Egyoldali próba

A szegedi paprika aflatoxin-szennyezése

Határérték 5μg/kg

Bács-Kiskun Megyei Állategészségügyi és Élelmiszerellenőrző Állomás nem akkreditált laborjának mérése szerint 4.8 μg/kg ±25%

OÉTI: génkárosító, rákkeltő, a határértékhez közeli eredmény miatt meg kellett volna ismételni a vizsgálatot

laborvezető: a ±25% azt jelenti, hogy 5μg/kg +25%=6.25 μg/kg megengedhető

$$\mathrm{H}_0 : \mu \le \mu_0 = 5\,\mu\mathrm{g/kg} \qquad \mathrm{H}_1 : \mu > \mu_0 = 5\,\mu\mathrm{g/kg}$$

Ha elutasítjuk H_0-t, azt látjuk bizonyítva, hogy a megengedettnél több van benne (a hatóság szempontja).
Ha elfogadjuk H_0-t, semmit nem látunk bizonyítva.
Itt elfogadták, tehát nem bizonyított, hogy a határértéket meghaladja.

$$\mathrm{H}'_0 : \mu \ge \mu_0 = 5\,\mu\mathrm{g/kg} \qquad \mathrm{H}'_1 : \mu < \mu_0 = 5\,\mu\mathrm{g/kg}$$

Ha elutasítjuk $H`_0$-t, azt látjuk bizonyítva, hogy a megengedettnél kevesebb van benne (a kibocsátó kötelezettsége).
Ha elfogadjuk $H`_0$-t, semmit nem látunk bizonyítva.
Itt elfogadták, tehát nem bizonyított, hogy a határérték alatt van.

Egymintás *t*-próba

1. példa
Egy analitikai módszer torzítatlanságának vizsgálatára 5 ismételt mérést végeztek egy 3.25% ismert koncentrációjú munka-standarddel.
Az eredmények: 3.25, 3.27, 3.24, 3.26 és 3.24.
Kérdezzük, hogy a módszer torzítatlan-e.
(Elfogadva, hogy az adatok közelítőleg normális eloszlásúak, ellenőrizzük 5%-os szignifikanciaszinten a torzítatlanság hipotézisét!)

$\mathrm{H}_0 : \mu = \mu_0 = 3.25$

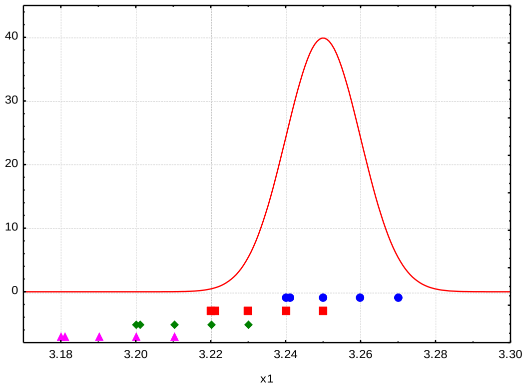
nullhipotézis

$\mathrm{H}_1 : \mu \neq \mu_0 = 3.25$

ellenhipotézis

A statisztikai próba filozófiája

A feltételezett várható érték 3.25, ehhez rajzoltuk meg a Gauss-görbét.



Elhisszük? Adhatja a kék pontokat ez az eloszlás? És a pirosakat?...

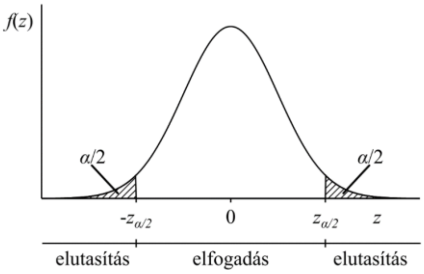
Milyen valószínűséggel?

$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu \neq \mu_0$

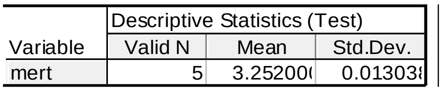
$$z = \frac{\bar{x} - \enclose{circle}{\mu}}{\sigma/\sqrt{n}} \longrightarrow z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$ próbastatisztika

standard normális eloszlás

Ha H_0 igaz, $z_0 \sim z$ (z-próba)
Ha z_0 olyan értékeket vesz föl, amilyeneket z szokott, elfogadjuk.

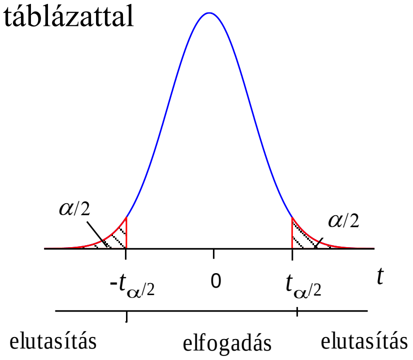


A próba elvégzéséhez σ ismerete lenne szükséges, ezért *t*-próbát használunk.



