A 3.2.5 pont példájának adatai

C1	C2	C3	C4
0.00000	0.00000	0.00000	0.00000
0.00000	0.11916	0.00131	0.14327
0.11458	0.27336	0.00785	0.34957
0.15625	0.31308	0.00916	0.43840
0.19792	0.32243	0.01440	0.44699
0.30208	0.32710	0.01832	0.44986
0.31250	0.35280	0.03010	0.51576
0.34375	0.37850	0.04450	0.61032
0.35417	0.41121	0.04581	0.63897
0.40625	0.42056	0.07068	0.65330
0.42708	0.42056	0.07592	0.68481
0.46875	0.42523	0.09031	0.72493
0.47917	0.45093	0.09162	0.73926
0.56250	0.45561	0.10864	0.78223
0.58333	0.47196	0.14136	0.82521
0.62500	0.57944	0.14660	0.86819
0.66667	0.60047	0.14921	0.88539
0.75000	0.62617	0.19372	0.89685
0.79167	0.64019	0.19503	0.89685
0.82292	0.66122	0.19634	0.94842
0.83333	0.70093	0.19634	0.95415
0.92708	0.89019	0.29450	0.98567
0.98958	0.95093	0.64398	0.99140
1.00000	1.00000	1.00000	1.00000

4. Statisztikai jellemzõk megbízhatósága

4.1 Konfidencia tartomány, konfidencia szint

A mintákból meghatározott becslõk magukban is érdekesek lehetnek, különösen, ha minták összehasonlításáról van szó. Természetes azonban, hogy a jellemzõk akkor értékesek igazán, ha azok megbízhatóságáról is van képünk. Ez a kívánalom egyenértékü azzal, hogy többé kevésbé ismerjük a becslõ statisztikák eloszlását, de legalábbis alkalmazhatunk néhány valószínûségszámításból ismert egyenlõtlenséget. Emlékeztetünk arra, hogy a *mintákból számított becslések valószínûségi változók függvényei lévén maguk is valószínûségi változók*.

A gyakorlatban a kérdések általában így vetõdnek fel:

(a) mi a valószínûsége annak, hogy a valószínûségi változó egy realizációja (= a következõ megfigyelt adat) *elõírt* határok közé essék (pl *a £ x £ b*)?

(b) melyek azok a határok, amelyek közé a következõ megfigyelt adat *elõírt* valószínûséggel esik?

A két kérdés lényegében ugyanaz, egyik feladat a másik inverze. Az (a) kérdéssel egy-egy megfigyelést értékelünk, a (b) kérdéssel követelményeket fogalmazunk meg, pl. pontosságot írunk elõ.

Ha ismerjük a szóbanforgó valószínûségi változó eloszlásfüggvényét, mindkét kérdésre választ kaphatunk:

$$p(a \le x < b) = F(b) - F(a) = \int_a^b f(x)dx \qquad (4.1)$$

A továbbiakban az általánosság kedvéért folytonos valószínûségi változók esetére mutatjuk be a megoldások gondolatmenetét.

4.2 Nevezetes eloszlások

Természettudományos gyakorlatunkban egyik leggyakoribb eloszlásfüggvény a *normális eloszlás*. Ha valamely vizsgált változóra számos, önmagában kis hatású, a változó értékét egyforma eséllyel növelõ vagy csökkentõ tényezõ is hat, számíthatunk arra, hogy megfigyelt értéke normális eloszlású lesz.

4.2.1 A normális eloszlás

A (4.1) integrál ezesetben:

$$p(a \le x < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \qquad (4.2)$$

ahol μ az *x* változó várható értéke, σ pedig annak szórása. (4.2) függvénynek nincs analitikusan megadható integrálja, értékeit numerikusan számítják ki az

$$u = \frac{x - m}{s}$$

standardizált, 0-közepû és 1 szórású változóra, -∞ és x határok között. (Ezt az eloszlást szokás N(0,1) röviditéssel jelölni). Miután a

$$\Phi(x) = \frac{1}{\sqrt{2p}} \int_{-\infty}^{x} \exp\left(-u^2 / 2\right) du \qquad (4.3)$$

függvény szimmetrikus, táblázatokban csak az eloszlás (második) felét adják meg, 0 és + ∞ határok között, ahol a $\Phi(x)$ valószínûség 0.5.tõl 1-ig nõ.

