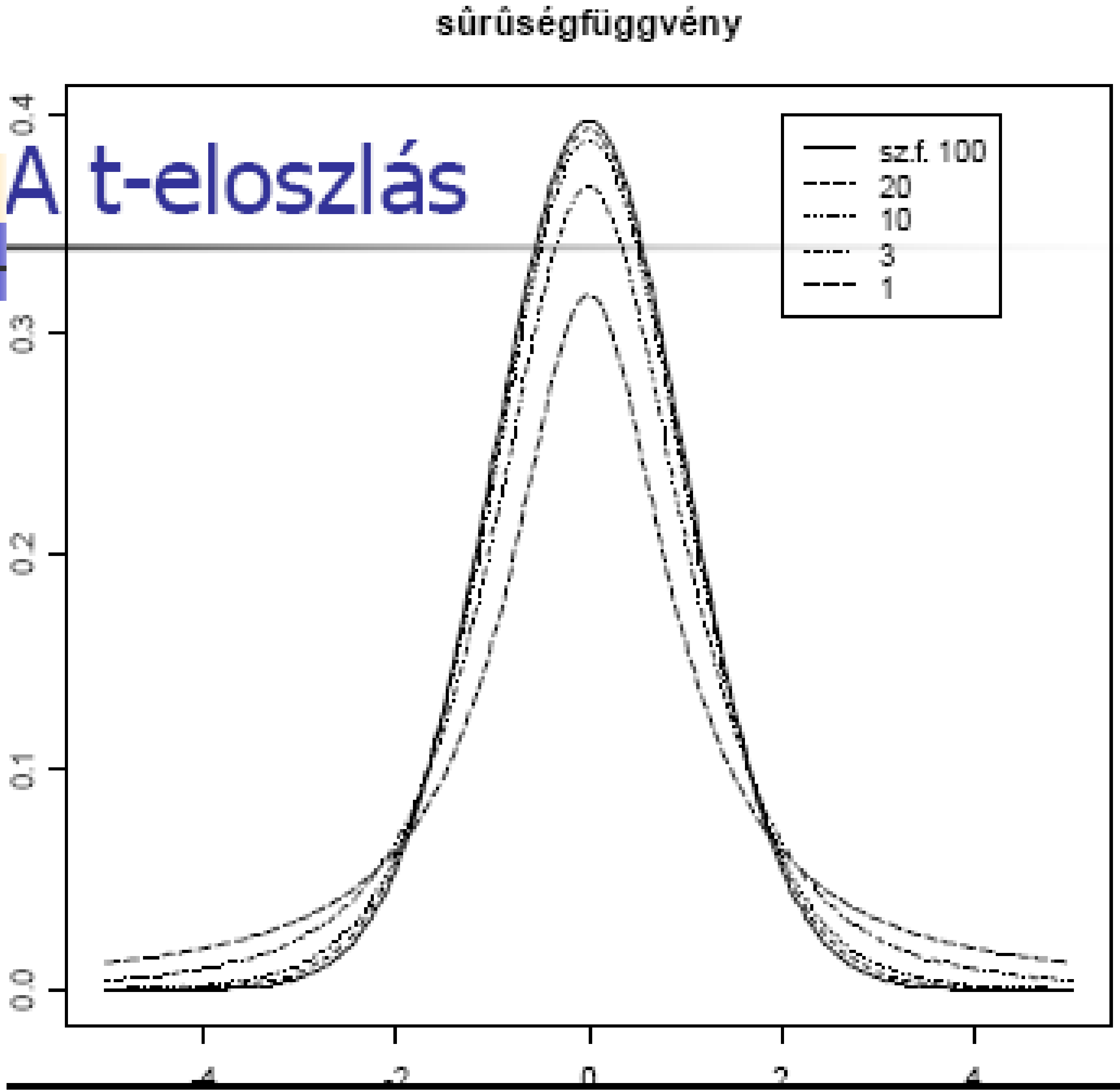
Konfidencia intervallum normális eloszlás várható értékére (ismeretlen szórás esetén)

- Ha a szórás nem ismert, becsüljük
- Tétel (biz. nélkül): normális eloszlású minta esetén a mintaátlag és a tapasztalati szórás független
- *n* szabadságfokú t (Student) eloszlás:

$$X_0, X_1, \ldots, X_n \ \boldsymbol{független}\ N(0,1)$$

$$\frac{X_0}{\sqrt{\left(X_1^2 + \ldots + X_n^2\right)/n}} \sim t_n$$

A t-eloszlás



Konfidencia intervallum normális eloszlás várható értékére (ismeretlen szórás esetén) (folyt.)

$$\xi_1,\dots,\xi_n \sim N(m,\sigma^2), \tilde{\sigma}^2 = \left(\left(\xi_1-\bar{\xi}\right)^2+\dots+\left(\xi_n-\bar{\xi}\right)^2\right)/(n-1) \Rightarrow$$

$$\frac{\sqrt{n}\left(\bar{\xi}-m\right)}{\sqrt{\tilde{\sigma}^2}} \sim t_{n-1}$$

$$P(t_{n-1} < t_{n-1,y}) = y$$

$$P\left(\bar{\xi}-t_{n-1,1-\alpha/2}\frac{\tilde{\sigma}}{\sqrt{n}} < m < \bar{\xi}+t_{n-1,1-\alpha/2}\frac{\tilde{\sigma}}{\sqrt{n}}\right)=1-\alpha,$$

$$P\left(m > \bar{\xi}-t_{n-1,1-\alpha}\frac{\tilde{\sigma}}{\sqrt{n}}\right)=1-\alpha,$$

$$P\left(m < \bar{\xi}+t_{n-1,1-\alpha}\frac{\tilde{\sigma}}{\sqrt{n}}\right)=1-\alpha$$

Példa (kenyér. folyt.)

- Tegyük fel most, hogy nem ismerjük Gyorskenyér Kft kenyereinek szórását. Az átlag 990 g volt.
- Ismert 10 szórásnál 991,6 g volt a 95%-os megbízhatóságú felső konfidencia határ.
- Amennyiben a korrigált tapasztalati szórás is 10, akkor ez a határ csak kis mértékben változik (991,8 g).
- Azonban 50-es korrigált tapasztalati szórásnál ez az érték 999 g-ra változik.

u és *t* együtthatók összehasonlítása

$$u_{1-5\%} = 1,64 \qquad (\Phi(1,64) = 95\%)$$

n	$t_{n-1,1-5\%}$
2	6,31
3	2,92
4	2,35
5	2,13
10	1,83
20	1,73
50	1,68
100	1,66
1000	1,65