Variable	Descriptive Statistics (Test)		
	Valid N	Mean	Std.Dev.
mert	5	3.25200	0.01303

táblázattal



elfogadási tartomány

$$P\left(-t_{\alpha/2} < \overset{t_0}{\frac{\bar{x} - \mu_0}{s/\sqrt{n}}} \leq t_{\alpha/2} \right) = 1 - \alpha$$

$$t_{0.05/2}(4) = 2.776$$

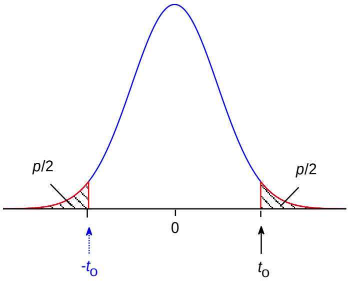
$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{3.252 - 3.25}{0.0134/\sqrt{5}} = 0.34$$

$$-2.776 < 0.34 < 2.776$$ elfogadjuk

programmal

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{3.252 - 3.25}{0.0134/\sqrt{5}} = 0.34$$

próbastatisztika



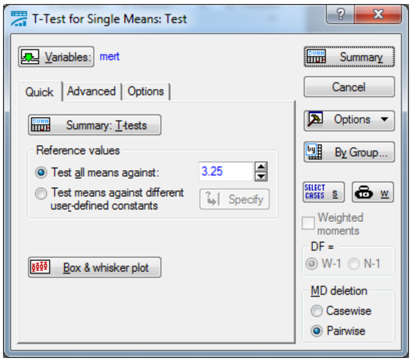
p annak valószínűsége, hogy a próbastatisztika a talált (t_0) értéket vagy annál szélsőségesebbet vegyen föl, ha H_0 igaz.

$$H_0 : \mu = \mu_0 = 3.25$$

Ha *p* elég nagy, elfogadjuk.

---

programmal: Statistics > Basic Statistics/Tables > t-test, single sample



Variables: mert

Reference values

● Test all means against: 3.25

○ Test means against different user-defined constants

Specify

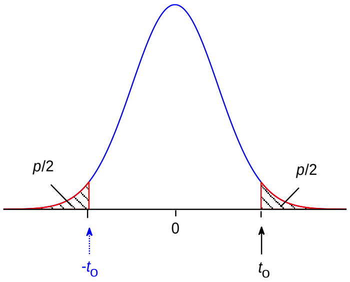
☐ Weighted moments

DF = ● W-1 ○ N-1

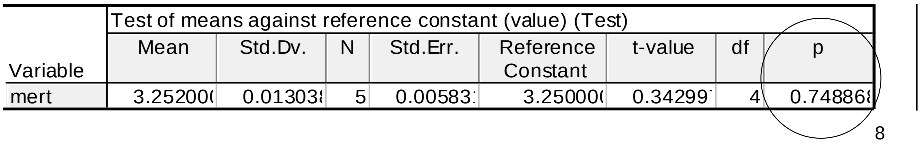
MD deletion

○ Casewise

● Pairwise



A *p* elég nagy, elfogadjuk.



	Test of means against reference constant (value) (Test)							
Variable	Mean	Std.Dv.	N	Std.Err.	Reference Constant	t-value	df	p
mert	3.25200(	0.01303ł	5	0.00583:	3.25000(	0.34299ˈ	4	0.74886ł

Ez az eljárás a „Null Hypothesis Significance Testing” (NHST).

p értelmezése

A *p* annak valószínűsége, hogy a kapott vagy még szélsőségesebb adatok adódjanak, ha H_0 igaz.

Amit szeretnénk: $P(\mathrm{H}_0|\text{adatok}) = ?$
(gyakori féreértelmezés)

Ami kiszámítható: $P(\text{adatok}|\mathrm{H}_0) = ?$

A nullhipotézis elfogadása azt jelenti, hogy az adatok nem mondanak ellent a nullhipotézisnek (vagy mert igaz, vagy kevés az információ) → nem interpretálható „akár igaz is lehet”.

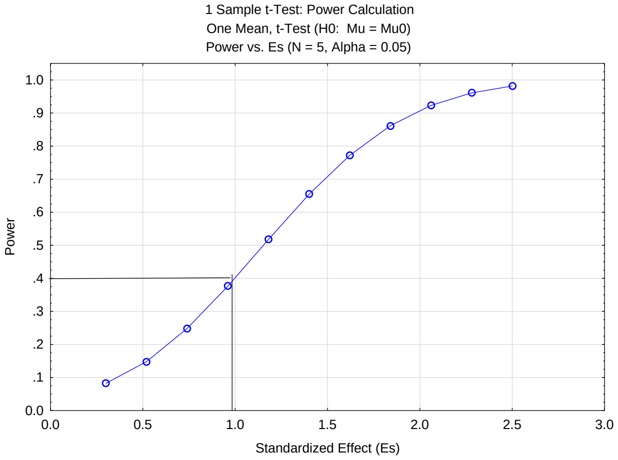
	döntés	
	a H_0 hipotézist	
	elfogadjuk	elutasítjuk
H_0 igaz	helyes döntés	elsőfajú hiba (α)
H_0 hamis	másodfajú hiba (β)	helyes döntés

Általában az elsőfajú hiba α valószínűségét rögzítjük úgy, hogy ha igaz, kis valószínűséggel utasítsuk el (szignifikanciaszint).

