

STATISZTIKA

A STATISZTIKA ALAPJAI

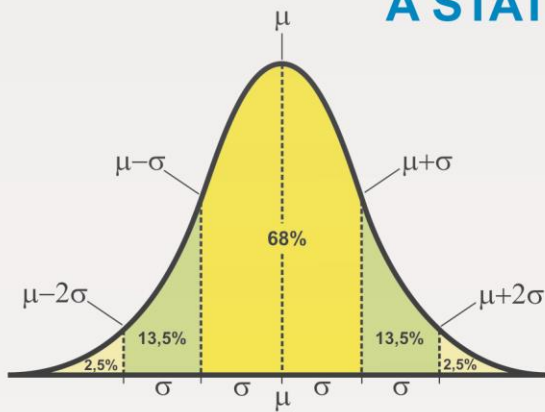


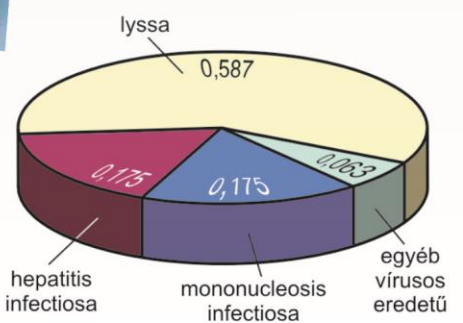
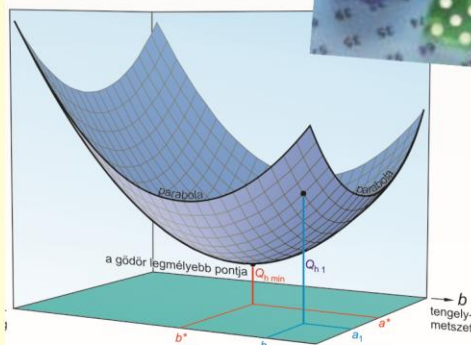
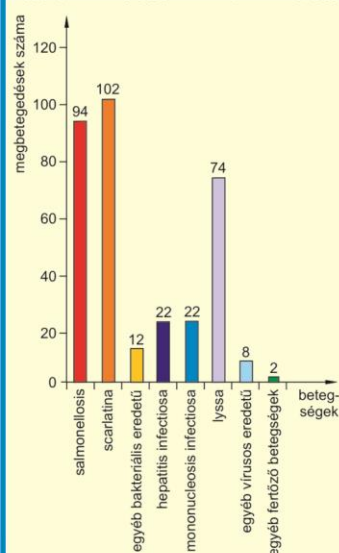
TABLE OF CRITICAL VALUES FOR T

One Tailed Significance

0.1	0.05	0.025	0.005	0.0025	0.0005	0.0002
-----	------	-------	-------	--------	--------	--------

Two Tailed Significance

0.2	0.1	0.05	0.01	0.005	0.001	0.0005
1.89	2.92	4.30	9.92	14.09	31.60	44.70
1.64	2.35	3.18	5.84	7.45	12.92	16.33
1.53	2.13	2.78	4.60	5.60	8.61	10.31
1.48	2.02	2.57	4.03	4.77	6.87	7.98
1.44	1.94	2.45	3.71	4.32	5.96	6.79
1.41	1.89	2.36	3.50	4.03	5.41	6.08
1.40	1.86	2.31	3.36	3.83	5.04	5.62
1.38	1.83	2.26	3.25	3.69	4.78	5.29
1.37	1.81	2.23	3.17	3.58	4.59	5.05
1.36	1.80	2.20	3.11	3.50	4.44	4.86
1.36	1.78	2.18	3.05	3.43	4.32	4.72
1.35	1.77	2.16	3.01	3.37	4.22	4.60
1.35	1.76	2.14	2.98	3.33	4.14	4.50
1.34	1.75	2.13	2.95	3.29	4.07	4.42
1.34	1.75	2.12	2.92	3.25	4.01	4.35
1.33	1.74	2.11	2.90	3.21	3.97	4.29
1.33	1.73	2.10	2.88	3.17	3.92	4.23
		2.09	2.86	3.14	3.88	4.19
		2.09	2.84	3.11	3.85	4.15
		2.08	2.82	3.08	3.82	4.11
		2.07	2.80	3.05	3.79	4.08
		2.07	2.78	3.02	3.77	4.05
		2.06	2.76	2.99	3.75	4.02
		2.06	2.74	2.96	3.72	4.00



ÖSSZEFOGLALÁS

ADAT: Minőségi vagy mennyiségi jellemző.