Negativ *x* argumentumok esetén a valószínûséget

$$\int_{-\infty}^{-x} f(u)du = 1 - \int_{-\infty}^{x} f(u)du = 1 - \Phi(x) \qquad (4.4)$$

módon kell keresni. Ha arra vagyunk kiváncsiak, mi annak valószínûsége, hogy *x* a *-b* és *+b* határok között lép fel, a táblázat *b* argumentumához tartozó érték kétszeresébõl ki kell vonnunk 1-et. Ugyanis:

$$\int_{-b}^{b} f(u)du = \int_{-\infty}^{b} f(u)du - \int_{-\infty}^{-b} f(u)du = \int_{-\infty}^{b} f(u)du - \left(1 - \int_{-\infty}^{b} f(u)du\right) = 2\int_{-\infty}^{b} f(u)du - 1 =$$
$$= 2\Phi(b) - 1 \qquad (4.5)$$

Érdemes megjegyezni, hogy normális eloszlás esetén

p$(-1 \le x \le 1)$	$= 2\Phi(1) - 1 =$	0.6826
p$(-2 \le x \le 2)$	$= 2\Phi(2) - 1 =$	0.9545
p$(-1.96 \le x \le 1.96)$	$= 2\Phi(1.96) - 1 =$	0.9500
p$(-3 \le x \le 3)$	$=2\Phi(3) - 1 =$	0.9973

illetve *nem standardizált* változóra:

p$((m - s) \le x \le (m + s))$	= 0.6826
p$((m - 2s) \le x \le (m + 2s))$	= 0.9545
p$(m - 1.96s) \le x \le m + 1.96s)$	= 0.9500
p$((m - 3s) \le x \le (m + 3s))$	= 0.9973

Azt a tartományt amelybe a valószínûségi változó várhatóan p valószínûséggel esik, a változó p *szintû megbízhatósági* vagy *konfidencia tartományának* nevezik. A változó természetesen α = 1 - p valószínûséggel a konfidencia tartományon kívül is realizálódhat. Ezt az α értéket *tévedési valószínûségnek* szokás nevezni. A konfidencia tartományt gyakran α szintû tartománynak is nevezik.

A bevezetésben feltett (b) kérdés, azaz az, hogy megkívánt, rendszerint kerek konfidencia szinthez milyen $\pm k\sigma$ konfidencia határok tartoznak, alkalmasan átrendezett táblázatokkal válaszolható meg.

p	$\alpha = 1$ - p	k
0.99	0.01	2.5758
0.95	0.05	1.9600
0.90	0.1	1.6449
0.70	0.3	1.0364

A normális eloszlás -és a továbbiakban tárgyalt Student és χ^2 eloszlások számértékeit kézikönyvekben vagy pl. a http://math.uc.edu/~brycw/148/tables.htm internetcímen lehet megtalálni.

4.2.2 A Student eloszlás

Ha egy normális eloszlású sokaságból vett minta sok elemû (n > 120), akkor a mintából számított *s standard deviáció* jól becsüli az elméleti szórást, σ-t. Ha azonban nem ez a helyzet, a kevesebb elemû mintából becsült *s* standard deviációval *szélesebb* konfidencia tartományt kell megadnunk ahhoz, hogy biztonságunk megmaradjon.

A helyes összefüggéseket ezekben az esetekben a normális eloszlás helyett a *Student eloszlás* adja meg, amelynél a konfidencia tartományok szélességét megadó *t* szorzók a minta elemszámától, pontosabban a minta *szabadsági fokától* függenek.

