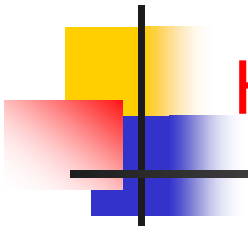


Statisztikai próbák

Gerzson Miklós

Pannon Egyetem, Villamosmérnöki és Információs
Rendszerek Tanszék

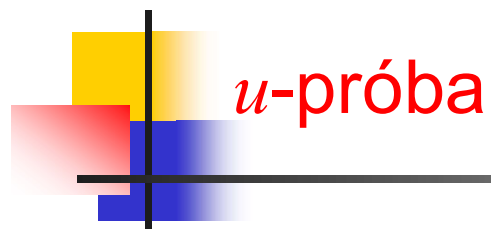
2019/2020 II. félév



Hipotézisvizsgálat

- cél: sokaság megismerése (paraméterek meghatározása)
- megoldás:
 - alapsokaságra feltevással élünk (például μ és/vagy σ értékére), és ezt statisztikai próbával ellenőrizzük,
 - ehhez mintát veszünk, és ennek segítségével próbáljuk igazolni a feltevést

- paraméteres próbák
 - u -próba
 - χ^2 -próba
 - F -próba
 - t -próba (egymintás, kétmintás, páros)
- nemparaméteres próbák
 - Mann-Whitney U-próba
 - Wilcoxon próba
 - Kruskal-Wallis H-próba
 - Friedman Fr-próba
 - Spearman-féle r_s -rangkorreláció



u -próba

- legyen adott egy normális eloszlású sokaság, melynek ismert a varianciája
- cél: a sokaság várható értékére vonatkozó feltevés, azaz a nullhipotézis igazolása
- a nullhipotézis:

$$H_0 : \mu = \mu_0$$

- lehetséges ellenhipotézisek:

$$H_1 : \mu \neq \mu_0$$

$$H_1 : \mu < \mu_0$$

$$H_1 : \mu > \mu_0$$

$$H_1 : \mu = \mu_1$$



u -próba

- legyen $x_1, \dots, x_n$ a sokaságból vett, n elemű minta
- meghatározás menete

1. a minta átlaga alapján meghatározzuk u_0 próbastatisztika értékét

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Az u_0 próbastatisztika kifejezése nem feltétlenül azonos az $N(0,1)$ eloszlású u standard normális eloszlású valószínűségi változóéval, mert μ helyett μ_0 szerepel benne. Csak akkor lesz az, ha $\mu = \mu_0$, azaz H_0 igaz.

átalakítva ezt a kifejezést:

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$$

az első tag definíció szerint u -eloszlású

a második tag az attól való eltérést fejezi ki

2. Az u -eloszlás táblázata alapján kiszámítjuk, hogy az u_0 próbastatisztika nagy $(1-\alpha)$ valószínűséggel milyen intervallumba esik.

Ha H_0 igaz (azaz a 2.tag zérus), akkor ez lesz az elfogadási tartomány.

Két oldali (ellen)hipotézis esetén

$$P\left(-u_{\alpha/2} < \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq u_{\alpha/2}\right) = P\left(\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \leq u_{\alpha/2}\right) = 1 - \alpha$$

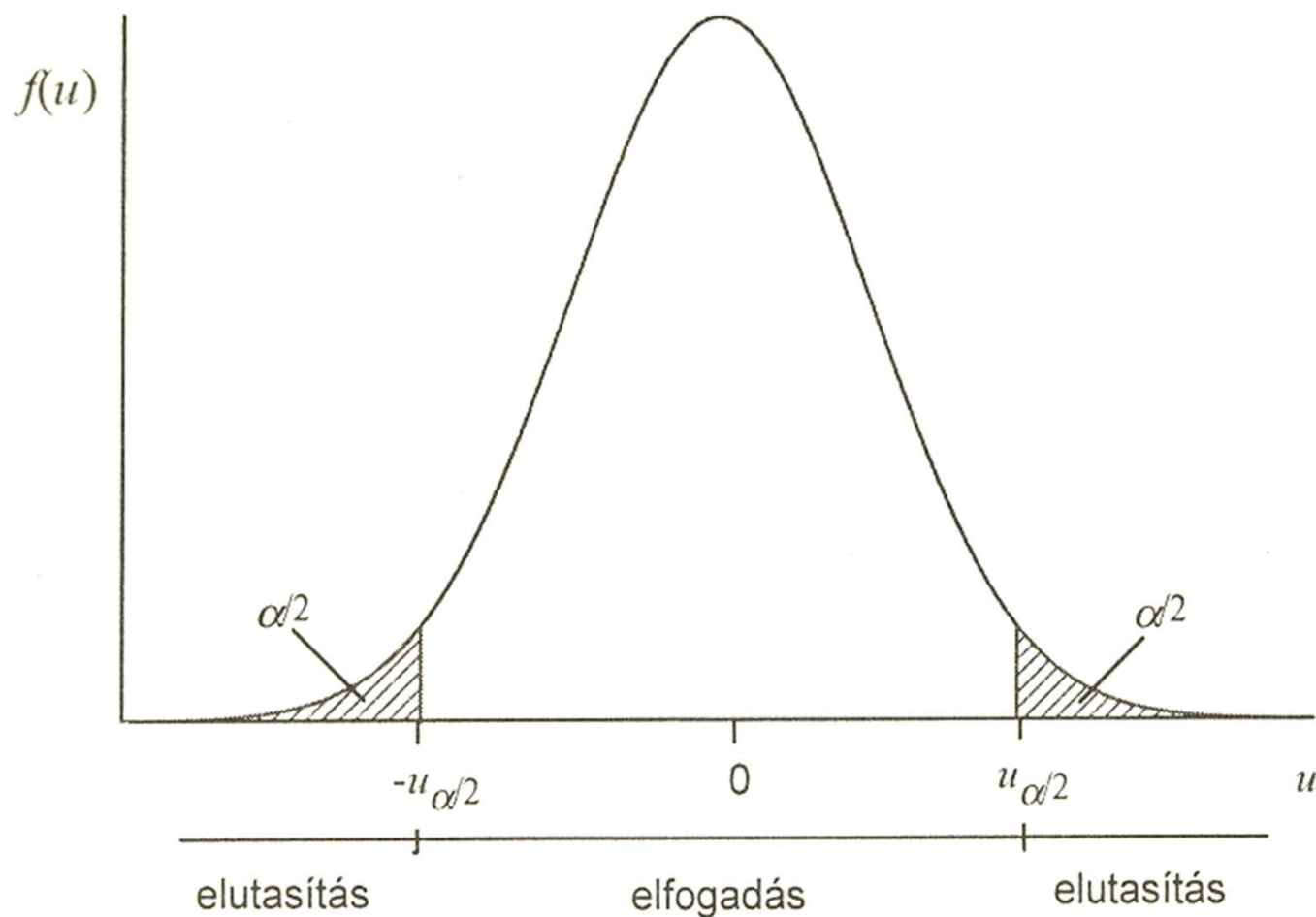
3. Megvizsgáljuk, hogy a számolt u_0 (a próbastatisztika értéke) az elfogadási tartományon belül van-e!

Ha H_0 igaz, akkor u_0 nagy (pl. $1-\alpha = 0.95$) valószínűséggel az $(-u_{\alpha/2}, u_{\alpha/2})$ elfogadási tartományban van, és csak kis ($\alpha = 0.05$) valószínűséggel esik azon kívülre, az elutasítási tartományba.

4. Ha az u_0 számított érték az $1-\alpha$ valószínűséghez tartozó elfogadási tartományba esik, akkor H_0 hipotézist elfogadjuk, míg ha a próbastatisztika értéke azon kívülre, az elutasítási tartományba esik, akkor elutasítjuk.

u -próba

- A nullhipotézis elfogadási tartománya



- az u -próba lényege: a próbastatisztika számlálójában szereplő minta átlag és a várható érték közti különbség és a nevezőben szereplő ingadozás összehasonlítása:

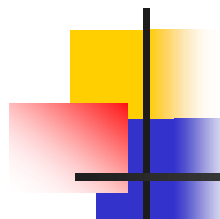
$$u_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- ha számlálóbeli eltérés csak véletlen ingadozással magyarázható, akkor az eredmény az elfogadási intervallumba esik, és H_0 hipotézist elfogadjuk



u -próba

- ha az eltérés ingadozása meghaladja az elfogadható mértéket, azaz az eredmény az elutasítási intervallumba esik, akkor elutasítjuk H_0 hipotézist
- α valószínűség a próba **szignifikancia szintje** ezt az eredménnyel együtt kell megadni, mert lehet, hogy a különbség 0.05-os szinten szignifikáns, az 0.01-os szinten nem
- példa

