Statisztikai próbák 1: ~~param~~ paraméteres próbák (u-próba, t-próba)

Legyen $X \sim N(m, \sigma^2)$ val. változó, σ ismert, m ismeretlen. 1/13

Legyen $\mu \in \mathbb{R}$ hipotetikus várható érték: valaki azt állítja, hogy $\boxed{m = \mu}$ ← ez a „null-hipotézis": $H_0: m = \mu$

~~Pl~~: Mintát veszünk X-ből: $x_1, x_2, \ldots, x_n$, és ez alapján szeretnénk eldönteni, hogy elhisszük-e a null-hipotézist, vagy sem.

Pl: a Piripócsi Tojás Manufaktúra azt állítja, hogy tyúkjaik átlagosan 45 gramm-os tojásokat tojnak. Lemértünk 10 tojást, és az jött ki (gramm-ban), hogy

43.6 ; 43.9 ; 42.7 ; 41.5 ; 42.8 ; 48.5 ; 44.6 ; 44.0 ; 46.0 ; 43.5

Elhisszük-e, hogy a várható érték tényleg 45?

1. lépés: számoljunk átlagot: esetünkben $\bar{x} \approx 44.11$.

Ha $\bar{x} = 45.0$ lenne, akkor H_0-t simán elhinnénk.

De $\bar{x} \neq 45.0$, hanem 0.89-cel kevesebb.

Kérdés: Ez gyanús-e, vagy sem? Elég nagy-e az eltérés ahhoz, hogy H_0-t ne higgyük el?

Avagy: Statisztikailag szignifikáns-e ez az eltérés?

Egyszerűsítő feltevés: Tudjuk, hogy a tojások méretének eloszlása $X \sim N(m, \sigma^2)$, ahol $\sigma = 2$ (grammban), m ismeretlen.

FONTOS: Bármennyi is az m, a fenti $x_1, x_2, \dots, x_{10}$ minta mindenképpen lehetséges, legfeljebb nagyon valószínűtlen. Pontosabban: az adott minta val.sége mindig nulla, ezért inkább:

Ha $m = \mu = 45$, akkor $\bar{x} = m - 0.89$, ami persze lehetséges. Kérdés: $\mathbb{P}(|\bar{x} - m| \geq 0.89)$ elég kicsi-e ahhoz, hogy ne higgyük el, hogy megesett?

Köv.: a „statisztikailag szignifikáns-e az eltérés" kérdés csak adott konfidencia-szinten értelmes: Ha H0-t elutasítom, mennyire akarok biztos lenni a dolgomban (hogy nem vádolom a céget alaptalanul)?

$\Rightarrow$ Feladat pontosabban: Döntsünk 95%-es konfidenciaszinten arról, hogy H0-t elfogadjuk-e!

1. megoldás:

- Konstruáljuk az $x_1, x_2, \ldots, x_n$ minta alapján 95%-os konfidencia-intervallumot m-re: $[\bar{x}-\Delta, \bar{x}+\Delta]$.

Az előző részből tudjuk, hogy ez azt jelenti, hogy

$$\mathbb{P}(m \in [\bar{x}-\Delta, \bar{x}+\Delta]) = 0.95$$

- Ha μ benne van, ~~akkor~~ vagyis $\mu \in [\bar{x}-\Delta, \bar{x}+\Delta]$, akkor H0-t $\boxed{\text{elfogadjuk}}$,

ha pedig nem, akkor H0-t $\boxed{\text{elutasítjuk}}$ ÉS bizonyítottnak tekintjük a $H1: \cancel{\mu}\, m \neq \mu$ ellenhipotézist.

95%-os szinten

FONTOS: ① Ha ~~μ~~ $m=\mu$, akkor $\mathbb{P}(\text{elutasítás}) = 0.05$, vagyis ha H0-t elutasítjuk, akkor 2 eset lehetséges:

1.) $m \neq \mu$ valóban, vagy
2.) Pechünk volt, és bekövetkezett egy kicsi (5%) val.ségű esemény.

② Ha $m \neq \mu = 45$, hanem $m = 44.999$, még mindig igaz, hogy $\mathbb{P}(\text{elutasítás})$ kicsi $\Longrightarrow$

Ha H0-t elfogadjuk, az semmit sem bizonyít, csak annyit jelent, hogy a rendelkezésre álló adatok nem mondanak ellent neki.