POPULÁCIÓ, ALAPSOKASÁG: Azoknak az egyedeknek (elemeknek), objektumoknak az összessége (jellemzőikkel együtt), amelynek tulajdonságairól egy vizsgálat során információt kívánunk nyerni. Elemszáma N , ami végtelen is lehet.

MINTA: A populációnak megvizsgálásra kiválasztott alkalmas része, amelynek alapján következtetünk a populáció tulajdonságaira. Elemszáma n , általában $n \ll N$.

VÁLTOZÓ: A populáció (egy adott tulajdonságára jellemző) általános eleme, x .

ABSZOLÚT GYAKORISÁG: A csoportosított, osztályokba sorolt adatok esetén az egy osztályba eső adatok száma.

RELATÍV GYAKORISÁG: Az egyes osztályokban lévő elemek számának és a minta elemszámának hányadosa.

GYAKORISÁGI ELOSZLÁS: Az osztályonkénti gyakoriságok, vagy a relatív gyakoriságok megadása, illetve ennek grafikus ábrázolása.

HISZTOGRAM: A gyakorisági eloszlás olyan grafikus megjelenítése, amikor minden osztály fölé olyan téglalapot („oszlopot”) rajzolunk, melynek területe arányos az osztályba eső gyakorisággal.

ELMÉLETI ELOSZLÁS: A populáció eloszlása.

NORMÁLIS, vagy GAUSS-ELOSZLÁS, $N(\mu, \sigma)$: Az elméleti eloszlás egy fajtája. Szimmetrikus, két paraméterrel, a várható értékkel (μ) és az elméleti szórással (σ) jellemzett eloszlás. A sok, egymástól független hatás eredményeképpen kialakuló értékek normális eloszlást követnek.

HARANGGÖRBE vagy GAUSS-GÖRBE: Gauss-eloszlású populáció hisztogramjának burkoló görbéje, amennyiben $N = \infty$ és $\Delta x \rightarrow 0$, azaz, ha végtelen elemszámú a populáció és az osztályszélességgel 0-hoz közelítünk.

VÁRHATÓ ÉRTÉK (GAUSS-ELOSZLÁSÉ): Az eloszlás egyik paramétere (μ), a Gauss-görbe maximumához tartozó változó érték. Vagy másképpen a változónak az az értéke, amely a populációt két egyenlő részre osztja, azaz a populációban az ennél nagyobb illetve kisebb elemek számaránya megegyezik.

ELMÉLETI SZÓRÁS (GAUSS-ELOSZLÁSÉ): A Gauss-görbe szélességét mérő paraméter (σ). A haranggörbe szélessége körülbelül a magassága felénél mérve 2σ -val egyenlő.

ÁTLAG: A várható érték becslésére szolgáló legstabilabb középérték (\bar{x}). Számítási közép.

TAPASZTALATI SZÓRÁS (s): Az elméleti szórás becslésére szolgál ($s = \sqrt{\sum (x_i - \bar{x})^2 / (n - 1)}$). Azt, hogy az adatok hogyan szóródnak, milyen variabilitást mutatnak, a szóródás mérőszámaival lehet jellemezni. A leggyakrabban használt mérőszám épp a tapasztalati szórás, vagy a minta szórása, amely az adatoknak az átlagtól való átlagos eltérését jellemzi.

VARIANCIA: A tapasztalati szórás négyzete (s^2).

STANDARD NORMÁLIS ELOSZLÁS, $N(0,1)$: A $\mu = 0$, $\sigma = 1$ paraméterű normális eloszlás.

STANDARD HIBA vagy AZ ÁTLAG SZÓRÁSA: A minták átlagának szóródását kifejező, a tapasztalati szórásból származtatott mennyiség ($s_{\bar{x}} = s / \sqrt{n}$).

KONFIDENCIA INTERVALLUMOK: Az átlag és a standard hiba segítségével megadható értéktartományok, amelyben a várható érték adott bizonyossággal megtalálható. Leggyakrabban a hibakorlátot ($\bar{x} \pm 2s_{\bar{x}}$) használják 95%-os konfidencia szint esetén.

KONFIDENCIA SZINTEK: Adott konfidencia intervallumokhoz tartozó, a bizonyosság mértékét kifejező szám %-ban kifejezve. Értéke leggyakrabban 95%.