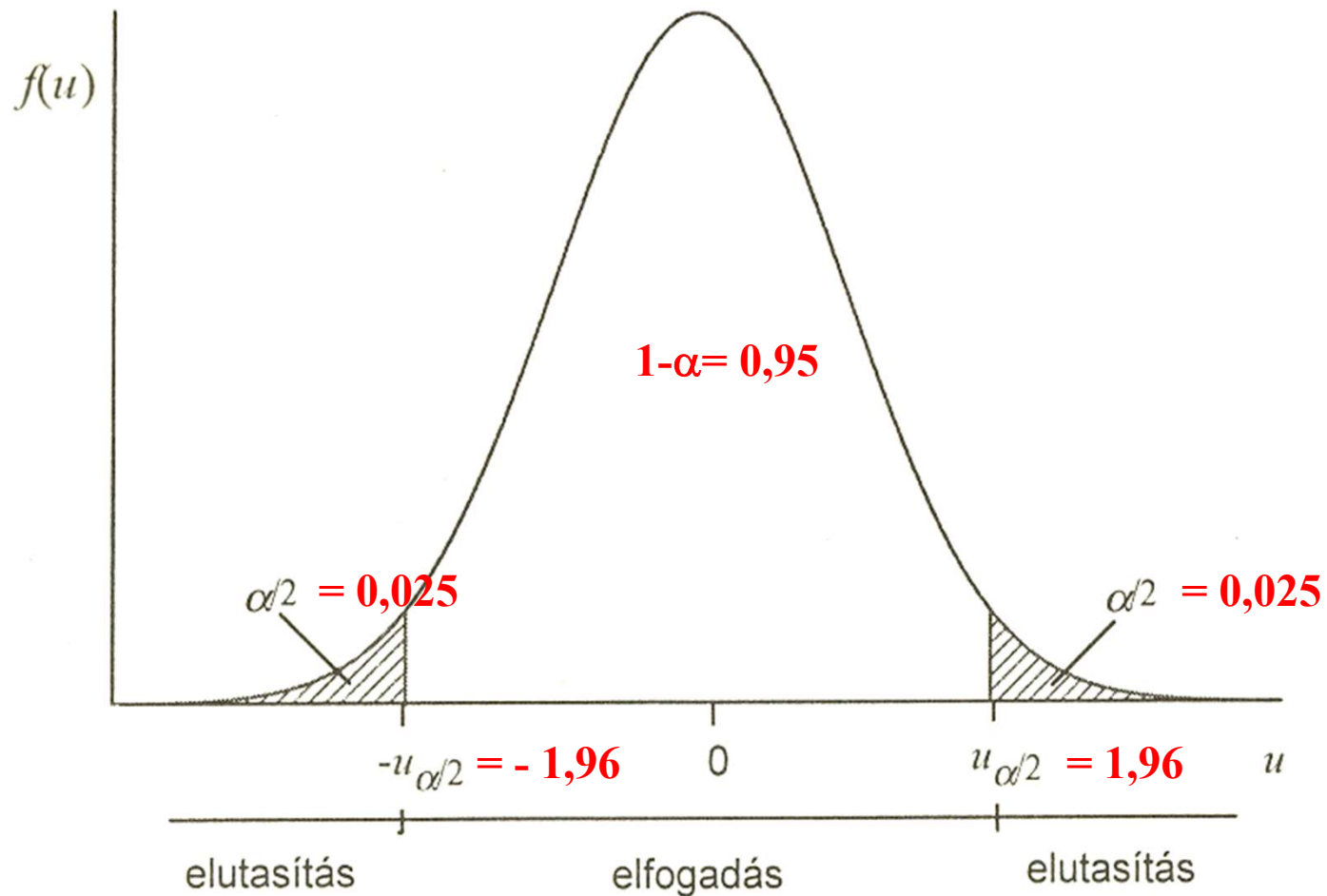


u-próba

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169

u -próba

- A nullhipotézis elfogadási tartománya a példában





u-próba

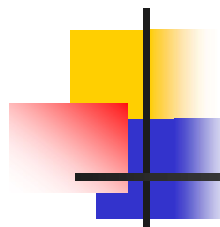
A példa kétoldalas határértékei

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58319	0.58709	0.59099	0.59487	0.59875	0.60261	0.60646	0.61030	0.61413
0.3	0.61794	0.62174	0.62553	0.62931	0.63307	0.63682	0.64056	0.64429	0.64799	0.65169
0.4	0.65538	0.65906	0.66273	0.66638	0.67002	0.67364	0.67725	0.68085	0.68443	0.68799
0.5	0.69154	0.69512	0.69868	0.70223	0.70576	0.70927	0.71277	0.71625	0.71971	0.72315
0.6	0.72657	0.73000	0.73341	0.73680	0.74017	0.74352	0.74685	0.75016	0.75345	0.75672
0.7	0.76000	0.76325	0.76648	0.76969	0.77288	0.77604	0.77918	0.78230	0.78540	0.78849
0.8	0.79156	0.79461	0.79764	0.80065	0.80364	0.80661	0.80956	0.81250	0.81542	0.81832
0.9	0.82120	0.82406	0.82690	0.82972	0.83251	0.83528	0.83803	0.84076	0.84347	0.84616
1.0	0.84883	0.85147	0.85409	0.85669	0.85927	0.86183	0.86437	0.86689	0.86939	0.87187
1.1	0.87433	0.87677	0.87919	0.88160	0.88398	0.88634	0.88868	0.89100	0.89330	0.89558
1.2	0.89784	0.90008	0.90230	0.90450	0.90668	0.90884	0.91098	0.91310	0.91520	0.91728
1.3	0.91934	0.92138	0.92340	0.92540	0.92738	0.92934	0.93128	0.93320	0.93510	0.93698
1.4	0.93884	0.94069	0.94252	0.94433	0.94612	0.94789	0.94964	0.95137	0.95308	0.95478
1.5	0.95646	0.95812	0.95976	0.96138	0.96298	0.96456	0.96612	0.96766	0.96918	0.97069
1.6	0.97228	0.97385	0.97540	0.97692	0.97842	0.97990	0.98136	0.98280	0.98422	0.98562
1.7	0.98701	0.98838	0.98973	0.99106	0.99236	0.99364	0.99490	0.99614	0.99736	0.99856
1.8	0.99975	0.99993	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
1.9	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.0	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.1	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.2	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.3	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.4	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.5	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.6	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.7	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.8	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.9	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
3.0	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999

$\alpha = 0,05$

$\alpha = 0,01$

$\alpha = 0,005$

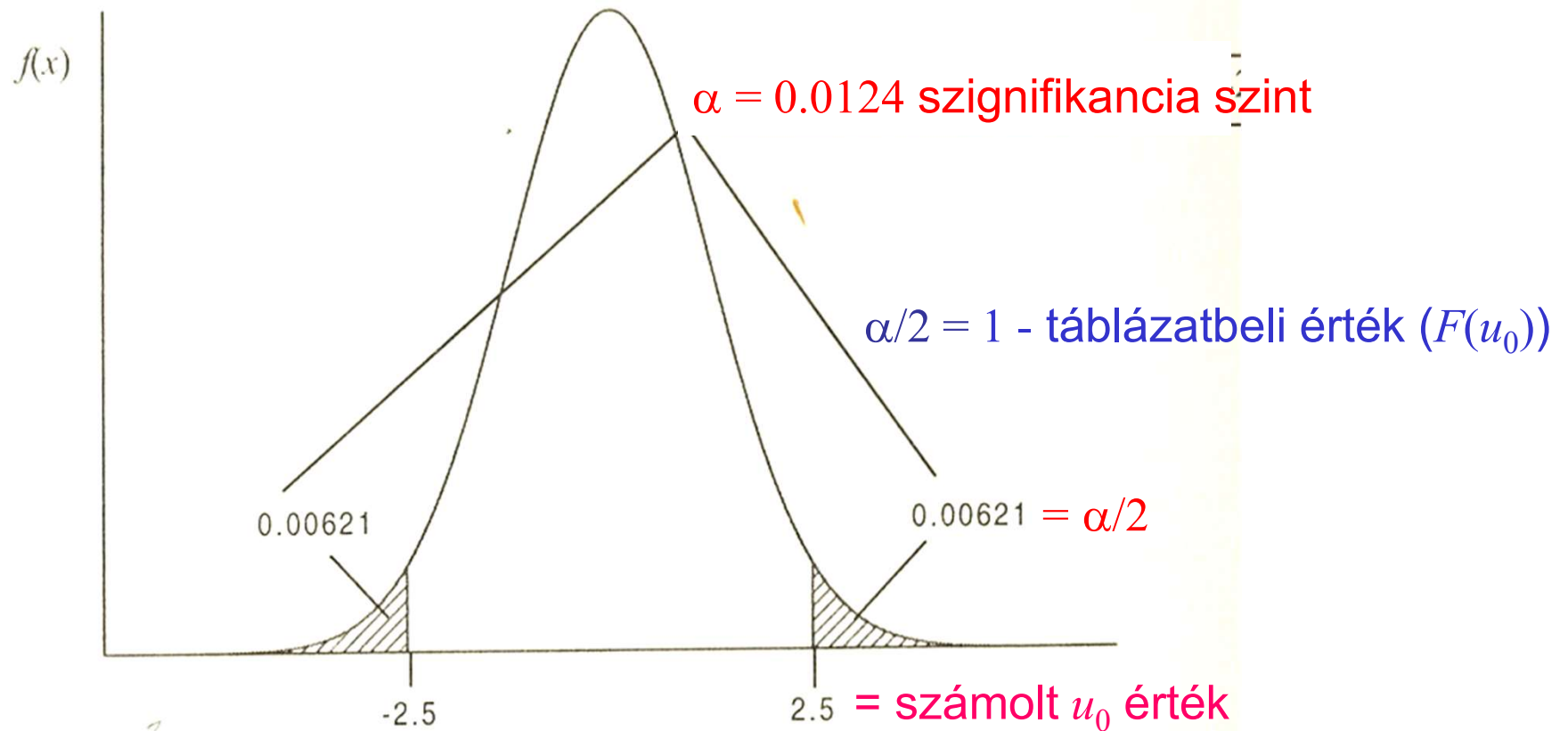


u-próba

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900

u -próba

- A példabeli értékhez tartozó elfogadási tartomány és szignifikancia szint – kétoldalas ellenhipotézis esetén



- Egy oldalas próba
- hipotézispárok: $H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$ (1. eset)
vagy $H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$ (2. eset)

- ekkor

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$$

kifejezés 2. tagja, ha H_0 igaz

- nulla vagy negatív (1. eset)
- nulla vagy pozitív (2. eset)

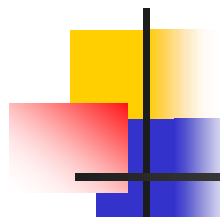
ha a H_0 nem igaz, azaz H_1 igaz, akkor a 2. tag pozitív(1. eset) / negatív (2. eset)

- a próba menete hasonló a kétoldalas próbához:
 - kiszámoljuk a példa adatai alapján u_0 -t
 - a táblázatból az α -hoz határértéket kikeressük,
 - H_0 elfogadási tartománya:
 - 1. eset

$$P\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq u_\alpha\right) = 1 - \alpha \text{ azaz } u_0 \in (-\infty, u_\alpha]$$