A szintén szimmetrikus

$$F(\nu,t)=\frac{1}{\sqrt{\nu\pi}}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\int_{-\infty}^{t}\frac{du}{\left(1+\frac{u^2}{2}\right)^{\frac{\nu+1}{2}}} \qquad (4.6)$$

Student eloszlás szintén táblázatoltan található. Leghasználatosabbak azok a táblázatok, amelyekkel az α tévedési valószínüséghez és a ν szabadsági fokhoz tartozó konfidencia tartomány határai kereshetõk ki. (Minta. 4.1 táblázat)

4.1 táblázat. Student eloszlás *t* értékei, különbözõ mérésszámnál

	a = 1 - p	*T*		
		n = 3	*n*=15	*n* = ∞
0.99	0.01	9,925	2.977	2.5758
0.95	0.05	4.303	2.145	1.96
0.90	0.10	2.92	1.761	1.6449
0.70	0.30	1.386	1.076	1.0364

4.2.3 A χ^2 eloszlás

Mivel a valószínûségi változók négyzetei (pl. a szórás négyzete, a *variancia*) gyakorlatunkban igen jelentõsek, fontos szerepû az a függvény, amelyik a független, külön-külön N(0,1) eloszlású változók

$$\chi^2 = X_1^2 + X_2^2 + \ldots + X_\nu^2 \qquad (4.7)$$

összegének eloszlását adja meg, a χ^2 *eloszlásfüggvény*:

$$F(\nu,,x) = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^x u^{\frac{\nu-1}{2}} e^{-\frac{u}{2}} du \qquad (4.7)$$

ahol ν a szabadsági fok, a független valószínûségi változók száma. A függvény láthatóan két változótól függõen adja meg azt, mi a valószínûsége annak, hogy a változók négyzetösszege x-nél kisebb.

4.2.4 Az F eloszlás

A normális és Student eloszlást sikeresen alkalmazzák normális változók *különbségeinek* vizsgálatára. Valószínûségi változók négyzetösszegei esetén hasznosabbnak bizonyult azok *hányadosainak* kritikus megítélése.

Erre a feladatra (nevezetesen annak eldöntésére,.hogy egyezõnek vagy eltérõnek tekinthetõ-e két változó négyzetösszege) az Fisher féle F *eloszlás* alkalmas. Ez a függvény két független, χ^2 *eloszlású* változó hányadosának eloszlásáról tájékoztat. Az F függvény az

$$F = \frac{\sum_{i=1}^{\nu 1} X_i^2 / \nu_1}{\sum_{i=1}^{\nu 2} Y_i^2 / \nu_2} \qquad (4.8)$$

hányados adott határok közötti elõfordulási valószínûségét adja meg, ahol ν_1 és ν_2 a számláló és nevezõ szabadsági foka. *F* számlálójában és nevezõjében varianciákat ismerhetünk fel. Az F eloszlás 0 és + ∞ között értelmezett. Ebbõl következõen az F törtben a számlálónak kell kisebbnek lennie.

A gyakorlatban használt F táblázatokban a választott α tévedési valószínûségnek, továbbá a számláló és a nevezõ szabadsági fokának ismeretében lehet megtalálni azt a kritikus $\hat{F}$ értéket, amelynél egyezõnek feltételezett változók esetén a kisérletileg megkapott F érték nem lehet nagyobb.

5. Statisztikai hipotézisek, statisztikai döntések.

5.1 Alapelvek

Az olyan becslések, mint a *középérték*, a *szórás* valószínûségi változók, amelyeknek megvan a maguk *eloszlása*, *várható értéke*, *szórása*. Ha ez így van, feltehetõk olyan kérdések, hogy *két becslés véletlenül tér-e el egymástól* vagy az eltérésnek *jelentõs* oka van? Más szavakkal fogalmazva kérdezzük:két becslés *ugyanahhoz a sokasághoz tartozik-e*, azaz, ha számértékük eltér, akkor ez annak tulajdonítható-e, hogy más sokasághoz tartoznak, vagy csak a véletlennek? Ezekre a kérdésekre válaszolnak a *statisztikai próbák*.

A válaszadás gondolatmenete ez:

(a) meg kell határozni két összehasonlítandó érték eltérését (különbségét).

(b) ha ismerjük a vizsgált valószínûségi változók eloszlását, akkor a két érték különbségérõl eldönthetõ, hogy mekkora annak fellépési valószínûsége. Az eltéréseket a különbség szórásához viszonyítjuk, azt vizsgáljuk, nagyobb-e az eltérés ennél a szórásnál, vagy annak két-, háromszorosánál.