A másik, ami érdekel, a másodfajú hiba β valószínűsége (elfogadjuk a nullhipotézist, pedig nem igaz).

	döntés	
	a H_0 hipotézist	
	elfogadjuk	elutasítjuk
H_0 igaz	helyes döntés	elsőfajú hiba (α)
H_0 hamis	másodfajú hiba (β)	helyes döntés

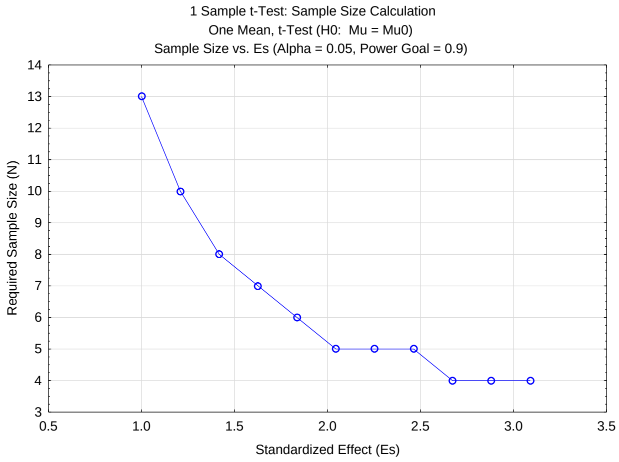
A próba 1-β ereje (elutasítjuk a nullhipotézist, amikor nem igaz, vagyis észrevesszük, hogy nem igaz)



$$Es = \frac{\mu_1 - \mu_0}{\sigma}$$

$\mu_1 - \mu_0 = \sigma$ nagyságú torzítást csak ~0.4 valószínűséggel veszünk észre 5 elemű mintából

α és β is rögzíthető → n (a szükséges mintaelemszám)



$$Es = \frac{\mu_1 - \mu_0}{\enclose{circle}{\sigma}}$$

„apró” probléma

13 ismétlés kell ahhoz, hogy a (feltételezett) σ szórással egyenlő mértékű torzítást 0.9 valószínűséggel kimutassuk ☹.

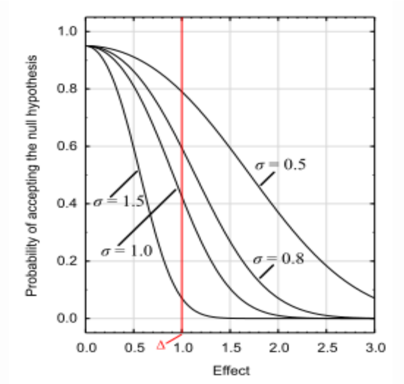


Fig. 4: Effect of σ - the value of β is clearly different at the same Δ (n = 7)

Ha σ nem ismert (márpedig …), csak hisszük, hogy β annyi!

Ismert, hogy a próba egyenértékű a konfidencia-intervallum alapján való döntéssel.

elfogadási tartomány

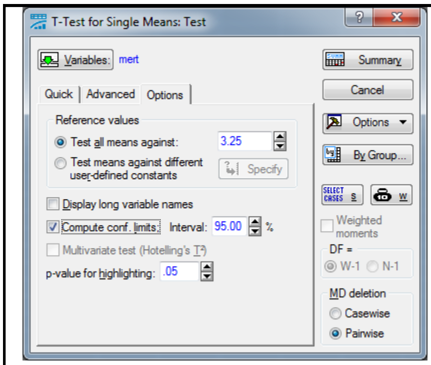
$$P\left(-t_{a/2} < \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \le t_{a/2} \right) = 1-\alpha$$

$$P\left(\bar{x} - t_{a/2}\, s/\sqrt{n} < \mu_0 \le \bar{x} + t_{a/2}\, s/\sqrt{n}\right) = 1-\alpha$$

konfidencia-intervallum

$$P\left(\bar{x} - t_{a/2}\, s/\sqrt{n} < \mu \le \bar{x} + t_{a/2}\, s/\sqrt{n}\right) = 1-\alpha$$

A nullhipotézist akkor fogadjuk el, ha a nullhipotézis szerinti μ_0 várható érték benne van a konfidencia-intervallumban.



T-Test for Single Means: Test

Variables: mert

Reference values

● Test all means against: 3.25

○ Test means against different user-defined constants

☐ Display long variable names

☑ Compute conf. limits; Interval: 95.00 %

☐ Multivariate test (Hotelling's T²)

p-value for highlighting: .05

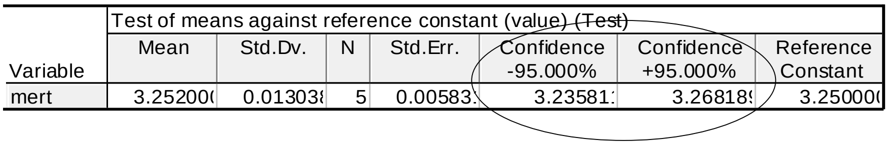
☐ Weighted moments

DF = ● W-1 ○ N-1

MD deletion

○ Casewise

● Pairwise

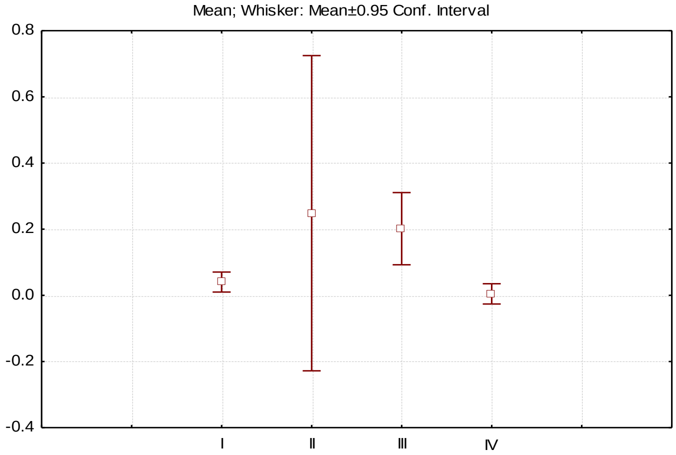


Variable	Test of means against reference constant (value) (Test)						
	Mean	Std.Dv.	N	Std.Err.	Confidence -95.000%	Confidence +95.000%	Reference Constant
mert	3.25200	0.01303	5	0.00583	3.23581	3.26818	3.25000

A 95%-os konfidencia-intervallum tartalmazza a nullhipotézis szerinti értéket (3.25), tehát elfogadjuk.