- 2. eset

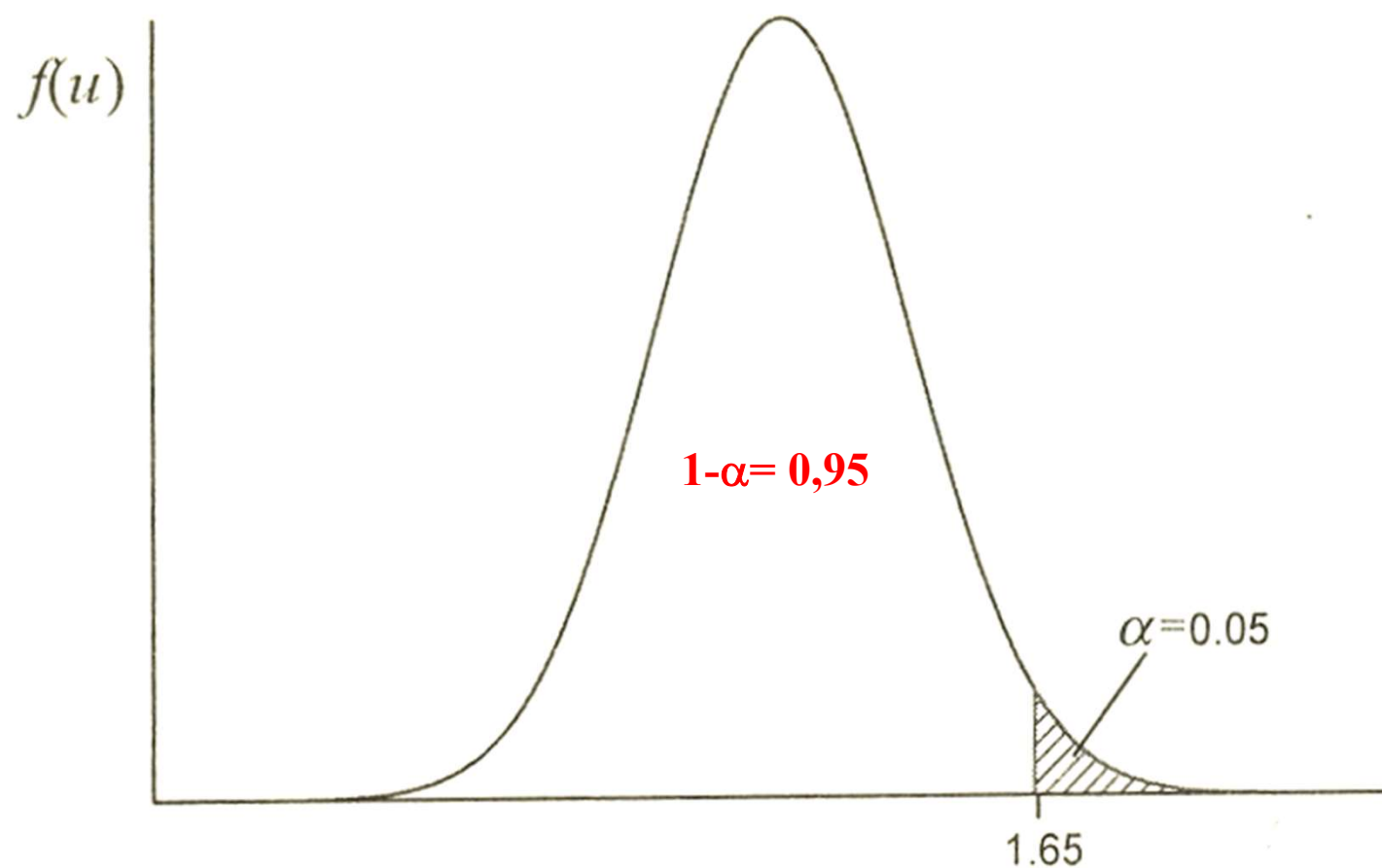
$$P\left(-u_\alpha \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha \text{ azaz } u_0 \in [-u_\alpha, \infty)$$



u -próba

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169

- A jobb oldali ellenhipotézis elfogadási tartománya





u-próba

A példa egyoldalas határértékei

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58319	0.58709	0.59099	0.59487	0.59875	0.60261	0.60646	0.61030	0.61413
0.3	0.61794	0.62174	0.62553	0.62931	0.63307	0.63682	0.64056	0.64427	0.64797	0.65166
0.4	0.65534	0.65901	0.66266	0.66629	0.66990	0.67350	0.67708	0.68064	0.68419	0.68772
0.5	0.69124	0.69474	0.69822	0.70168	0.70513	0.70856	0.71197	0.71536	0.71873	0.72208
0.6	0.72541	0.72872	0.73201	0.73528	0.73853	0.74176	0.74497	0.74816	0.75133	0.75448
0.7	0.75761	0.76073	0.76383	0.76690	0.76995	0.77298	0.77599	0.77898	0.78195	0.78490
0.8	0.78783	0.79074	0.79363	0.79650	0.79935	0.80218	0.80499	0.80778	0.81055	0.81330
0.9	0.81603	0.81874	0.82143	0.82410	0.82675	0.82938	0.83199	0.83458	0.83715	0.83970
1.0	0.84224	0.84476	0.84726	0.84974	0.85220	0.85464	0.85706	0.85946	0.86184	0.86420
1.1	0.86655	0.86888	0.87119	0.87348	0.87575	0.87799	0.88021	0.88241	0.88459	0.88675
1.2	0.88889	0.89099	0.89308	0.89515	0.89720	0.89923	0.90124	0.90323	0.90520	0.90715
1.3	0.90908	0.91100	0.91290	0.91478	0.91664	0.91848	0.92030	0.92210	0.92388	0.92564
1.4	0.92738	0.92910	0.93080	0.93248	0.93414	0.93578	0.93740	0.93899	0.94057	0.94213
1.5	0.94367	0.94519	0.94669	0.94817	0.94963	0.95107	0.95249	0.95389	0.95527	0.95663
1.6	0.95798	0.95931	0.96062	0.96191	0.96318	0.96443	0.96566	0.96687	0.96807	0.96925
1.7	0.97041	0.97156	0.97269	0.97380	0.97489	0.97596	0.97699	0.97801	0.97901	0.98000
1.8	0.98097	0.98192	0.98285	0.98376	0.98465	0.98551	0.98636	0.98719	0.98800	0.98879
1.9	0.98956	0.99032	0.99106	0.99178	0.99248	0.99316	0.99382	0.99446	0.99509	0.99569
2.0	0.99628	0.99685	0.99740	0.99793	0.99844	0.99893	0.99940	0.99985	0.99999	0.99999
2.1	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.2	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.3	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.4	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.5	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.6	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.7	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.8	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
2.9	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
3.0	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999

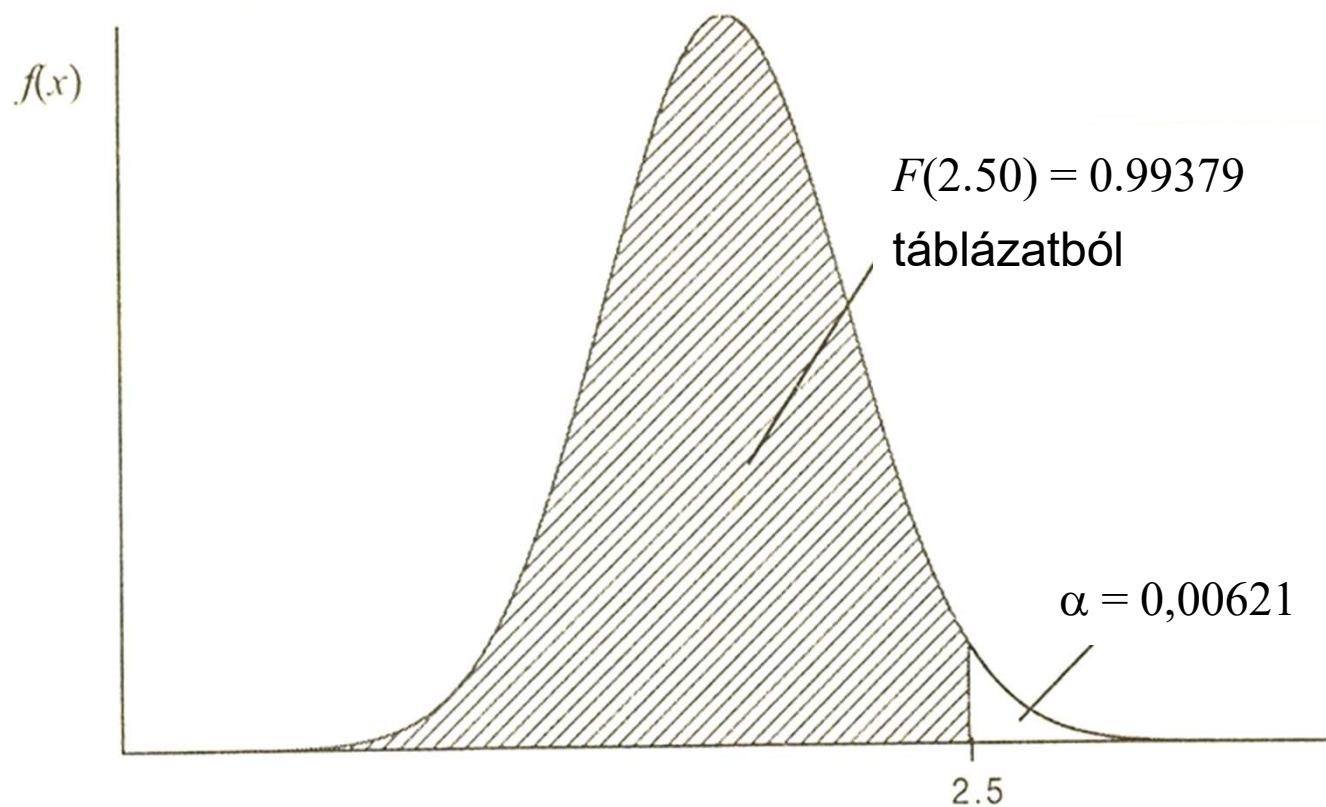
$\alpha = 0,05$

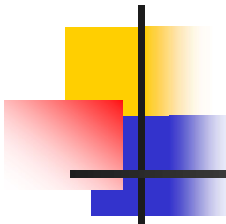
$\alpha = 0,01$

$\alpha = 0,005$

$\alpha = 0,001$

- A példabeli értékhez tartozó elfogadási tartomány és szignifikancia szint – jobb oldali ellenhipotézis esetén