(c) ha úgy itéljük, hogy ez a valószínûség kicsiny, akkor az eltérést *nem a véletlennek* tulajdonítjuk és a két értéket *jelentõsen, szignifikánsan* eltérõnek nyilvánítjuk, tudva azt, hogy tévedhetünk is. A kis valószínûség szokásosan az α tévedési valószínûséggel egyezik.

Hogy az eltérés mekkora valószínûségét tekintjük majd “kicsiny”-nek (mekkora tévedési valószínûséget vállalunk), az a feladat körülményeitõl függõ, *elõzetes* elhatározás kérdése. Belátható, hogy a választást a próba elött illik megejteni.

A gondolatmenetet és a használt szakkifejezéseket szemléljük meg egy példán. Tegyük fel, hogy arra vagyunk kiváncsiak, egy valószínûségi változó konkrét *x* értéke beletartozik-e egy, általunk ismert μ középértékû és σ szórású normális sokaságba vagy nem? Más szavakkal arra, hogy a μ -*x* különbség beesik-e

$$(-3\sigma \le x - \mu < +3\sigma)$$

tartományba? *Ha nem*, akkor egy 0.27% valószínûségû esemény következett be. Ha ezt kicsinynek ítéljük, akkor azt mondjuk, *x* nem tartozik a sokasághoz és ebben 0.27%

valószínûséggel tévedhetünk, hiszen elvben végtelen nagy vagy kicsi elem is lehetne a sokaság eleme.

Formálisan ezt tesszük: A

$$p(\mu - u\sigma \le x < \mu + u\sigma) \qquad (5.1)$$

valószínûség nem változik, ha a zárójeleken belüli eseményt leíró egyenlötlenséget szabályosan átalakítjuk:

$$(\mu - u\sigma \le x < \mu + u\sigma) =$$

$$(- u\sigma \le x - \mu < + u\sigma) =$$

$$(|x - \mu| \le u\sigma)$$

$$\frac{|x - m|}{s} \le u$$

Nos, ha az egyenlõtlenség baloldala, amit $\hat{u}$ módon is szoktak jelölni nagyobb, mint három, akkor a

$$p\left(\frac{|x - \nu|}{\sigma} > 3\right) \qquad (5.2)$$

valószínûség a normális eloszlás táblázata szerint 100 - 99.73 = 0.27%, ami kicsiny valószínûség, ezért *x*-et nem tartjuk, a sokaság elemének.

5.2 Statisztikai hipotézisek

Mivel statisztikai vizsgálattal az igazságot abszolut bizonyossággal nem sikerül megállapítani, az állításokat *hipotéziseknek* nevezzük, és nem azt mondjuk róluk, hogy igazak, vagy hamisak, hanem azt, hogy *elfogadjuk*-e õket, vagy *elvetjük*.

5.2.1 Nullahipotézis és alternatív (ellen)hipotézis

Feltevésünk általában az, hogy a vizsgált becslések *megegyeznek*, azaz *különbségük* 0. Innen a nullahipotézis elnevezés és a (H_0) jelölés

$$H_0: \mu_1 = \mu_2 \qquad (5.3)$$

A nullahipotézissel szemben alternatív hipotézist (H_A) szokás felállítani, amely lehet a nullahipotézis ellentéte, de nem szükségképpen az.