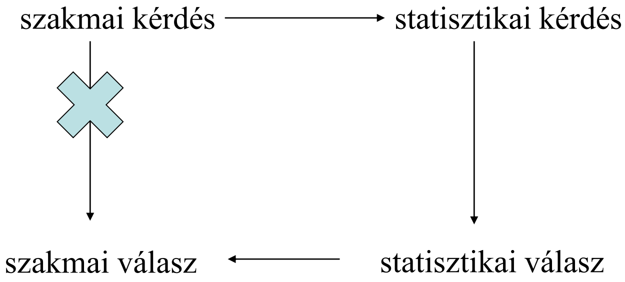
Ha két módszer van (próba és konfidencia-intervallum alapján), melyiket válasszuk?

A konfidencia-intervallum többet mond, mint a *p* érték:



$$H_0 : \mu = \mu_0 = 0$$

Probléma-megoldási vázlat a 6 szigma tréningeken



szakmai kérdés → statisztikai kérdés

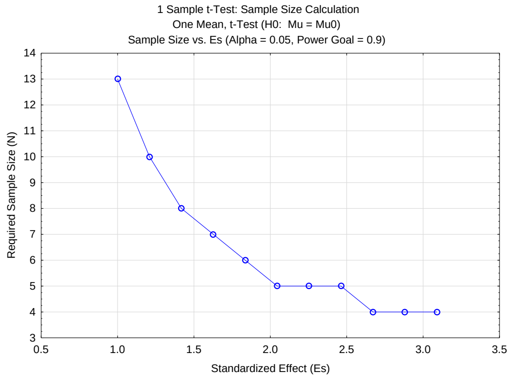
szakmai válasz ← statisztikai válasz

szakmai probléma: torzítatlan a mérés?
szakmai válasz: ????
statisztikai kérdés: $H_0 : \mu = \mu_0 = 3.25$

- csak a legegyszerűbb esetekben működik így.
- maga a szakmai kérdés sem fogalmazható meg egyszerűen, és szintén nem egyszerű azt statisztikai kérdéssé alakítani.
- együttműködésben alakul ki a releváns szakmai kérdés, egymás szakmáját „megtanulva”, iterálva alakítják át a megfelelő statisztikai kérdéssé.

Most ott tartunk, hogy a „Null Hypothesis Significance Testing” (NHST) módszer (amit használni szoktunk a döntésre) lényegében használhatatlan.

α és β rögzítésével → n (a szükséges mintaelemszám), pl. 13



$$Es = \frac{\mu_1 - \mu_0}{\sigma}$$

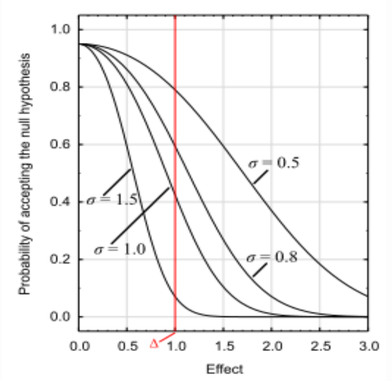
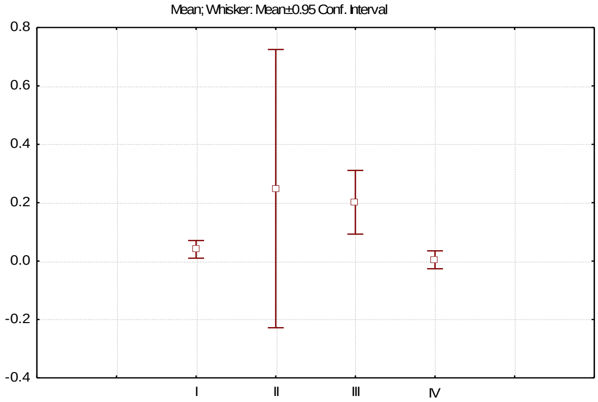


Fig. 4: Effect of σ - the value of β is clearly different at the same Δ (n = 7)

Ha σ nem ismert (márpedig …), csak hisszük, hogy β annyi!

A konfidencia-intervallumos szemlélet használható, de a formalizálás nehézkes.



Sokszor a statisztikai elemzésnél arra koncentrálunk, hogy megfelelő választ találjunk a föltett kérdésre. Legalább ilyen fontos pedig, hogy a megfelelő kérdést tegyük föl.

$H_0 : \mu = \mu_0$ Tényleg ezt kérdezzük?

pl. hogy az 3.25%-os ismert koncentrációjú anyagra mért értékek tényleg 3.25% (3.2500000…) körül ingadoznak-e? (ez jelentené a torzítatlanságot).

t_0

$$-t_{a/2} < \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \leq t_{a/2}$$

Minél bizonytalanabb a mérés (*s* nagy), és minél kevesebbet mérünk (*n* kicsi), annál közelebb van a t_0 próbastatisztika zérushoz, tehát annál biztosabban elfogadjuk, hogy torzítatlan ☺.