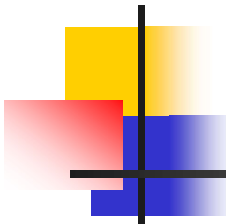




Első- és másodfajú hiba

- minden statisztikai próbánál kétféle hiba követhető el:
 - elvetjük H_0 -t, pedig igaz *elsőfajú hiba*
 - elfogadjuk H_0 -t, pedig nem igaz *másodfajú hiba*

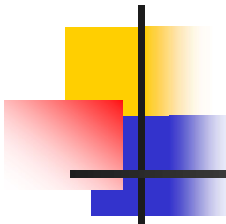
	Döntés: H_0 -t	
	elfogadjuk	elutasítjuk
H_0		
igaz	helyes	elsőfajú hiba
nem igaz	másodfajú hiba	helyes



Első- és másodfajú hiba

- annak valószínűsége, hogy elsőfajú hibát követünk el: α
 - mert α valószínűsége annak, hogy H_0 igazsága mellett a próbastatisztika az elutasítási tartományba esik: $P\{|u_0| > \mu_{\alpha/2} \mid H_0\} = \alpha$
- a másodfajú hiba valószínűségét egy olyan H_1 alternatív hipotézisre lehet megadni, amely a H_0 hipotézistől a feladat megszabta műszaki szempontból már észrevehetően különböző állítást tartalmaz:

$$H_1: \mu = \mu_1$$

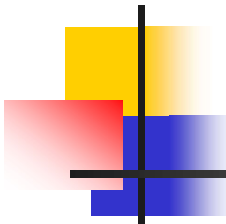


Első- és másodfajú hiba

- Ha H_0 helyett H_1 igaz, akkor az u_0 próbastatisztika sűrűségfüggvénye az u -eloszláshoz képest a $\mu_1 - \mu_0$ különbség nagyságának függvényében eltolódik:

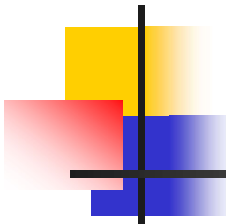
$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu_1}{\sigma/\sqrt{n}} + \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}$$

- ezt sűrűségfüggvények segítségével ábrázolva



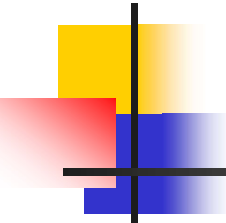
Első- és másodfajú hiba

- másodfajú hiba:
 - H_0 -t elfogadom, pedig nem igaz,
 - vagyis H_1 igaz, de nem fogadom el
- β elkövetésének valószínűsége annál kisebb, minél távolabb van μ_0 μ_1 -től
 - azaz minél nagyobb az eltérés, annál kisebb a valószínűsége, hogy észrevétlen maradjon



Első- és másodfajú hiba

- β nagysága függ
 - a varianciától – egyenes arányban
 - mintaszámtól – fordított arányban
 - α csökkentésével nő!
- másképpen: minél kisebb a minta információtartalma:
 - nagy szórás
 - kicsi mintaszámannál nagyobb lesz β
- példa



Első- és másodfajú hiba

- másodfajú hiba β valószínűsége tehát csak egy adott $H_1: \mu = \mu_1$ ellenhipotézishez számítható ki és β lesz a valószínűsége annak, hogy a $\mu_1 - \mu_0$ különbséget nem vesszük észre:

$$\beta = P(-u_{\alpha/2} < u_0 < u_{\alpha/2} | H_1)$$

$$\beta = P\left(-u_{\alpha/2} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} < u_{\alpha/2} \mid \mu = \mu_1\right)$$

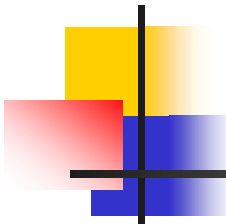
$$\beta = P\left(-u_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} < u < u_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\right)$$

az $N(0,1)$ normális eloszlás $F(u)$ eloszlásfüggvénye

u	0	1	2	3	4	5/	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327

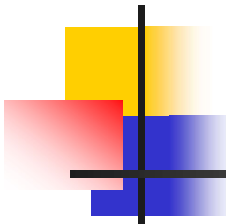
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99157
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900

$\alpha = 0,01$



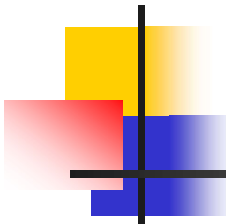
Első- és másodfajú hiba

- Összefoglalva:
 - ha u_0 kívül van az α -hoz meghatározott elfogadási tartományon, akkor H_0 -t elutasítjuk
 - α annak valószínűsége, hogy H_0 -t elutasítjuk, pedig igaz
 - α -t csökkentve ez a kockázat is csökkenthető
 - ha H_0 -t elfogadjuk, az nem jelenti azt, hogy igaz is, inkább azt, hogy *nincs elégséges információnk az elutasításhoz!*
 - ennek kockázatát adja meg a másodfajú hiba



Első- és másodfajú hiba

- ha nem ilyen típusú ellenhipotézisről van szó, hanem pl.: $H_1: \mu > \mu_0$ azaz alternatívák sorozatáról, akkor a másodfajú hiba β valószínűsége függvény lesz, melynek maximuma a $\mu_1 = \mu_0$ helyen van, és ez a próba **erőfüggvénye**
- β értékét szokás a $\mu_1 - \mu_0$ függvényében ábrázolni, ez a próba működési jelleggörbéje (OC-görbe)
- példa

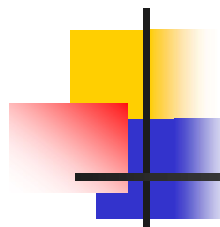


Első- és másodfajú hiba

- jelölje Δ azt a különbséget, amely a feladat műszaki szempontjából lényeges, vagyis kimutatandó

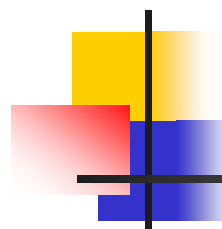
$$\Delta = \mu_1 - \mu_0$$

- határozzuk meg ehhez β -t, vagyis annak valószínűségét, hogy Δ nagyságú eltérést nem veszünk észre
- adott α mellett β az alternatív hipotézistől és $\sigma/\sqrt{n}$ -től függ
- ha adott α , β , Δ és σ , akkor meghatározható a mérések száma



u-próba

u	0	1	2	3	4	5 /	6	7	8	9
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	$\beta = 0,02$	0.98077	0.98124
2.1	0.98169	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537
2.2	0.98574	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870
2.3	0.98899	0.98928	0.98956	0.98983	0.99010	$\alpha = 0,01$	0.99036	0.99061	0.99086	0.99111
2.4	0.99134	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343
2.5	0.99361	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506
2.6	0.99520	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632
2.7	0.99643	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728
2.8	0.99736	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801
2.9	0.99807	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856
3.0	0.99861	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896



χ^2 -próba

- χ^2 -próba a variancia vizsgálatára
- cél: *normális eloszlású* sokaság ismeretlen σ^2 varianciájára vonatkozó H_0 hipotézis ellenőrzése
- megoldás:
 - vegyünk n számú mintát
 - határozzuk meg a minta tapasztalati szórásnégyzetét, s^2 -t
 - ez alapján ellenőrizzük, hogy $\sigma^2 = \sigma_0^2$ azaz

$$H_0: \sigma^2 = \sigma_0^2$$

- legyen az ellenhipotézis

$$H_1: \sigma^2 > \sigma_0^2$$

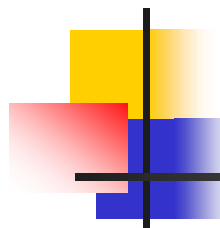
- de ekkor

$$H_0: \sigma^2 \leq \sigma_0^2$$

- (Ha $H_1: \sigma^2 < \sigma_0^2$, akkor $H_0: \sigma^2 \geq \sigma_0^2$!)