LINEÁRIS REGRESSZIÓ: A hibával terhelt mérési pontokból két mennyiség (pl. x , y) között fennálló lineáris kapcsolatra való következtetés módszere.

LEGKISEBB NÉGYZETEK MÓDSZERE: Lineáris regresszió esetén a mérési pontokra legjobban illeszkedő egyenes az, amelyre a mérési pontok és az egyenes függőlegesen mért távolságainak négyzetösszege a lehető legkisebb.

KORRELÁCIÓS EGYÜTTHATÓ: A változók közötti kapcsolatot, illetve annak szorosságát jellemző (-1 és 1 közé eső) szám. ($r = 0 \rightarrow$ nincs kapcsolat, $r = \pm 1 \rightarrow$ maximális szorosságú a kapcsolat).

STATISZTIKAI PRÓBÁK: Egy hipotézis elfogadásának vagy elvetésének eldöntésére használt módszerek.

NULLHIPOTÉZIS (H_0): A hipotézisvizsgálatok során, a statisztikai próbák kiindulási hipotézise, amely az ismert eloszlású alapállapotból indul ki. Általában a feltett kérdésre megfogalmazott egyértelmű tagadó válasz.

ALTERNATÍV HIPOTÉZIS (H_1): A nullhipotézis elvetésekor életbe lépő, a feltett kérdésre megfogalmazott igenlő válasz.

KRITIKUS TARTOMÁNY: Általánosan azt mondhatjuk, hogy ez az a tartomány, amely ha a nullhipotézis igaz, akkor csak előre megadott igen kis valószínűséggel (leggyakrabban 5%) tartalmazza a mintát. Ezért, ha a minta ide esik, a nullhipotézist elvetjük.

ELFOGADÁSI TARTOMÁNY: A kritikus tartomány „ellentettje” jobban mondva kiegészítője, komplementere.

ELSŐFAJÚ HIBA: A valójában igaz nullhipotézist tévesen elvetjük.

MÁSODFAJÚ HIBA: A valójában hamis nullhipotézist tévesen elfogadjuk.

t-PRÓBÁK: Egy vagy több normális eloszlású változó, várható értékére vonatkozó hipotézisek ellenőrzésére használt statisztikai próbák.

χ^2 -(khi-négyzet) PRÓBÁK: Leggyakrabban kategoriális jellemzők függetlenségének vizsgálatára használható statisztikai próbák.

SZABADSÁGFOK: Az egymástól függetlenül választható tagok, mintaelemek számával egyenlő. Nyilván, ezek nem lehetnek függetlenek akkor, ha érvényesül köztük egy, vagy több összefüggés. Ilyenkor az összefüggés(ek) számát le kell vonni a mintaelemszámból: a különbség értéke lesz a szabadságfok.

BEVEZETÉS

Ismerjük annak az embernek a történetét, aki hétfőn whiskyt ivott szódával, kedden gint szódával, szerdán pedig rumot szódával. Mivel az eredmény mindig azonos volt, arra a következtetésre jutott, hogy a szódától rúgott be.

Ha nem is a whisky és szóda viszonylatában, de nagyon hasonlóan viselkedünk mindannyian. Sokunkra jellemző, hogy **könnyen vonunk le megalapozatlan következtetéseket**, és így hozunk döntéseket. Egyedi esetek alapján általánosítunk, a rendelkezésünkre álló ismeretekből is önkényesen válogatunk vélt igazunk alátámasztására. Az így kialakított álláspontunkhoz, véleményünkhöz pedig sokszor megingathatatlanul ragaszkodunk.

Első megközelítésként azt mondhatjuk, hogy **a statisztika** ennek az általános „kórnak” a leküzdésére alkalmas tudomány. **Segítséget nyújt ahhoz, hogy** kritikusabban rendezzük el gondolatainkat, és **kétkedésünket** — ami minden igényes értelmiségi tevékenység alapja — mindig **éberen tartsuk**.

Mivel nap mint nap találkozunk statisztikai kimutatásokkal, úgy tűnhet, hogy ismerjük az előállításukra szolgáló módszereket is. Ez részben igaz, hiszen például, **aki** Magyarországon az utóbbi néhány évtizedben **járt iskolába**, az életében már **néhányszor biztosan használt statisztikai módszereket**. Elég csak azt az esetet említeni, **amikor** egy kisdíák **kiszámolja** az egyes tantárgyakból szerzett **érdemjegyeinek átlagait** abból a célból, hogy megtudja, milyen osztályzatokra számíthat a félévi vagy év végi bizonyítványában. Ilyenkor a tanuló a statisztikai **becslés** egy tipikus esetét alkalmazta anélkül, hogy tudott volna róla.