Minél jobb a mérés, annál kevésbé fogadjuk el ☹ .

$H_0 : \mu = \mu_0$ Tényleg ezt kérdezzük?

pl. hogy a 3.25 %-os ismert koncentrációjú anyagra mért értékek tényleg 3.2500000… % körül ingadoznak-e? (ez jelentené a torzítatlanságot). Az ettől való szignifikáns eltérés a statisztikai szignifikancia.
Inkább ezt kérdezzük (intervallum-hipotézis)

$$\mu_0 - \Delta < \mu < \mu_0 + \Delta$$

Biztosak lehetünk-e abban, hogy az eltérés nem halad meg egy megengedett Δ mértéket? Az ettől való szignifikáns eltérés a szakmai szignifikancia.

A megengedett Δ mérték szakmai adat, nem statisztikai. Ha például a koncentrációra a tűrésmező ±0.1%, Δ=0.01% biztosan (realisztikusan) megengedhető torzítás.

$\mu_0 - \Delta < \mu < \mu_0 + \Delta$ intervallum-hipotézis (két ellenhipotézis)

$\mathrm{H}_{0L} : \mu \le \mu_0 - \Delta$ $\mathrm{H}_{1L} : \mu > \mu_0 - \Delta$ alsó

Ezt szeretnénk bizonyítva látni (a módszer torzítatlan, a várható érték legalább egy hajszállal meghaladja az alsó határt)

$\mathrm{H}_{0U} : \mu \ge \mu_0 + \Delta$ $\mathrm{H}_{1U} : \mu < \mu_0 + \Delta$ fölső

Ezt szeretnénk bizonyítva látni (a módszer torzítatlan, a várható érték legalább egy hajszállal a fölső határ alatt van)

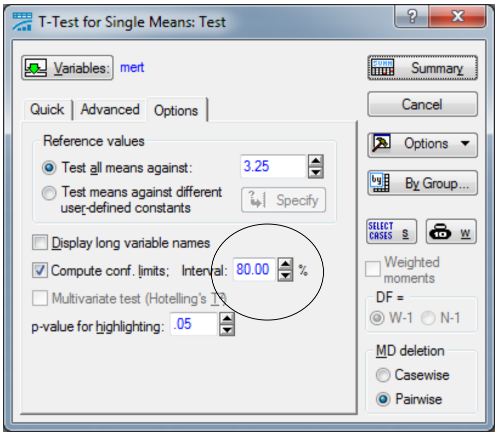
Ezt is lefordíthatjuk konfidencia-intervallumra.

A várható érték 1-2β valószínűségű konfidencia-intervalluma:

$$P\left(\bar{x} - t_\beta\, s/\sqrt{n} < \mu < \bar{x} + t_\beta\, s/\sqrt{n}\right) = 1 - 2\beta$$

$$-\Delta < \bar{x} - t_\beta\, s/\sqrt{n} < \mu < \bar{x} + t_\beta\, s/\sqrt{n} < \Delta$$

A szakmai hipotézist (a módszer torzítatlan, a várható érték az elfogadható tartományban van) elfogadjuk, ha a várható érték 1-2β valószínűségű konfidencia-intervalluma a megengedett (μ_0-Δ, μ_0+Δ) tartományon belül van.
β rögzített.



Variables: mert

Reference values

● Test all means against: 3.25

○ Test means against different user-defined constants

☐ Display long variable names

☑ Compute conf. limits; Interval: 80.00 %

☐ Multivariate test (Hotelling's T²)

p-value for highlighting: .05

☐ Weighted moments

DF = ● W-1 ○ N-1

MD deletion

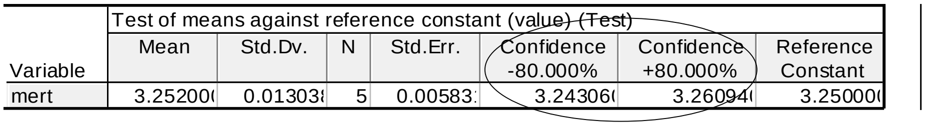
○ Casewise

● Pairwise

ha pl. Δ=0.01% (ez nem az adatokból következik)

$$3.24 < 3.243 < \mu < 3.261 < 3.26$$

csak majdnem



	Test of means against reference constant (value) (Test)						
Variable	Mean	Std.Dv.	N	Std.Err.	Confidence -80.000%	Confidence +80.000%	Reference Constant
mert	3.25200	0.01303	5	0.00583	3.24306	3.26094	3.25000

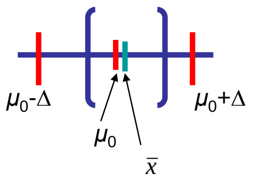
Két egyoldali t-próba (two one-sided t-test, TOST), β rögzített.

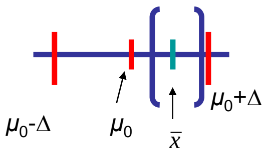
D.J. Schuirmann: Journal of Pharmacokinetics and Biopharmaceutics, 15 657-680 (1987) (+előzmények)

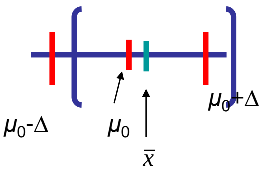
Ez a megközelítés a vevőt/fogyasztót védi: ha nem tudjuk bizonyítani, hogy a tűrésmezőn belül van a jellemző, nem nyilváníthatjuk jónak.