- Ha H_0 igaz, akkor a következő kifejezés χ^2 -eloszlású $\nu = n-1$ szabadsági fokkal:

$$\chi_0^2 = \frac{s^2(n-1)}{\sigma_0^2}$$



χ^2 -próba

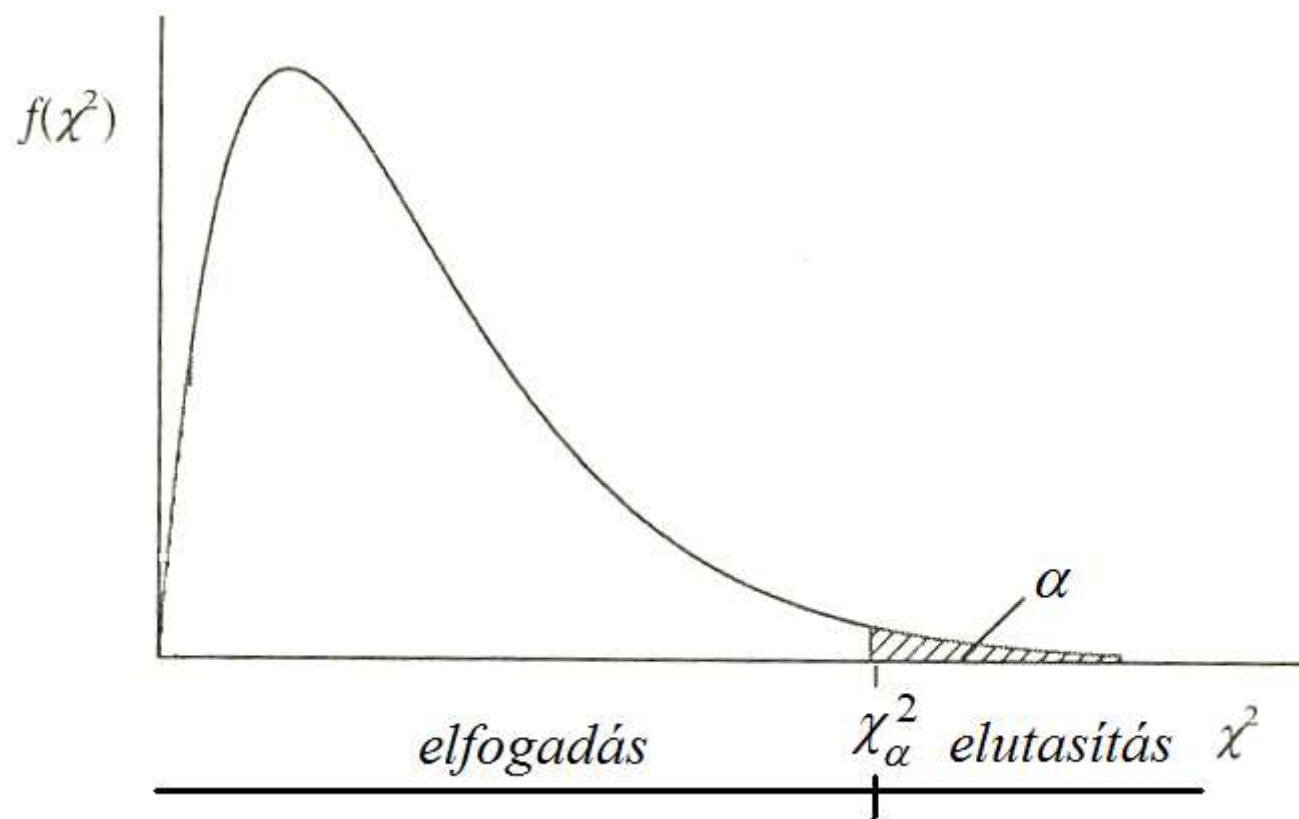
- adott α szignifikancia szint mellett annak valószínűsége, hogy a χ_0^2 próbastatisztika értéke az elfogadási tartományba esik $1-\alpha$:

$$P(\chi_0^2 \leq \chi_\alpha^2) = P\left(\frac{s^2(n-1)}{\sigma_0^2} \leq \chi_\alpha^2\right) = 1 - \alpha$$

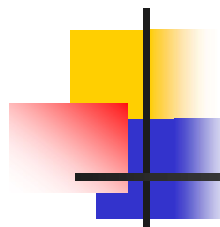
- azaz H_0 -t elfogadom, ha $\frac{s^2(n-1)}{\sigma_0^2} < \chi_\alpha^2$
- H_0 -t elutasítom, ha $\frac{s^2(n-1)}{\sigma_0^2} > \chi_\alpha^2$

χ^2 -próba

- elfogadási/elutasítási tartomány a sűrűségfüggvény alapján



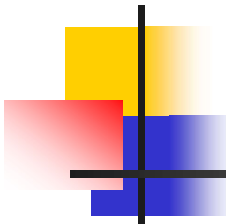
- ha H_0 igaz, akkor α valószínűsége annak, hogy $\chi_0^2 > \chi_\alpha^2$
- ha H_1 igaz, akkor χ_0^2 eloszlása $\chi^2 \sigma^2 / \sigma_0^2$ és mivel $\sigma^2 > \sigma_0^2$, így általában χ_α^2 -nél nagyobb értéket vesz fel
- mindenesetre kis α -nál nagyobb a valószínűsége, hogy χ_α^2 -től jobbra eső értéket vegyen fel, tehát a $H_1: \sigma^2 > \sigma_0^2$ ellenhipotézis jobb oldali ellenhipotézis
- példa



χ^2 -próba

- χ^2 -eloszlás kritikus értékei α és ν függvényében

ν	α									
	0.999	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.001
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.706	3.841	5.024	6.635	10.827
2	0.0020	0.0201	0.0506	0.1026	0.2107	4.605	5.991	7.378	9.210	13.815
3	0.0243	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.345	16.266
4	0.0908	0.2971	0.4844	0.7107	1.064	7.779	9.488	11.143	13.277	18.466
5	0.2102	0.5543	0.8312	1.145	1.610	9.236	11.070	12.832	15.086	20.515
6	0.3810	0.8721	1.237	1.635	2.204	10.645	12.592	14.449	16.812	22.457
7	0.5985	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	24.321
8	0.8571	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	26.124
9	1.152	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	27.877
10	1.479	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	29.588
11	1.834	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	31.264
12	2.214	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	32.909
13	2.617	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	34.527
14	3.041	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	36.124
15	3.483	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	37.698
16	3.940	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	39.252



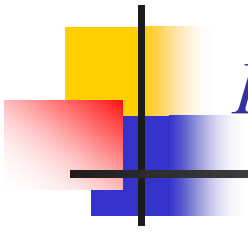
F-próba

- két szórásnégyzet összehasonlítása: *F*-próba
- cél: két, *normális eloszlású* sokaságból vett minta szórásnégyzetének összehasonlításával kell eldönteni, hogy a varianciák egyeznek-e
- ha a két variancia azonos, akkor a szórásnégyzetek eloszlása *F*-eloszlású:

$$H_0: \sigma_1^2 = \sigma_2^2$$

- próbastatisztika

$$F_0 = \frac{s_1^2}{s_2^2} = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \cdot \frac{\sigma_1^2}{\sigma_2^2} = F(n_1 - 1, n_2 - 1) \frac{\sigma_1^2}{\sigma_2^2}$$



F -próba

- azaz F_0 csak akkor F eloszlású, ha $\sigma_1^2 = \sigma_2^2$, ellenkező esetben attól lefelé ($\sigma_1^2/\sigma_2^2 < 1$), vagy felfelé ($\sigma_1^2/\sigma_2^2 > 1$) eltér

- egyoldalas ellenhipotézis

$$H_1: \sigma_1^2 > \sigma_2^2$$

- H_0 -t elutasítom, ha $s_1^2/s_2^2 > F_\alpha$

- kétoldalas ellenhipotézis

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

- H_0 -t elutasítom, ha $s_1^2/s_2^2 < F_{1-\alpha/2}$ vagy $s_1^2/s_2^2 > F_{\alpha/2}$

- írjuk a nagyobb tapasztalati szórást a számlálóba, azaz legyen $s_1^2 > s_2^2$, azaz $s_1^2/s_2^2 > 1$, ekkor elég csak a felső határt ellenőrizni:

$$s_1^2/s_2^2 < F_{\alpha/2}(v_1, v_2)$$

mivel ekkor biztos, hogy

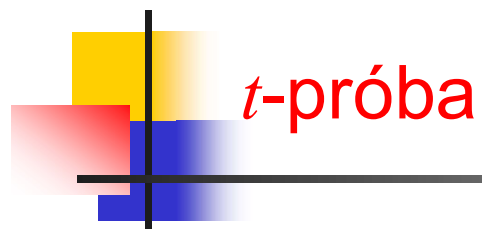
$$s_2^2/s_1^2 > F_{1-\alpha/2}(v_2, v_1), \text{ ha } s_1^2 > s_2^2$$

- a kétoldalas próba szignifikancia szintje nem $\alpha/2$ hanem α lesz!
- példa

F-eloszlás táblázat

- F_α kritikus értékei $\alpha = 0.05$ egyoldali szinten

v_1	nevező szabadsági foka										v_2
	1	2	3	4	5	6	7	8	9	10	
számláló szabadsági foka	1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96
	2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10
	3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71
	4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48
	5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33
	6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22
	7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14
	8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07
	9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02
	10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98



t -próba

- az u -próbával ismert varianciájú sokaság esetében a minta értékek átlaga alapján következtettünk a sokaság várható értékére
- a variancia meghatározásához nagy elemszámú mérési adatsor kell – nem mindig áll rendelkezésre
- a t -eloszlás hasonló az u -eloszláshoz, csak a variancia helyett a tapasztalati szórás szerepel benne
- így, ha nem áll rendelkezésre a minták háttérében lévő sokaság elméleti szórásnégyzete, akkor az u -próba helyett a t -próba alkalmazható

- Egymintás t -próba

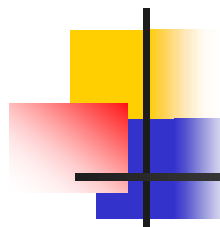
- analóg az u -próbával
- cél: a várható érték megegyezik-e egy adott (pl. a mérések alapján meghatározott) értékkel
- számolásnál σ^2 helyett s^2
- hipotézisek:

$$H_0 : \mu = \mu_0 \text{ vagy } H_0 : \mu - \mu_0 = 0$$

$$H_1 : \mu \neq \mu_0 \text{ vagy } H_1 : \mu - \mu_0 \neq 0$$

- próbastatisztika

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$



t -próba

- t -eloszlás $t_{\alpha/2}$ kritikus értékei α és ν függvényében

ν	α						
	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	3.078	6.314	12.706	25.452	63.656	127.321	636.578
2	1.886	2.920	4.303	6.205	9.925	14.089	31.600
3	1.638	2.353	3.182	4.177	5.841	7.453	12.924
4	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	1.476	2.015	2.571	3.163	4.032	4.773	6.869
6	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	1.415	1.895	2.365	2.841	3.499	4.029	5.408
8	1.397	1.860	2.306	2.752	3.355	3.833	5.041
9	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	1.341	1.753	2.131	2.490	2.947	3.286	4.073

- Kétmintás t -próba

- cél: két, *egymástól független* minta mögött álló sokaság várható értékeinek egyezőségének vagy különbözőségének kimutatása

- hipotézisek:

$$H_0 : \mu_1 = \mu_2 \text{ vagy } H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 \neq \mu_2 \text{ vagy } H_1 : \mu_1 - \mu_2 \neq 0$$

- szükséges adatok:

n_1, n_2 – a két minta elemszáma

s_1^2, s_2^2 – a két minta tapasztalati szórása

- elsőként a két minta mögötti két sokaság varianciájának egyezőségét kell ellenőrizni – F -próba
 - ha teljesül, akkor mehet tovább
 - ha nem teljesül, akkor más eljárást kell alkalmazni

- Legyen d a következő v.v.

$$d = \bar{x}_1 - \bar{x}_2$$

- belátható, hogy d is normális eloszlású v.v.,
paraméterei

$$E(d) = E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2) = \mu_1 - \mu_2$$

$$\begin{aligned} Var(d) &= Var(\bar{x}_1 - \bar{x}_2) = Var(\bar{x}_1) + Var(\bar{x}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \\ &= \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \end{aligned}$$

egymástól független minták!