ε: Komoly súlya az elutasításnak van.

Avagy: Statisztikailag „bizonyítani" csak az ellenhipotézist van esélyünk.

Persze: Ha milliárdos kártérítési perről van szó, akkor a 95%-es konfidencia kevés lehet, de mondjuk 99.999999% már meggyőző:

$$\mathbb{P}(\text{hibásan megítélt kártérítés}) \leq 10^{-8}$$

Példánkban: a konfidencia-szint $95\% = 1-\varepsilon$, $\varepsilon = 0.05$,

$\sigma = 2$, $n = 10 \Rightarrow$

$\Downarrow$ előző rész

a konfidencia-intervallum sugara

$$\Delta = \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1-\frac{\varepsilon}{2}\right) = \frac{2}{\sqrt{10}} \Phi^{-1}(0.975) \approx \frac{2}{\sqrt{10}} \cdot 1.96 \approx 1.24$$

Mivel $\bar{x} \approx 44.11$, a konfidencia-intervallum

$$[\bar{x}-\Delta, \bar{x}+\Delta] = [42.87, 45.35] \ni 45$$

$\Rightarrow$ $\boxed{H_0\text{-t ELFOGADJUK}}$

2. megoldás: ekvivalens az előzővel, de $\boxed{\text{jobban általánosítható}}$:

1.) Az $x_1, x_2 \ldots x_n$ minta és a μ hipotetikus várható érték alapján számoljuk ki az

$$u := \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$ teszt-statisztikát

Tény (türi): HA H_0 IGAZ, akkor $u \sim N(0,1)$.

2.) Keressünk $K>0$ küszöbértéket úgy, hogy ha $u \sim N(0,1)$, akkor $\mathbb{P}(|u| \le K) = 1-\varepsilon$ legyen:

$$\mathbb{P}(|u| \le K) = \mathbb{P}(-K \le u \le K) = \phi(K) - \phi(-K) = 2\phi(K) - 1 = 1-\varepsilon$$

$$\Updownarrow$$

$$\boxed{K := \phi^{-1}\left(1-\frac{\varepsilon}{2}\right)}$$

3.) Ha $|u| \le K$, akkor H_0-t elfogadjuk,
ha nem, akkor elvetjük.

Példánkban: $\mu = 45$, $\bar{x} = 44.11$, $n = 10$, $\sigma = 2$, $\varepsilon = 0.05$

$$\Rightarrow \quad u = \frac{44.11-45}{2/\sqrt{10}} \approx -1.41 \;; \quad K = \phi^{-1}(0.975) \approx 1.96$$

Döntés: $|u| \le K \Rightarrow H_0$-t ELFOGADJUK.

Ez volt az EGYMINTÁS, KÉTOLDALI u-próba

u-próbák, t-próbák

8-féle statisztikai próbát tárgyalunk $8 = 2 \cdot 2 \cdot 2$.

Mindegyikben normális eloszlásból vesszünk mintát, és a ~~kape~~ H_0 nullhipotézis a várható értékre vonatkozik.

A próba lehet

a.) • 1-mintás: $x_1, x_2 \ldots x_n$ mintát hasonlítunk össze egy μ hipotetikus várható értékkel

• 2-mintás: $x_1, x_2 \ldots, x_n$ és $y_1, y_2 \ldots, y_m$ (esetleg $n \neq m$) mintákat hasonlítunk össze egymással.

b.) • Kétoldali: a null-hipotézis egyenlőség: $m = \mu$ (1 minta)

vagy $m_x = m_y$ (2 minta).

Egy ilyen null-hipotézis kétféleképpen tud megbukni:

$\bar{x}$ túl kicsi ~~(mondjuk μ-höz képest)~~, vagy

$\bar{x}$ túl nagy

[és pedig μ-höz képest (1 mintás eset)
vagy $\bar{y}$-hoz képest (2 mintás eset)]

• egyoldali: a null-hipotézis egyenlőtlenség ~~pl. $m \geq \mu$~~

pl $m \geq \mu$ vagy $m_x \geq m_y$. ~~vagy $m_x \geq m_y$~~

Ez csak 1-féleképpen tud megbukni: ha $\bar{x}$ nagyon nagy, az nem baj.

c) • u-próba, ha σ ismert

• t-próba, ha σ ismeretlen, így őt is a mintából kell becsülni.