Egyébként a „statisztika” szó több jelentéssel is bír. Az itt használt értelmét leginkább a latin eredetre visszanyúlva a „status” szóból ismerhetjük meg. A status eredeti jelentése: állapot, helyzet, a dolgok állásának módja. Ennek megismeréséhez, illetve leírásához biztosítanak lehetőséget az **adatok**. Általánosságban a környező világ **minőségi és mennyiségi** jellemzőit nevezhetjük adatoknak. Mindennapi életünkben leggyakrabban előforduló adatok például a személyi adataink: nevünk, születési helyünk, születésünk időpontja, ... ; vagy akármilyen boltban az árucikkek neve, az árucikkek ára, ... ; de egészségi állapotunkkal kapcsolatban is mondhatunk példát: arcunk sápadtsága, vérnyomásunk, hőmérsékletünk, de akármilyen laboratóriumi diagnosztikai vizsgálat eredménye is adat.

Normális esetben adatokat csak valamilyen cél érdekében gyűjtünk. Például, azért kérdezzük meg valakinek a telefonszámát, hogy később fel tudjuk hívni. Nem szerencsés az a hozzáállás, hogy gyűjtsünk adatokat, majd csak jó lesz valamire, és utólag próbálunk célokat kitalálni. (Ez legfeljebb a titkosrendőröknél szokásos.) Az összegyűjtött, de rendezetlen adatok önmagukban, sok esetben teljesen használhatatlanok. Gondoljunk például arra, hogy egy telefonkönyv adatait a központba való beérkezésük sorrendjében adnák közre, mire mennénk vele? Sokszor kell az adatokat értékelni, például fontosságuk szerint. Ezt teszi az orvos is, amikor felállít egy diagnózist, vagy amikor a beteg állapotáról nyilatkozik.

Az **adatokat** tehát **össze kell gyűjteni**, legtöbbször **fel kell dolgozni** őket, és szükség esetén **következtetéseket kell levonni** belőlük, illetve időnként **döntéseket kell hozni** azok alapján. A **statisztika** az a tudomány, ami mindezt megtanít bennünket. A „statisztika” szó előtt gyakran szerepelő „**bio-**”, előtag arra utal, hogy bizonyos statisztikai módszereket az **élővilággal kapcsolatos jelenségek** elemzésére használnak. Ehhez hasonlóan az **orvosi statisztika** módszerei a felhasználások még konkrétabb körére, nevezetesen orvosi problémák megoldására lettek kidolgozva.

Nem könnyű első éves orvos-egyetemisták számára meggyőző érveket hozni annak alátámasztására, hogy **a statisztika számukra is fontos**, mondhatni **nélkülözhetetlen** tudomány. Ezért most csak néhány példát említünk.

Az orvosi fizikai gyakorlatok nagy részén, de az elméleti modul többi tárgyának részeként is, sőt az **egyetemi évek során** később is a hallgatók egyéni méréseket végeznek. A **mérési adatokból megbízható következtetéseket** pedig csak bizonyos statisztikai ismeretek birtokában lehet **levonni**.

A **kórtörténeti leírásokban**, a laboratóriumi naplókban rengeteg adat szerepel. Nagyon fontos, hogy ezen **adatok helyes értékelését, a helyes következtetések levonásának módszereit és az eredmények megbízhatóságának vizsgálatát** az orvosok, fogorvosok, gyógyszerészek **megismerjék, és munkájukban alkalmazzák**. A statisztika óvhat meg bennünket az új eljárások, gyógyszerek reklámáradatának esetleges hazugságaitól is.

Megemlíthetjük **az orvosi irodalom értő olvasásának** nehézségeit is. Példaként álljon itt egy rövid kivonat egy orvosi közleményből. „*Az 1. csoportban az átlagos $-3,94 \pm 1,3$ dioptriás preoperatív fénytörési hiba az 1 éves követése során $-0,47 \pm 0,54$ dioptriára ... csökkent.*”; majd később „*A statisztikai eredményeket kétmintás t-próba és regressziós analízis segítségével értékeltük.*” A kérdés ugye csak az: mit jelentenek a számok, illetve mik ezek a módszerek?