- legyen s_d^2 egy d -től független v.v.:

$$s_d^2 = s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

ahol s^2 a minták közös tapasztalati szórásnégyzete, meghatározása:

$$s^2 = \frac{1}{n_1 + n_2 - 2} \left(s_1^2(n_1 - 1) + s_2^2(n_2 - 1) \right)$$

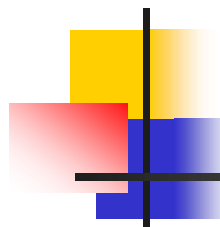
- s_d^2 kiszámításához célszerű az egyesített s^2 -t alkalmazni, mivel ennek szabadsági foka nagyobb, mint akár s_1^2 , akár s_2^2 -nek, így kisebb kritikus érték tartozik hozzá

- az alábbi kifejezés t -eloszlású $\nu = n_1 + n_2 - 2$ szabadsági fokkal:

$$t = \frac{d - E(d)}{s_d} = \frac{d - E(d)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- legyen a H_0 hipotézis

$$H_0 : \mu_1 = \mu_2 \Rightarrow E(d)=0$$



t -próba

- a próbastatisztika

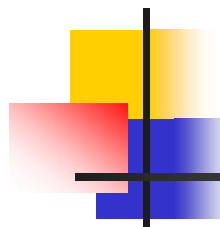
$$t_0 = \frac{d - 0}{s_d} = \frac{d}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \nu = n_1 + n_2 - 2$$

- ha a próbastatisztika t -eloszlású, azaz

$$-t_{\alpha/2} < t_0 \leq t_{\alpha/2},$$

akkor az átlagértékek különbözősége α szinten nem szignifikáns

- példa

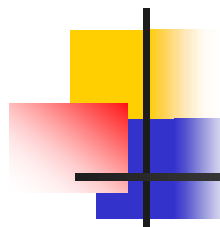


F -eloszlás táblázat

- F_α kritikus értékei $\alpha = 0.05$ egyoldali szinten

IV. táblázat. Az F -eloszlás F_α kritikus értékei $\alpha = 0.05$ egyoldali szinten

		v_2 nevező szabadsági foka														
v_1		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
számláló szabadsági foka	1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96	4.84	4.75	4.67	4.60	4.54
	2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10	3.98	3.89	3.81	3.74	3.68
	3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71	3.59	3.49	3.41	3.34	3.29
	4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48	3.36	3.26	3.18	3.11	3.06
	5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33	3.20	3.11	3.03	2.96	2.90
	6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22	3.09	3.00	2.92	2.85	2.79
	7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14	3.01	2.91	2.83	2.76	2.71
	8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07	2.95	2.85	2.77	2.70	2.64
	9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02	2.90	2.80	2.7	2.65	2.59
	10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98	2.85	2.75	2.67	2.60	2.54



t -próba

- t -eloszlás $t_{\alpha/2}$ kritikus értékei α és ν függvényében

ν	α						
	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	3.078	6.314	12.706	25.452	63.656	127.321	636.578
2	1.886	2.920	4.303	6.205	9.925	14.089	31.600
3	1.638	2.353	3.182	4.177	5.841	7.453	12.924
4	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	1.476	2.015	2.571	3.163	4.032	4.773	6.869

16	1.337	1.746	2.120	2.473	2.921	3.252	4.015
17	1.333	1.740	2.110	2.458	2.898	3.222	3.965
18	1.330	1.734	2.101	2.445	2.878	3.197	3.922
19	1.328	1.729	2.093	2.433	2.861	3.174	3.883
20	1.325	1.725	2.086	2.423	2.845	3.153	3.850
21	1.323	1.721	2.080	2.414	2.831	3.135	3.819
22	1.321	1.717	2.074	2.405	2.819	3.119	3.792
23	1.319	1.714	2.069	2.398	2.807	3.104	3.768
24	1.318	1.711	2.064	2.391	2.797	3.091	3.745
25	1.316	1.708	2.060	2.385	2.787	3.078	3.725

- Páros t -próba
 - legyen x és y két normális eloszlású valószínűsége változó, a következő feltétellel
 - a két minta, amelyből x és y származik, *nem független egymástól, de a mérésakor elkövetett hibák igen*
 - példa
 - Nullhipotézis:

$$H_0 : E(x_i) = E(y_i)$$

- legyen d valószínűségi változó:

$$d_i = x_i - y_i$$

- paraméterei:

- várható érték: $E(d_i) = E(x_i) - E(y_i)$
- variancia (a mérési hibák függetlensége miatt):

$$Var(d_i) = Var(x_i) + Var(y_i)$$

- a páronkénti eltérés átlagértéke:

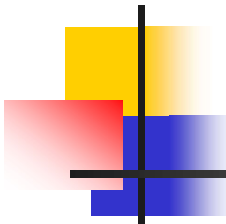
$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$$

- szórásnégyzete

$$s_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n - 1} = \frac{\sum_{i=1}^n d_i^2 - n\bar{d}^2}{n - 1}$$

- a következő kifejezés t -eloszlású:

$$t_0 = \frac{\bar{d} - E(\bar{d})}{s_d / \sqrt{n}}$$



t -próba

- Ha H_0 igaz:

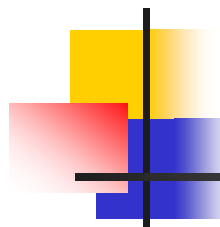
$$E(\bar{d}) = 0$$

- a próbastatisztika:

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}}$$

- elfogadom H_0 -t α szignifikancia szinten, ha

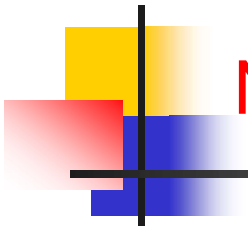
$$-t_{\alpha/2} < t_0 \leq t_{\alpha/2}$$



t -próba

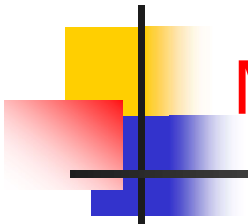
- t -eloszlás $t_{\alpha/2}$ kritikus értékei α és ν függvényében

ν	α						
	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	3.078	6.314	12.706	25.452	63.656	127.321	636.578
2	1.886	2.920	4.303	6.205	9.925	14.089	31.600
3	1.638	2.353	3.182	4.177	5.841	7.453	12.924
4	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	1.476	2.015	2.571	3.163	4.032	4.773	6.869
6	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	1.415	1.895	2.365	2.841	3.499	4.029	5.408
8	1.397	1.860	2.306	2.752	3.355	3.833	5.041
9	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	1.341	1.753	2.131	2.490	2.947	3.286	4.073



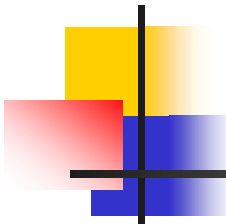
Nemparaméteres statisztikai eljárások

- paraméteres próbák alkalmazásához kellenek a minták számszerű értékei és bizonyos feltevéseknek (pl. normális eloszlás) kell teljesülnie az alapsokaságokra
- a nemparaméteres próbák esetében a elegendő, ha az adatokat rangsorolni tudjuk és nincsenek előfeltételek az alapsokaságokra



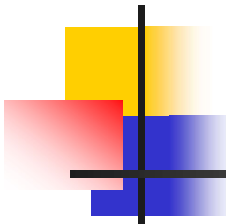
Mann-Whitney U-próba

- a két mintás t -próba nemparaméteres párja, azaz legyen adott *két független sokaság*, és ezek várható értékeinek egyezőségét vizsgáljuk mintavétel segítségével
- akkor célszerű alkalmazni, ha a kétmintás t -próba előfeltételei nem biztos, hogy teljesülnek
 - kétmintás t -próba alkalmazható, ha mindkét sokaság normális eloszlású és a szórásuk megegyezik



Mann-Whitney U-próba

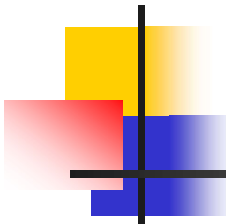
- az alkalmazás menete
 - legyen n_1 számú mintánk az 1. adathalmazból és n_2 a 2. adathalmazból
 - legyen
 - H_0 : a nincs különbség a két adathalmaz eloszlásában a várható értéket tekintve
 - H_1 : van különbség a várható értékekben (kétoldalas próba) vagy az egyik adathalmaz a másikhoz képest jobbra eltolódik (egyoldalas próba)



Mann-Whitney U-próba

- végezzük el a rangsorolást valamennyi adat figyelembe vételével (azonos értékek esetén az átlagrang alkalmazandó)
- határozzuk meg a rangösszegeket az egyes adathalmazokból vett minták esetében $\rightarrow T_1, T_2$
- Mann-Whitney U-próba képletei:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - T_1 \quad U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - T_2$$



Mann-Whitney U-próba

- $U_1 < U_2$, ha T_1 nagy, azaz ha az 1. adathalmaz adatainak eloszlása jobbra tolódik el a 2. adathalmaz eloszlásához képest
- egy oldalas próba esetén elvetjük H_0 -t, ha U_1 kisebb, mint egy meghatározott U_0 küszöbérték
- ugyanez levezethető U_2 -re nézve is
- az U_0 küszöbértéket, a minták elemszámai alapján, táblázatból határozzuk meg
- a táblázat megfelelő cellája megadja annak valószínűségét, hogy a kapott U_{1v2} érték kisebb, mint U_0 , de a H_0 hipotézis igaz $\rightarrow \alpha$ érték

Mann-Whitney U-próba

- U_0 meghatározása

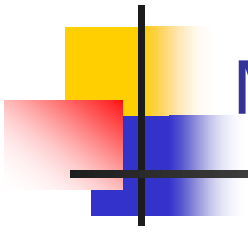
Table II
Distribution function
of U, $P(U \leq U_0)$;
 U_0 is the argument;
 $n_1 \leq n_2$;
 $3 \leq n_2 \leq 10$.