$$H_A: \mu_1 \ne \mu_2 \qquad (5.4)$$

vagy például

$$H_A: \mu_1 > \mu_2 \qquad (5.5)$$

5.2.2 Egyoldalas és kétoldalas hipotézisek

Az (5.3) ("egyenlõ") hipotézissel szembeállított (5.4) ("nem egyenlõ") és (5.5) ("nagyobb") hipotéziseket meg kell különböztetnünk! Az elsõ esetben elvetjük a hipotézist akkor is, ha a μ_1 - μ_2 különbség túl nagy *negatív,* és akkor is, ha ha túl nagy *pozitív* szám. Ha 5% tévedési valószínûséget választottunk, 2.5% valószínûséget kell adni annak, hogy a különbség a "haranggörbe" egyik végére, 2.5%-ot annak, hogy a másik végére essék. A kérdésfeltevést ezért is nevezik "kétoldalas" (two sided) feltevésnek. Ha viszont a H_A: $\mu_1 > \mu_2$ alternatív hipotézissel foglalkozunk, csak az a határ érdekel, amelynél μ_1 5% valószínûséggel nagyobb. mint μ_2. (Egyoldalas, one sided kérdésfeltevés.) Más szavakkal: ha kétoldalas a feltevés, azokat az *u* határokat figyeljük, amelyek a sûrûségfüggvény alatti 2.5% - 97.5% valószínûségû területet határolják, egyoldalas esetben pedig a -∞ - 95% valószínûségterületet. A gyakorlatban ez azt jelenti, hogy a kritikus *u* értékeket kétoldalas próbánál a 0.025 ($\alpha/2$), egyoldalasnál a 0.05 (α) oszlopban kell keresni.

5.2.3 Elsõfajú és másodfajú hibák

Statisztikai hipotézisek elfogadásánál vagy elvetésénél kétféle hibát lehet véteni: elsõfajú és másodfajú hibákat:

Egy igaz hipotézis elfogadása	nincs hiba
Egy igaz hipotézis elvetése	elsõfajú, vagy α hiba.-
Egy hamis hipotézis elvetése	nincs hiba
Egy hamis hipotézis elfogadása	másodfajú vagy β hiba.

A kétféle hiba jelentõségét csak az adott helyzetben lehet mérlegelni. A körülmények döntik el, hogy mi okoz nagyobb kárt: egy jobb növényvédõszer elvetése, vagy egy rossz bevezetése, egy beteg kezelésének elhagyása, vagy egy egészséges megoperálása. Az elsõfajú hiba valószínûségét a tévedési valószínûség csökkentésével lehet kisebbiteni. A másodfajú hiba valószínûségének beállítása bonyolultabb kérdés.

5.3 Gyakori statisztikus próbák

A továbbiakban két gyakran használt példát mutatunk be. A példák több szempontból egyszerûek, de jó megjegyezni, hogy a matematikai statisztikának a gyakorlatban felvetõdõ nehezebb feladatokra (nem normális, vagy ismeretlen eloszlású adatok, különbözõ mérteû minták stb.) is számos megoldása van.

5.3.1 Két számtani közép egyezésének vizsgálata

Két mérési eredményt akarunk összehasonlítani. A mérési eredmények véges n_1 és n_2 párhuzamos mérés átlagai, számtani közepek, $\bar{x}_1$.és $\bar{x}_2$ értékek. Tudni szeretnénk, eltér-e egymástól a két eredmény. Egyszerûség kedvéért tételezzük fel, hogy a két eredményt ugyanannyi párhuzamos mért értékbõl számították, és azt is, hogy a mérési módszer pontossága a két mérés között nem változott. Tegyük fel továbbá, hogy a mért értékek normális eloszlásúak.

A nullahipotézis:

H_0: $\mu_1 = \mu_2$ Feltevés : $\sigma_1 = \sigma_2$ $n_1 = n_2$

Az ellenhipotézis:

H_A: $\mu_1 \neq \mu_2$

A nullahipotézisbõl következik, hogy a vizsgált valószínûségi változónk a $\mu_1 - \mu_2$ küllönbség. Kérdés, mi ennek a különbségnek a szórása? Tudjuk, hogy az $\bar{x}$ számtani közép varianciáját az s^2/n mennyiség becsli. A varianciák összeadhatóságából következik, hogy az $\bar{x}_1 - \bar{x}_2$ különbség szórása becslése: $\sqrt{s_1^2/n_1 + s_2^2/n_2}$, esetünkben: $\sqrt{(s_1^2 + s_2^2)/n}$. A szabadsági fok: 2*(*n*-1). Ismerve ezeket a mennyiséget

A számított *t*: $$\hat{t} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(s_1^2 + s_2^2\right)/n}}$$

Ezt a mennyiséget kell a táblázati kritikus $t(\alpha,\nu)$-értékkel összemérni.