KÖZÖS RECEPT:

1) A mintá(k)ból kiszámolunk egy teszt-statisztikát, ami olyan, hogy • Ha H0 igaz, akkor • közel van a 0-hoz nagy val.séggel

• és az eloszlása jól ismert

[Konkrétan: u-próba esetén normális eloszlás

t-próba esetén „Student-féle t-eloszlás"]

2.) A konfidencia szint alapján számolunk egy K küszöbértéket

3.) Ha a teszt-statisztika eltérése 0-tól K-nál nagyobb,

akkor H0-t elutasítjuk;

ha nem, akkor elfogadjuk

eltérés: • bármerre a kétoldali esetben

• a „rossz irányba" (amerre nem szabadna) az 1-oldali esetben.

Pl: Tudjuk, hogy $X \sim N(m_x, \sigma_x^2)$; $Y \sim N(m_y, \sigma_y^2)$

ahol $\sigma_x = 3$, $\sigma_y = 5$ ismert, viszont m_x, m_y ismeretlen.

Mintát vettünk X-ből és Y-ból, és az kaptuk, hogy

X : 12.7 ; 8.8 ; 11.9 ; 6.4 ; 8.4 ; 5.3 ; 7.9 ; 14.0

Y : 9.2 ; 11.8 ; 15.4 ; 17.9 ; 19.3 ; 14.2

Döntsünk 99%-es szinten arról a null-hipotézisről, hogy

a.) feladat: $m_x \le m_y$

b.) feladat: $m_y \le m_x$

Megoldás: Kétmintás egyoldali u-próbát végzünk:

↳ szórások ismertek

↳ H_0 egyenlőtlenség

↳ két mintál kell összehasonlítani.

Az adatokból $\bar{x} = 9.425$, $n_1 = 8$

$\bar{y} \approx 14.63$, $n_2 = 6$

a.) Mivel $\bar{x} < \bar{y}$, persze, hogy elfogadjuk: a mérés megerősíti a hipotézist: kár lenne tovább számolni.

b.) Mivel $\bar{x} < \bar{y}$, az adatok ellentmondani látszanak a hipotézisnek: kérdés, hogy szignifikánsan kisebb-e $\bar{x}$ mint $\bar{y}$, 99%-os szinten?

Ehhez: LÁSD A KÉPLETGYŰJTEMÉNYT ($\sigma_1 := \sigma_x$, $\sigma_2 := \sigma_y$)

- a teszt-statisztika

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

[Megj.: Tábel (könnyű): Ha H0 éppen hogy csak igaz, vagyis $m_x = m_y$, akkor $u \sim N(0,1)$]

esetünkben $u = \dfrac{9.425 - 14.63}{\sqrt{\frac{3^2}{8} + \frac{5^2}{6}}} \approx -2.26$

- A küszöbérték

$$K = u_\varepsilon = \phi^{-1}(1-\varepsilon) \underset{\substack{\varepsilon = 0.01 \\ 1-\varepsilon = 0.99}}{\overset{\text{esetünkben}}{=\!=}} \phi^{-1}(0.99) \approx 2.33$$

Döntés: $u < 0$, de nem eléggé: $u > -K$

$\Rightarrow$ H0-t ELFOGADJUK.

$\boxed{\text{t-próbák}}$ nagyon hasonlóak az u-próbákhoz. Annyit kell pluszban tudni, hogy

(1) A szórás ismeretlen $\Rightarrow$ a t teszt-statisztika képletében korrigált tapasztalati szórások szerepelnek

- S_n^* - gal jelölve az 1-mintás esetben
- S_1^*, S_2^* - gal jelölve a 2-mintás esetben.

WARNING: ezek mindig négyzeten vannak!