$n_2 = 3$

U_0	n_1		
	1	2	3
0	.25	.10	.05
1	.50	.20	.10
2		.40	.20
3		.60	.35
4			.50

$n_2 = 4$

U_0	n_1			
	1	2	3	4
0	.2000	.0667	.0286	.0143
1	.4000	.1333	.0571	.0286
2	.6000	.2667	.1143	.0571
3		.4000	.2000	.1000
4		.6000	.3143	.1714
5			.4286	.2429
6			.5714	.3429
7				.4429
8				.5571



Mann-Whitney U-próba

- Egy oldalas teszt elvégzése
 - H_0 : a nincs különbség a két adathalmaz eloszlásában a várható értéket tekintve
 - H_1 : az egyik adathalmaz – legyen ez az 1. adathalmaz a másikhoz képest jobbra eltolódik
- rangsorolás elvégzése
- rangösszegek meghatározása – T_1
- képletek alkalmazása – U_1 kiszámítása
- U_0 meghatározása a táblázatból adott α szignifikancia szinthez

Mann-Whitney U-próba

- U_0 meghatározása

Table II
Distribution function
of U, $P(U \leq U_0)$;
 U_0 is the argument;
 $n_1 \leq n_2$;
 $3 \leq n_2 \leq 10$.

$n_2 = 3$

U_0	n_1		
	1	2	3
0	.25	.10	.05
1	.50	.20	.10
2		.40	.20
3		.60	.35
4			.50

például legyen

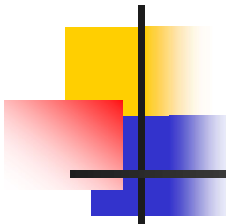
2. halmaz elemszám $n_2 = 4$ 1 $\rightarrow$ $n_2 = 4$

1. halmaz elemszám $n_1 = 3$ 2 $\rightarrow$

U_0 küszöbérték meghatározása 4 $\rightarrow$

szignifikancia szint $\alpha \approx 0,05$ 3 $\rightarrow$

U_0	n_1			
	1	2	3	4
0	.2000	.0667	.0286	.0143
1	.4000	.1333	.0571	.0286
2	.6000	.2667	.1143	.0571
3		.4000	.2000	.1000
4		.6000	.3143	.1714
5			.4286	.2429
6			.5714	.3429
7				.4429
8				.5571



Mann-Whitney U-próba

- Két oldalas teszt elvégzése
 - H_0 : a nincs különbség a két adathalmaz eloszlásában a várható értéket tekintve
 - H_1 : a két adathalmaz egymástól különbözik
- rangsorolás elvégzése
- rangösszegek meghatározása – T_1 és T_2
- képletek alkalmazása – U_1 és U_2 kiszámítása
- a teszthez a kisebb értéket használjuk
- U_0 meghatározása a táblázatból adott α szignifikancia szinthez

Mann-Whitney U-próba

- U_0 meghatározása

Table II
Distribution function
of U , $P(U \leq U_0)$;
 U_0 is the argument;
 $n_1 \leq n_2$;
 $3 \leq n_2 \leq 10$.

$n_2 = 3$

	n_1		
U_0	1	2	3
0	.25	.10	.05
1	.50	.20	.10
2		.40	.20
3		.60	.35
4			.50

például legyen

2. halmaz elemszám $n_2 = 4$ 1 $\rightarrow$

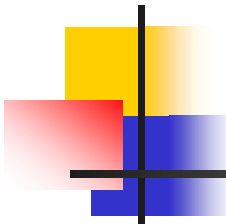
1. halmaz elemszám $n_1 = 3$ 2 $\rightarrow$

$n_2 = 4$

	n_1			
U_0	1	2	3	4
0	.2000	.0667	.0286	.0143
1	.4000	.1333	.0571	.0286
2	.6000	.2667	.1143	.0571
3		.4000	.2000	.1000
4		.6000	.3143	.1714
5			.4286	.2429
6			.5714	.3429
7				.4429
8				.5571

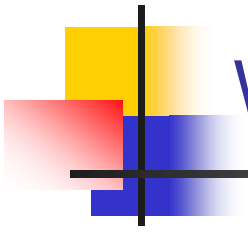
U_0 küszöbérték meghatározása 4 $\rightarrow$

ha szignifikancia szint $\alpha \approx 0,05$
akkor a határértéket $\alpha/2 \approx 0,025$ -hez
keressük 3 $\rightarrow$



Wilcoxon-próba

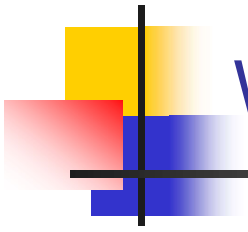
- két összetartozó minta várhatóértéke egyenlőségének ellenőrzésére → páros t -próba nemparaméteres párja
- páronkénti összehasonlításon alapul
- ha nincs különbség a két eloszlás között, akkor a különbségek fele részben pozitívak, fele részben negatívak lesznek, és nagyságrendileg megegyeznek egymással
- azaz a rangsorolt különbségeknél a pozitív és a negatív különbségek rangösszegei megegyeznek



Wilcoxon-próba

- az alkalmazás menete
 - legyen n számú mintánk az 1. adathalmazból és a 2. adathalmazból egyaránt
 - legyen
$$H_0 : \text{nincs különbség a két adathalmaz eloszlásában a várható értéket tekintve}$$
$$H_1 : \text{van különbség a várható értékekben (kétoldalas próba) vagy}$$

az egyik adathalmaz a másikhoz képest jobbra eltolódik (egyoldalas próba)



Wilcoxon-próba

- határozzuk meg páronként a különbségeket
 - ha két adat megegyezik, akkor azokat hagyjuk figyelmen kívül, de csökkentsük n értékét
- rangsoroljuk az abszolút érték alapján a különbségeket (legkisebb különbség kapja az 1 rangot)
- határozzuk meg a rangösszegeket külön a pozitív és a negatív különbségekre $\rightarrow T^+, T^-$

- **kétoldalas teszt**nél: a két rangösszeg közül a kisebbet használjuk a H_0 hipotézis ellenőrzésére, és elvetjük H_0 -t, ha ez az érték kisebb a táblázatból meghatározott T_0 küszöbértéknél
- **egyoldalas teszt**nél:
 - ha az 1. sokaság jobbra tolódását vizsgáljuk a 2.-hez képest a minták alapján, akkor a T^- -t használjuk, és elvetjük H_0 -t, ha $T^- \leq T_0$
 - ha a 2. sokaság jobbra tolódását vizsgáljuk a 1.-hez képest a minták alapján, akkor a T^+ -t használjuk, és elvetjük H_0 -t, ha $T^+ \leq T_0$

Wilcoxon-próba

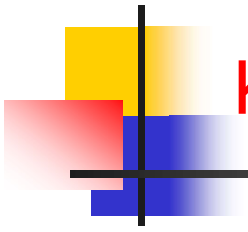
- T_0 meghatározása

Table 12
Critical values of T in
the Wilcoxon signed-
rank test; n = 5(1)50

One-sided	Two-sided	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$P = .05$	$P = .10$	1	2	4	6	8	11
$P = .025$	$P = .05$		1	2	4	6	8
$P = .01$	$P = .02$			0	2	3	5
$P = .005$	$P = .01$				0	2	3

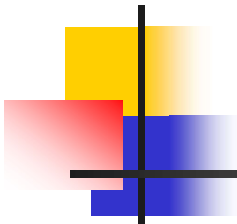
One-sided	Two-sided	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$P = .05$	$P = .10$	14	17	21	26	30	36
$P = .025$	$P = .05$	11	14	17	21	25	30
$P = .01$	$P = .02$	7	10	13	16	20	24
$P = .005$	$P = .01$	5	7	10	13	16	19

One-sided	Two-sided	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$P = .05$	$P = .10$	41	47	54	60	68	75
$P = .025$	$P = .05$	35	40	46	52	59	66
$P = .01$	$P = .02$	28	33	38	43	49	56
$P = .005$	$P = .01$	23	28	32	37	43	49



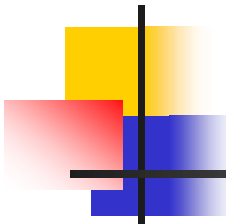
Kruskal-Wallis H-próba

- alkalmazható *kettőnél több, független sokaság* szórás egyezőségének vizsgálatára véletlen mintavétel esetén
- az F -próba több alapsokaság esetén alkalmazható módosításának (Cohran-próba, Levene-próba) nemparaméteres változata
- analógia: Mann-Whitney próba a kétmintás t -próba nemparaméteres párja, a Kruskal-Wallis próba az F -próba nemparaméteres párja



Kruskal-Wallis H-próba

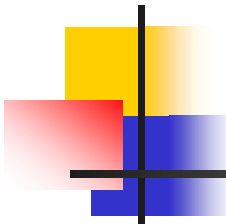
- az alkalmazás menete
 - legyen a sokaságok száma p , és ezekből vegyünk rendre $n_1, n_2, \dots, n_p$ számú mintát, n legyen a minták elemszámának összege
 - legyen
 - H_0 : a nincs különbség az adathalmazok eloszlásában a szórás értékét tekintve
 - H_1 : legalább két adathalmaz szórása eltér egymástól a p számú adathalmaz közül