Végül, de nem utolsósorban a **statisztika** egy **sajátos szemléletet**, gondolkodási módot **közvetít**, ami nagyon **hasonlít az orvosi gondolkodáshoz** is, így megismerése ebből a szempontból is kívánatos. Ennek illusztrálására szintén bemutatunk egy példát. Tegyük fel, hogy *páciensünk fejfájásra panaszkodik, a kérdés az, hogy mi lehet ennek az oka*. Orvosi tanulmányaink alapján eszünkbe kell, hogy jusson az „összes” lehetséges ok. Ezek közül néhány:

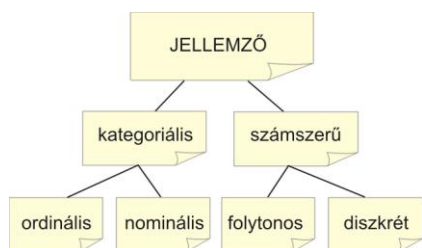
1. Magas a vérnyomása.
2. Rossz a szemüvege.
3. Túl magas a szemnyomása.
4. Koponyaüri daganata van.
5. A nyaki csigolyái meszesek.
6. Érzékeny a frontokra.
7. Oxigénszegény környezetben dolgozik.

Mivel itt általában több helyes válasz is létezhet egyszerre, ezért a válaszokat egyesével vizsgáljuk meg, és ellenőrizzük, hogy a gyanú alapos, vagy alaptalan.

Ez az eljárás alapjaiban megegyezik a statisztikában használatos, e rövid összefoglalóban is bemutatásra kerülő egyik módszerrel, a „hipotézisvizsgálattal”. Első közelítésben a statisztikai tevékenységeket négy csoportba

sorolhatjuk, de ezek között nincs éles határ: **adatgyűjtés, az adatok áttekinthetővé tétele, az adatok elemzése, és a következtetések.**

ADATGYŰJTÉS, AZ ADATOK FŐBB TÍPUSAI



1. táblázat. A jellemzők osztályozása.

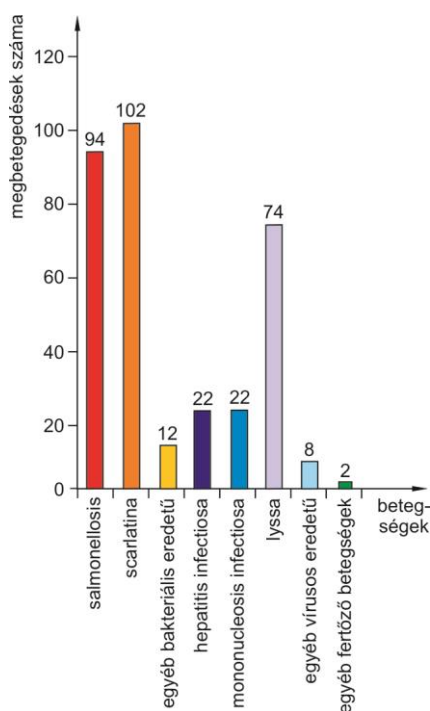
	abszolút gyakoriság
	absolute frequency
	absolute Häufigkeit

	relatív gyakoriság
	relative frequency
	relative Häufigkeit

Mint már említettük, az adatgyűjtés valamilyen cél eléréséhez szükséges. Vannak adatok, amik csak bizonyos dolgok azonosítására, megkülönböztetésére szolgálnak. Nagyobb számban adatokat akkor gyűjtünk, ha azt reméljük, hogy ezek segítségével valamilyen korábban feltett kérdésünkre feleletet kapunk. Az adatgyűjtés módját, az **adatokhoz való hozzájutást** a továbbiakban általánosan „kísérletnek”, ill. ha emberen végzik „vizsgálatnak” nevezzük. Az adatok egy része ismert, csak meg kell kérdezni valakitől, másik részét meg kell mérni valahogy, de a fenti értelemben kísérletnek tekinthető egy természeti jelenség megfigyelése, vagy egy dobókocka feldobása is. A **kísérlet, vagy vizsgálat eredménye**, az adat lehet **minőségi**, legfeljebb kategóriákba sorolható, azaz **kategorialis**, vagy **mennyiségi**, tehát számmal jellemzett, azaz **számszerű** jellemző. A betegségek neve, ami azonosításuknál többre nem nagyon alkalmas, minőségi (kvalitatív) jellemzést jelent. Ugyanígy a betegség súlyossága, a fertőzést előidéző kórokozó típusa szintén minőségi jellemző, de például az esetleges kiütések nagysága, vagy a betegség lefolyásának ideje már mérőszámmal (és mértékegységgel) is megadható, tehát mennyiségi (kvantitatív) jellemző. A minőségi vagy kategorialis jellemzőket két csoportba sorolhatjuk aszerint, hogy valamilyen természetes sorba rendezhetők-e vagy sem. Sorba rendezhető, azaz **ordinális** jellemző például egy betegség lefolyásának foka: *enyhe, közepes, súlyos*; de nem ilyen, azaz **nominális** például a vércsoport: *A, AB, B, 0*. A számszerű jellemzőkön belül megkülönböztetünk **folytonos**, azaz bizonyos határokon belül tetszőleges értéket felvehető, illetve **diszkrét**, azaz csak bizonyos értékeket felvehető jellemzőt. A testsúly, testmagasság, vérnyomás folytonos, de például egy családon belül a gyermekek száma csak egész szám lehet, tehát az diszkrét jellemző. (A jellemzők főbb típusait az 1. táblázatban foglaltuk össze.) Itt jegyezzük meg, hogy a folytonos jellemző csak elvileg folytonos, a gyakorlatban mindig diszkrét értékekkel dolgozunk (ellenkező esetben végtelen tizedes törtek alkalmazására is szükségünk lenne).