TOST szerint:

elfogadjuk elfogadjuk elutasítjuk







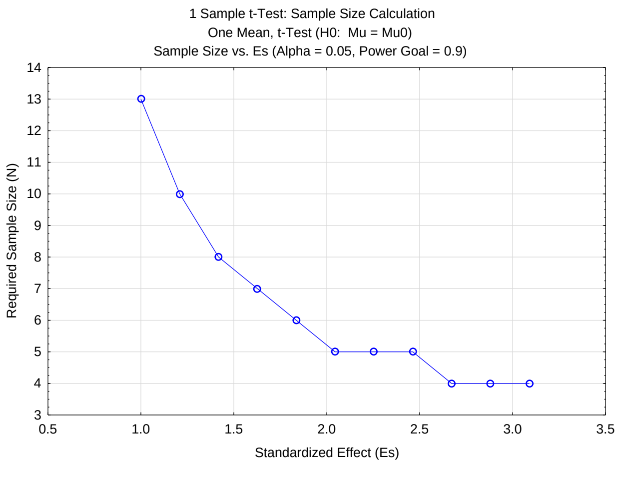
hagyományos szerint:

elfogadjuk elutasítjuk elfogadjuk

A hagyományos hipotézisvizsgálat (NHST) és a TOST viszonya

- a hagyományos hipotézisvizsgálatnál α értékét rögzítjük (kicsi legyen a valószínűsége, hogy elutasítsuk, amikor igaz)
- a TOST-nál β értékét rögzítjük (kicsi legyen a valószínűsége, hogy elfogadjuk, amikor nem igaz).

Lehet mindkettőt, de ...



$$Es = \frac{\mu_1 - \mu_0}{\sigma}$$

13 ismétlés kell ahhoz, hogy a (feltételezett) σ szórással egyenlő mértékű torzítást 0.9 valószínűséggel kimutassuk.

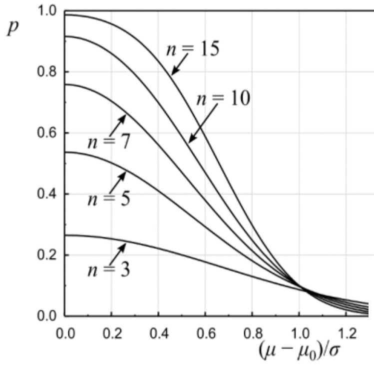
Ha nem vagyunk hajlandók a mintaelemszámot megfelelően megválasztani, vagy csak α, vagy csak β rögzíthető.

A hagyományos hipotézisvizsgálatnál α értékét rögzítjük (kicsi legyen a valószínűsége, hogy elutasítsuk, amikor igaz).

A TOST-nál β értékét rögzítjük (kicsi legyen a valószínűsége, hogy elfogadjuk, amikor nem igaz).

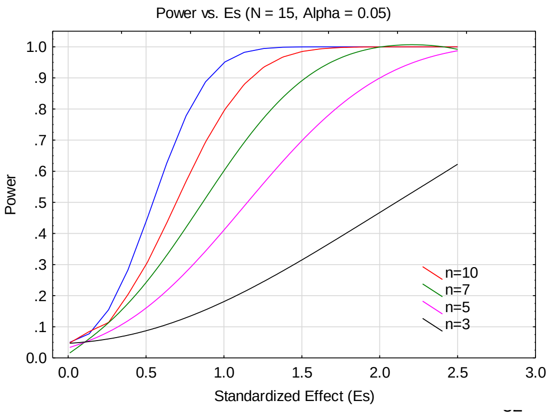
Tehát a TOST-nál inkább a fogyasztót védjük.

Ez az ellenkezője az ártatlanság vélelmének!



Az elfogadás („bizonyítva látjuk”) valószínűsége a TOST-nál

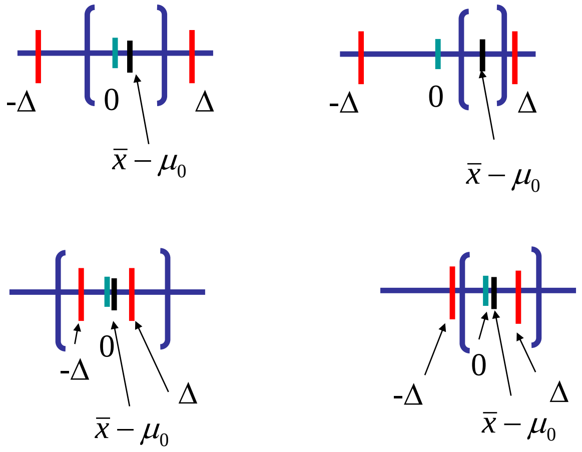
Az eltérés kimutatásának valószínűsége a hagyományos próbánál



a hagyományos ($H_0 : \mu - \mu_0$) szerint
elfogadjuk elutasítjuk
nem mondhatjuk,
hogy különböznek szignifikánsan különböznek