Kruskal-Wallis H-próba

- végezzük el a rangsorolást valamennyi adat figyelembe vételével (azonos értékek esetén az átlagrang alkalmazandó)
- határozzuk meg a rangösszegeket az egyes adathalmazokból vett minták esetében
 $\rightarrow T_1, T_2, \dots, T_p$
- Kruskal-Wallis H-próba képlete:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^p \frac{T_i^2}{n_i} - 3(n+1)$$



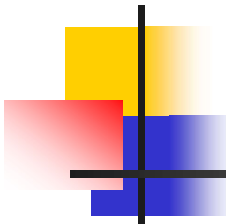
Kruskal-Wallis H-próba

- elvetjük H_0 hipotézist, ha kapott H érték egy küszöbértéknél nagyobb
- belátható, ha a minták elemszámai elegendően nagyok ($n_i \geq 5, \forall i$ -re) és H_0 igaz, akkor H közelítőleg χ^2 eloszlásnak felel meg $\nu = (p - 1)$ szabadsági fokkal
- így elvetjük H_0 -t, ha adott α érték esetén H érték meghaladja χ_α^2 küszöbértéket

Kruskal-Wallis H-próba

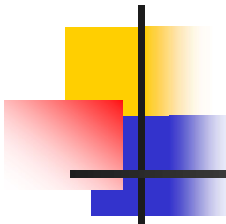
- χ^2 -eloszlás kritikus értékei α és ν függvényében

ν	α									
	0.999	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.001
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.706	3.841	5.024	6.635	10.827
2	0.0020	0.0201	0.0506	0.1026	0.2107	4.605	5.991	7.378	9.210	13.815
3	0.0243	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.345	16.266
4	0.0908	0.2971	0.4844	0.7107	1.064	7.779	9.488	11.143	13.277	18.466
5	0.2102	0.5543	0.8312	1.145	1.610	9.236	11.070	12.832	15.086	20.515
6	0.3810	0.8721	1.237	1.635	2.204	10.645	12.592	14.449	16.812	22.457
7	0.5985	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	24.321
8	0.8571	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	26.124
9	1.152	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	27.877
10	1.479	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	29.588
11	1.834	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	31.264
12	2.214	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	32.909
13	2.617	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	34.527
14	3.041	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	36.124
15	3.483	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	37.698
16	3.940	5.812	6.909	7.962	9.312	23.542	26.296	28.845	32.000	39.252



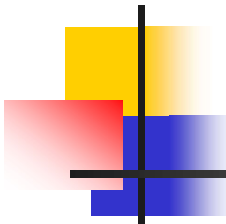
Friedman F_r -próba

- alkalmazható *kettőnél több, független sokaság* középértéke egyezőségének vizsgálatára véletlen mintavétel esetén
- összetartozó minták mediánjának összehasonlítása
- változók száma: b (>2)
- szempontok száma: p
- pl. p szempont szerint hasonlítunk össze b kísérleti személyt vagy esetet



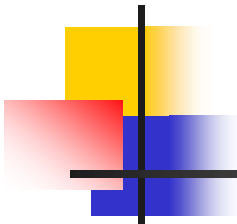
Friedman F_r -próba

- az alkalmazás menete
 - legyen az összehasonlítandó szempontok száma p , és az összehasonlítást végezzük el b számú kísérleti esetben (az összes megfigyelés száma bp lesz!)
 - legyen
 - H_0 : a nincs különbség az adathalmazok eloszlásában a várható értéket tekintve
 - H_1 : legalább két adathalmaz különbözik a p számú adathalmaz közül a várható értékeket nézve



Friedman F_r -próba

- rendezzük el az adatokat: célszerűen az összehasonlítható szempontokat az oszlopokban, az ugyanahhoz a kísérleti értékeket sorokban
- végezzük el a rangsorolást soronként, azaz az egyes kísérletekhez tartozó adatokat rangsoroljuk a szokásos módon
- határozzuk meg a rangösszegeket az oszlopok szerint, azaz a rangösszegek mindig ugyanahhoz a szemponthoz tartoznak $\rightarrow T_1, T_2, \dots, T_p$

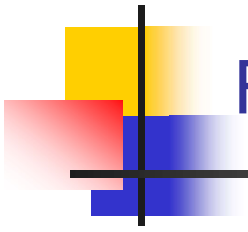


Friedman Fr -próba

- Friedman Fr -próba képlete

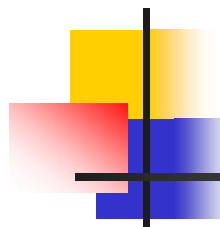
$$Fr = \frac{12}{bp(p+1)} \left(\sum_{i=1}^p T_i^2 \right) - 3b(p+1)$$

- a próba értéke (Fr) akkor lesz minimális, ha a rangösszegek közelítőleg megegyeznek
- ha a rangösszegek között eltérés van, akkor ez a próba értékének a növekedését okozza



Friedman Fr -próba

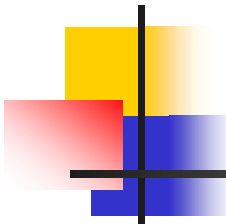
- elvetjük H_0 hipotézist, ha kapott Fr érték egy küszöbértéknél nagyobb
- belátható, ha vagy $p \geq 5$ vagy $b \geq 5$ és H_0 igaz, akkor Fr közelítőleg χ^2 eloszlásnak felel meg $\nu = (p - 1)$ szabadsági fokkal
- így elvetjük H_0 -t, ha adott α érték esetén Fr érték meghaladja χ_α^2 küszöbértéket



Friedman Fr -próba

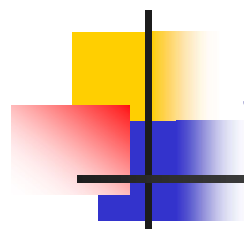
- χ^2 -eloszlás kritikus értékei α és ν függvényében

ν	α									
	0.999	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.001
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.706	3.841	5.024	6.635	10.827
2	0.0020	0.0201	0.0506	0.1026	0.2107	4.605	5.991	7.378	9.210	13.815
3	0.0243	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.345	16.266
4	0.0908	0.2971	0.4844	0.7107	1.064	7.779	9.488	11.143	13.277	18.466
5	0.2102	0.5543	0.8312	1.145	1.610	9.236	11.070	12.832	15.086	20.515
6	0.3810	0.8721	1.237	1.635	2.204	10.645	12.592	14.449	16.812	22.457
7	0.5985	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	24.321
8	0.8571	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	26.124
9	1.152	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	27.877
10	1.479	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	29.588
11	1.834	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	31.264
12	2.214	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	32.909
13	2.617	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	34.527
14	3.041	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	36.124
15	3.483	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	37.698
16	3.940	5.812	6.909	7.962	9.312	23.542	26.296	28.845	32.000	39.252



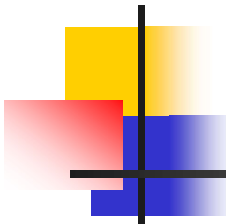
Spearman r_s rangkorreláció

- alkalmazható két változó közötti viszony / összefüggés kimutatására
 - pl. egyik változó növekedése okozza-e a másik változó növekedését ill. csökkenését, vagy a két változó független egymástól
 - a próbához két összefüggő adatsor kell, ezeknek az értékeit hasonlítjuk össze a rangsorolás segítségével
- a korreláció számítás nemparaméteres párja



Spearman r_s rangkorreláció

- az alkalmazás menete
 - legyen a két összehasonlítandó változóból n számú megfigyelés (mintapár)
 - hipotézisek
$$H_0 : \text{a nincs összefüggés az adatok között}$$
$$H_1 : \text{van összefüggés az adatok között}$$
 - elvégezzük a rangsorolást az egyes változókra, majd meghatározzuk páronként az eltérések négyzetösszegét
 - ha két rang megegyezik, akkor ezeket az adatokat figyelmen kívül hagyjuk (n -t csökkentjük!)



Spearman r_S rangkorreláció

- Spearman r_S -próba képlete

$$r_S = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n d_i^2$$

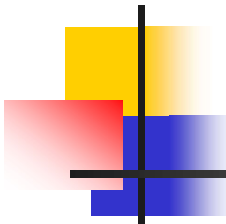
- a próba értéke $-1 \leq r_S \leq 1$ esik
- *kétoldalas próbánál* elvetjük a H_0 hipotézist (nincs összefüggés az adatok között), ha r_S értéke kívül van egy a próbához tartozó táblázatból meghatározható $[-r_0, r_0]$ tartományon ($r_S \leq -r_0 \vee r_S \geq r_0$)
- *egy oldalas próbánál* elvetjük a H_0 -t, ha $|r_S| \geq r_0$

Spearman r_s rangkorreláció

- táblázat értékei egy oldalas próbára a párok száma (n) és a szignifikancia szint függvényében (α)

Table I3
*Critical values of
Spearman's rank
correlation coefficient
for a one-tailed test*

n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	0.900	—	—	—
6	0.829	0.886	0.943	—
7	0.714	0.786	0.893	—
8	0.643	0.738	0.833	0.881
9	0.600	0.683	0.783	0.833
10	0.564	0.648	0.745	0.794
11	0.523	0.623	0.736	0.818
12	0.497	0.591	0.703	0.780
13	0.475	0.566	0.673	0.745
14	0.457	0.545	0.646	0.716
15	0.441	0.525	0.623	0.689



Spearman r_s rangkorreláció

- egyoldalas teszt: az egyik változó növekedése (csökkenése) a másik változó növekedését (csökkenését) okozza – azonos irányú hatás
- kétoldalas teszt: az egyik változó növekedése (csökkenése) a másik változó csökkenését (növekedését) okozza – ellentétes irányú hatás
- kétoldalas teszthez a szignifikancia szintet felezni kell - $\alpha/2$ -nél kell a küszöbértéket nézni