5.1 Numerikus példa (L. Sachs: Statistische Methoden, Springer, Berlin 1993. p. 77)

Legyen $\bar{x}_1 = 42.76, \bar{x}_2 = 40.21$, $s_1^2 = 33.44, s_2^2 = 22.55$, $n = 30$

$$\hat{t} = \frac{42.76 - 40.21}{\sqrt{(33.44 + 22.55)/30}} = \frac{2.55}{1.3661} = 1.8666$$

t kritikus értéke 95%-os megbízhatósági szinten, 58 szabadsági foknál:

$$t(\alpha/2 = 0.025, \nu = 58) = 2.002$$

A μ_1 - μ_2 különbség konfidencia tartománya:

$$2.55 - 2.002 * 1.3661 \leq \mu1 - \mu2 < 2.55 + 2.002 * 1.3661$$

$$-0.1847 \leq \mu1 - \mu2 < 5.2847$$

A két középérték *nem* tér el egymástól szignifikánsan, H_0-t megtartjuk, a különbség konfidencia tartománya 95% valószínûséggel tartalmazza 0-t.

5.3.2. Tapasztalati szórások összehasonlítása

Mint errõl a 4.2.3 és 4.2.4 pontban már szó volt, valószínûségi változók négyzetei összegének összehasonlítására célszerûen nem különbségük, hanem *hányadosuk eloszlásfüggvénye* használtatik. Végesszámú mintákból becsült varianciák ilyen mennyiségek, a döntõ függvény az F-eloszlás. Ha a szórások négyzetének hányadosa meghalad egy bizonyos, α-tól függõ értéket, akkor a két variancia 1 - α biztonsággal eltér egymástól. Az F eloszlás két másik változója a számláló és nevezõ szabadsági foka.

A próba lépései a következõk: Legyen adott 2 minta. A minták elemszáma legyen n_1 és n_2. A két mintából meghatározunk két standard deviációt: s_1-et és s_2-t. Kérdés: *szignifikánsan eltér-e a két szórás*?

1) Fogalmazzuk meg a hipotéziseket:

$H_0 : \sigma_1^2 = \sigma_2^2$ $\quad$ $H_A: \sigma_1^2 \neq \sigma_2^2$ $\quad$ (kétoldalas kérdésfeltevés)

$H_0: \sigma_1^2 = \sigma_2^2$ $\quad$ $H_A : \sigma_1^2 < \sigma_2^2$ $\quad$ (egyoldalas kérdésfeltevés)

2) Válasszunk tévedési valószínûséget (α)

3) Válasszuk ki a két szórás közül a nagyobbat. Kapja ez az 1 indexet.

4) Képezzük a számított $\hat{\mathrm{F}}$ hányadost:

$$\hat{F} = \frac{s_1^2}{s_2^2}.$$

5) Keressük meg F_{kritikus} értékét

F -nek három változója van: a *tévedési valószínûség* (α) és a két *szabadsági fok*: $\nu_1 = n_1 - 1$ és $\nu_2 = n_2 - 1$.

A kritikus *F* értékek a táblázatok α oldalán, a ν_1 oszlopban és a ν_2 sorban találhatók. Egyoldalas kérdésfeltevésnél az α valószínüséghez tartozó táblázatot, kétoldalasnál az $\alpha/2$ valószínüséghez tartozó táblázatot kell választani. Ha a számított $\hat{\mathrm{F}}$ nagyobb a kritikusnál, a nullahipotézist el kell vetni, a szórások szignifikánsan eltérnek egymástól, adott tévedési valószínûséggel.

5.2 Numerikus példa: Elfogadhatjuk-e azt az 5.1 példában megadott hipotézist, miszerint az abban szereplõ szórások megegyeznek? (L. Sachs: Statistische Methoden, Springer, Berlin 1993. p. 77)

$H_0 : \sigma_1^2 = \sigma_2^2$ $\quad$ $H_A: \sigma_1^2 \neq \sigma_2^2$ $\quad$ (kétoldalas kérdésfeltevés)

$\alpha = 0.05$

$$s_1^2 = 33.44, \quad s_2^2 = 22.55, \; n_1 = 30, \quad n_2 = 30, \quad \nu_1 = 29,. \quad \nu_1 = 29$$

$$\hat{\mathrm{F}} = \frac{33.44}{22.55} = 1.483 < 2.09 = \mathrm{F}(29,29,0.025)$$

A nullahipotézist elfogadjuk.