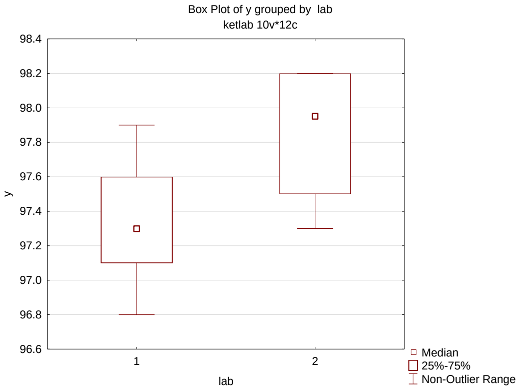
TOST szerint
elfogadjuk
mondhatjuk, hogy
az eltérés <Δ

TOST szerint
elutasítjuk
nem mondhatjuk,
hogy az eltérés
<Δ



2. példa
Két laboratórium összehasonlítása (módszerátadás)

	1. lab	2. lab
1	97.1	97.3
2	97.2	98.2
3	97.9	98.2
4	97.4	98.2
5	97.6	97.5
6	96.8	97.7
$\bar{x}$	97.33	97.85
s	0.388	0.404



	1. lab	2. lab
1	97.1	97.3
2	97.2	98.2
3	97.9	98.2
4	97.4	98.2
5	97.6	97.5
6	96.8	97.7
$\bar{x}$	97.33	97.85
s	0.388	0.404

Mellékfeltétel: a két sokaság varianciája megegyezik (*F*-próba!)?

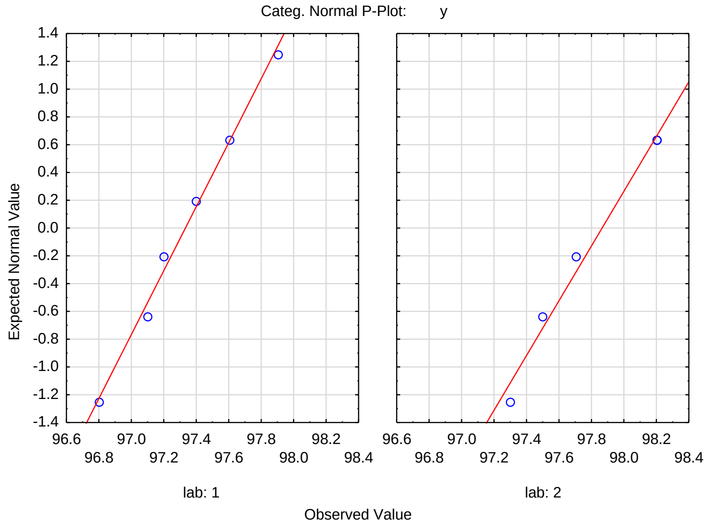
$$F_0 = \frac{s_1^2}{s_2^2} = \frac{0.388^2}{0.404^2} = 1.08$$

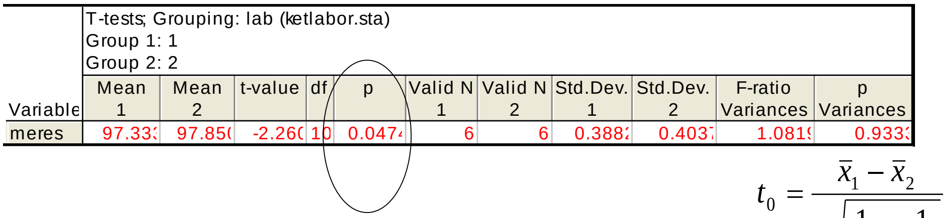
$$F_{0.05}(5,5) = 5.05$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{97.33 - 97.85}{0.396\sqrt{\frac{1}{6} + \frac{1}{6}}} = -2.260$$

$$t_{0.05/2}(10) = 2.228$$

Szignifikáns a különbség, de $|\bar{x}_1 - \bar{x}_2| = 0.52$ kicsi!





T-tests; Grouping: lab (ketlabor.sta)
Group 1: 1
Group 2: 2

Variable	Mean 1	Mean 2	t-value	df	p	Valid N 1	Valid N 2	Std.Dev. 1	Std.Dev. 2	F-ratio Variances	p Variances
meres	97.33	97.85	-2.26	10	0.047	6	6	0.388	0.403	1.081	0.933

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Kis szórás (vagy nagyon sok mérés) esetén statisztikailag szignifikánsnak találjuk a két labor közötti eltérést, holott az szakmailag nem jelentős. ☹

Nagy szórás (vagy kevés mérés) esetén könnyen elfogadjuk, hogy μ_1=μ_2, „a rosszul dolgozó analitikust jutalmazza”. ☺ ☹

.
Releváns kérdés: a két labor (várható értéke) közötti eltérés nem nagyobb-e a megengedettnél/kisebb-e a megengedettnél; nagy biztonsággal vegyük észre, ha nagyobb.

Mit is jelent az, hogy elfogadjuk a nullhipotézist?

$\mathrm{H}_0 : \mu_1 - \mu_2 = 0$

$\mathrm{H}_1 : \mu_1 - \mu_2 \neq 0$

Az adatok nem mondanak ellent. „Akár igaz is lehet”
Ezt kérdezzük? Nem!

intervallum-hipotézis: azt állítjuk, hogy a különbség nem haladja meg a megengedettet (TOST).

$$-\Delta < \mu_1 - \mu_2 < \Delta$$

$$-\Delta < \mu_1 - \mu_2 < \Delta$$

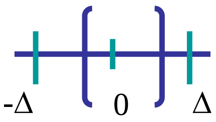
Δ nem a statisztikából következik, hanem a szakmai kérdésből!