(2) a t teszt-statisztika eloszlása mindig $\boxed{\text{Student } t\text{-eloszlás}}$. Erről nekünk annyit kell tudni, hogy

a) több is van belőle: 1 szabadsági fokú; $df = 2$ —"—; 3 —"—; ⋮ — Hogy ~~melyi~~ mikor melyik kell az a képletgyűjteményből látszik.

szabadsági fokok száma

b.) Ez is táblázatba van szedve, mint a normális eloszlás,

sőt: nem kell visszafelé keresni, mert a táblázat eleve $(df, \varepsilon) \longmapsto$ küszöbérték alakban van szerkesztve — „degrees of freedom"

(3) A képletgyűjtemény-beli próbák csak arra az esetre érvényesek, ha $\sigma_x = \sigma_y$ (bár mindkettő ismeretlen).

Példa: Egy hadseregben a katonák testmagassága $\sim N(m, \sigma^2)$, de m és σ is ismeretlen. Megmértünk néhány katonát, és a következő magasságokat kaptuk (cm-ben):

165; 191; 176; 188; 168; 185; 161; 175; 182; 174

A Vezérkari Tábornok azt állítja, hogy $m \geq 186$ (cm).

Döntsünk 99%-es szinten erről a hipotézisről!

Megoldás: 1-mintás, 1-oldali t-próbát végzünk

→ mert σ ismeretlen

→ mert $\boxed{H_0 \text{ egyenlőtlenség}: m \geq \mu}$

→ mert 1 mintát hasonlítunk össze ~~a~~ a $\boxed{\mu = 186}$ hipotetikus várható értékkel

Adatok: a konfidenciaszintből $99\% = 1-\varepsilon \Rightarrow \underline{\varepsilon = 0.01}$

A mintából: $\bar{x} = 176.5$; $s_n^* \approx 9.99$ (zsebszámológéppel számolva)

$n = 10$ $\boxed{\bar{x} < \mu \text{ GYANÚS, számolni kell ☹}}$

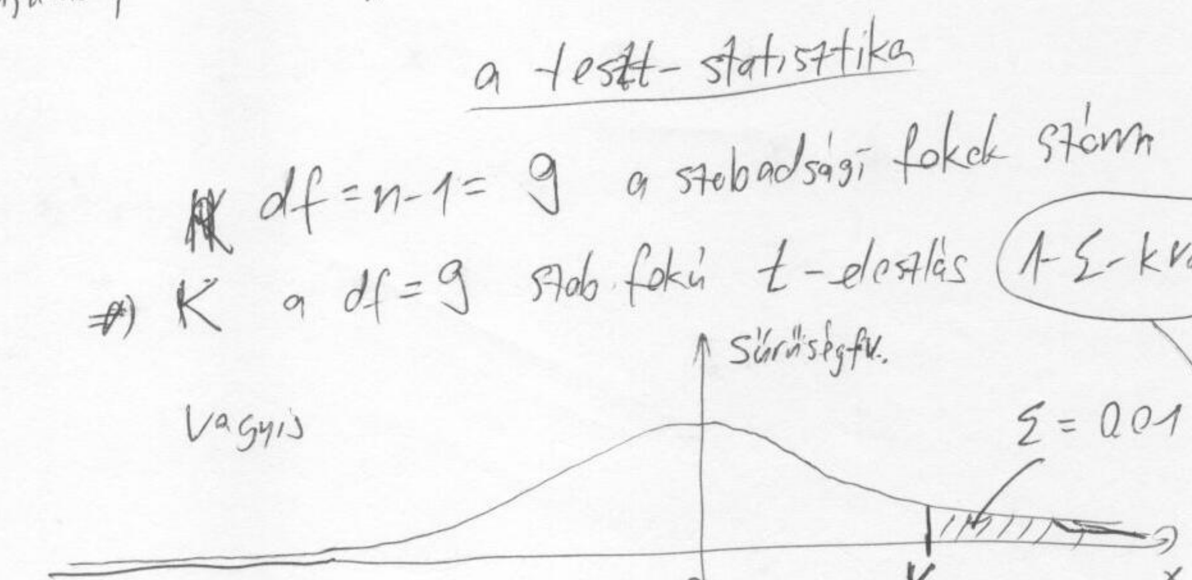
$$\xrightarrow[\text{gyűjteményből}]{\text{képlet-}} t = \frac{\bar{x}-\mu}{s_n^*}\sqrt{n} \approx \frac{176.5-186}{9.99}\sqrt{10} \approx -3.01$$

a teszt-statisztika

$df = n-1 = 9$ a szabadsági fokok száma

K a $df = 9$ szab. fokú t-eloszlás $1-\varepsilon$-kvantilise

Vagyis



ide esik a súly $(1-\varepsilon)$ hányada

Táblázatból a „$df = 9$" sorból keressük ki a „one-tail probability = 0.01" elemet:

12/13

PERCENTAGE POINTS OF THE T DISTRIBUTION

Tail Probabilities

DEGREES OF FREEDOM

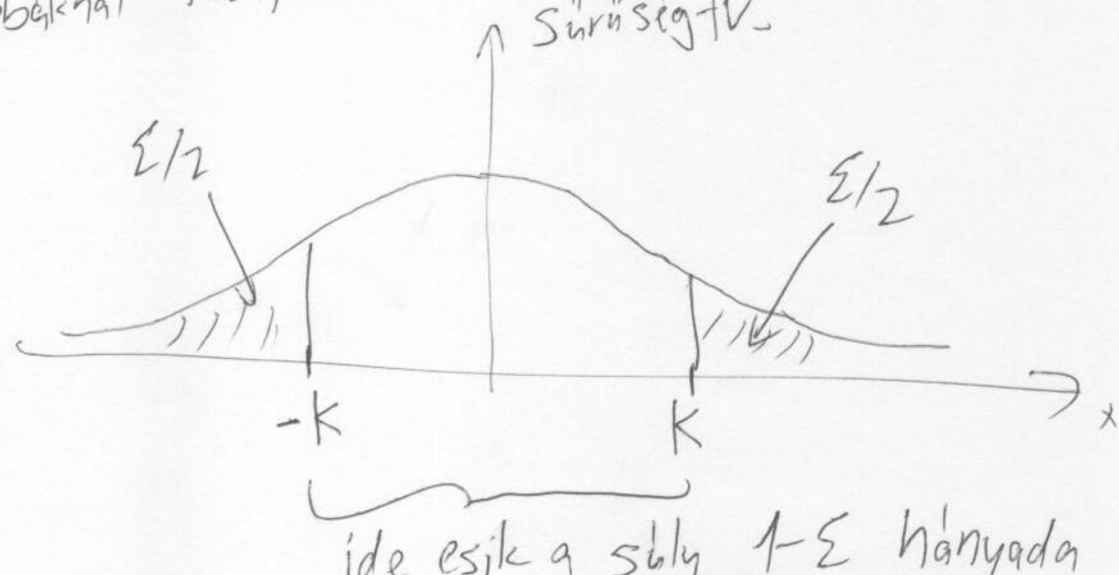
One Tail	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	
Two Tails	0.20	0.10	0.05	0.02	0.01	0.002	0.001	
1	3.078	6.314	12.71	31.82	63.66	318.3	637	1
2	1.886	2.920	4.303	6.965	9.925	22.330	31.6	2
3	1.638	2.353	3.182	4.541	5.841	10.210	12.92	3
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	4
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	5
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	6
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	8
9	1.383	1.833	2.262	$\boxed{2.821}$	3.250	4.297	4.781	9
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	10
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	11
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	12
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	13
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	14
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	15
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	16
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	17
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	18
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	19
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	20
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	21
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	22
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	23
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	24
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	25
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	26
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690	27
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	28
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659	29
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	30
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622	32
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601	34
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	36
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566	38
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551	40
42	1.302	1.682	2.018	2.418	2.698	3.296	3.538	42
44	1.301	1.680	2.015	2.414	2.692	3.286	3.526	44
46	1.300	1.679	2.013	2.410	2.687	3.277	3.515	46
48	1.299	1.677	2.011	2.407	2.682	3.269	3.505	48
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496	50
55	1.297	1.673	2.004	2.396	2.668	3.245	3.476	55
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460	60
65	1.295	1.669	1.997	2.385	2.654	3.220	3.447	65
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435	70
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416	80
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390	100
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357	150
200	1.286	1.653	1.972	2.345	2.601	3.131	3.340	200

K = 2.821

Döntés: $t < -K \Rightarrow H_0$-t elutasítjuk.

Helyette 99%-es ~~k~~ konfidencia-szinten bizonyítottuk, a ~~hogy m~~ $H_1: \underline{m < 186}$ ellenhipotézist.

[Megj: A táblázat „two tail probabilities" sora a kétoldali próbáknál kényelmes:



]