Pl. az analízis-eredményeket arra akarjuk használni, hogy elbíráljuk, hogy a hatóanyagtartalom a névleges ±5%-án belül van-e. Ehhez elég lehet, ha a 100 körüli tartalom 1%-a alatt van a torzítás, tehát

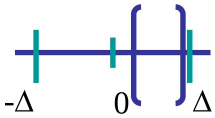
$$\Delta = 1\%$$

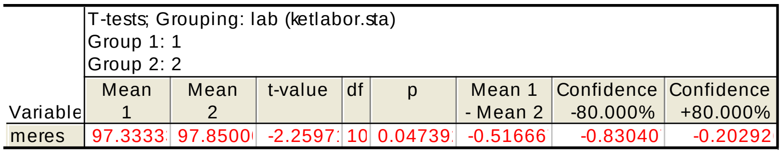
TOST szerint elfogadjuk, ha

$$\bar{x}_1 - \bar{x}_2 + t_\beta s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \Delta \qquad \bar{x}_1 - \bar{x}_2 - t_\beta s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} > -\Delta$$



±Δ tartalmazza a konfidencia-intervallumot





	T-tests; Grouping: lab (ketlabor.sta) Group 1: 1 Group 2: 2							
Variable	Mean 1	Mean 2	t-value	df	p	Mean 1 - Mean 2	Confidence -80.000%	Confidence +80.000%
meres	97.3333	97.8500	-2.2597	10	0.04739	-0.51666	-0.83040	-0.20292

Ha $\Delta = 1\%$ elfogadjuk

a hagyományos ($H_0 : \mu_1 - \mu_2 = 0$) szerint

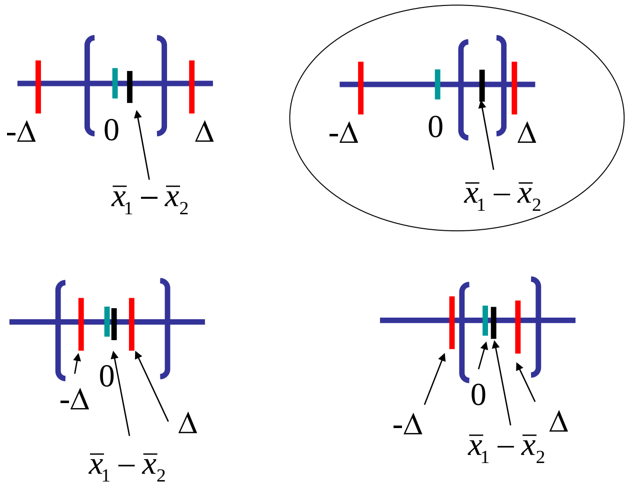
elfogadjuk elutasítjuk

nem mondhatjuk,
hogy különböznek

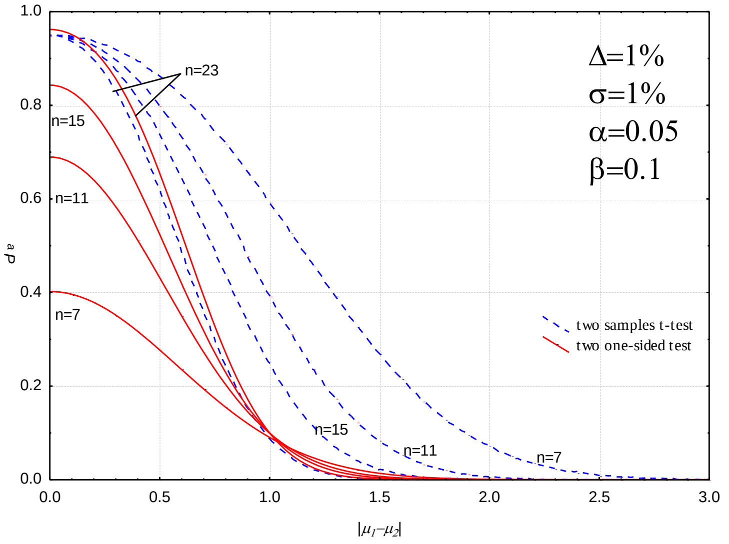
szignifikánsan különböznek

TOST szerint
elfogadjuk
mondhatjuk, hogy
az eltérés <Δ

TOST szerint
elutasítjuk
nem mondhatjuk,
hogy az eltérés
<Δ



módszerátadás $H_0 : \mu_1 - \mu_2 = 0$ $H_0' : -\Delta \le \mu_1 - \mu_2 \le \Delta$



Mi is az a paradigmaváltás, amiről beszéltem?

Hagyományos hipotézisvizsgálat (NHST):

$H_0 : \mu = \mu_0$

Ha p>0.05, elfogadjuk 5%-os szignifikanciaszinten („akár igaz is lehet”)

TOST

$\mu_0 - \Delta < \mu < \mu_0 + \Delta$

A két különböző kérdésre különböző lehet a válasz.
Melyik a releváns kérdés?

---

a hagyományos ($H_0 : \mu - \mu_0$) szerint

elfogadjuk
nem mondhatjuk,
hogy különböznek

elutasítjuk
szignifikánsan különböznek

TOST szerint
elfogadjuk
mondhatjuk, hogy
az eltérés <Δ

TOST szerint
elutasítjuk
nem mondhatjuk,
hogy az eltérés
<Δ