AZ ADATOK ÁTTEKINTHETŐVÉ TÉTELE, GYAKORISÁGOK ÉS ÁBRÁZOLÁSUK

A mindennapi életben is gyakran előfordul, hogy egy probléma kapcsán viszonylag sok adat áll rendelkezésünkre. Ilyen esetekben **szükséges, hogy az adatokról valamilyen áttekintésünk legyen.**



1. a. ábra. Oszlop diagram. Abszolút gyakoriságok a betegségek, mint kategóriák függvényében.

KÓROKOZÓ	BETEGSÉG	abszolút gyakoriság		relatív gyakoriság	
baktérium	salmonellosis (szalmonella fertőzés)	94	208	0,280	0,619
	scarlatina (skarlát)	102		0,304	
	egyéb bakteriális eredetű	12		0,036	
vírus	hepatitis infectiosa (fertőző májgyulladás)	22	126	0,065	0,375
	mononucleosis infectiosa (mirigyláz)	22		0,065	
	lyssa (veszetség)	74		0,220	
	egyéb vírusos eredetű	8		0,0238	
egyéb	egyéb fertőző betegségek	2	2	0,006	0,006
összesen:		336	336	1,000	1,000

2. táblázat. Egy összesítés a fertőző megbetegedésekről.

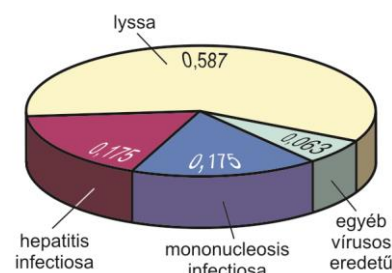
A fenti 2. táblázat a Budapesten, 2000 októberében bejelentett fertőző megbetegedések összesítését mutatja.

A táblázat első számoszlopában látható számok (94, 102, stb.) azt mutatják, hogy az egyes betegségtípusokból (Salmonellosis, Scarlatina, stb.) hányat észleltek az adott időszakban. Ezeket a számokat **abszolút gyakoriságoknak** nevezzük. A következő oszlopban a részösszegek (208, 126, 2) szerepelnek, tehát az, hogy az észlelt betegségek közül hány volt bakteriális, vírusos, vagy egyéb eredetű. Ezek szintén abszolút gyakoriságok.

Ha az abszolút gyakoriságokat elosztjuk az adott területen, adott időszakban előforduló összes fertőző betegségek számával (336), akkor megkapjuk a viszonylagos értékeket, a **relatív gyakoriságokat**. Ezek mindig 0 és 1 közé eső számok, és a táblázat következő két oszlopa tartalmazza őket. Százal való szorzással %-ban is kifejezhetők. A relatív gyakoriság, értelmezéséből kifolyólag, egy hányados. Ezért **nem csak azt kell tisztázni, hogy minek a relatív gyakoriságáról beszélünk, hanem azt is, hogy mihez viszonyítottunk**.

Ha például arra vagyunk kíváncsiak, hogy a bakteriális eredetű betegségeken belül milyen gyakori a szalmonella fertőzés, akkor a szalmonella fertőzések számát (94) az összes bakteriális eredetű betegségek számával (208) kell elosztani. Az így kapott hányados (0,452) is relatív gyakoriság, de most nem az összes betegséghez (336), hanem csak a bakteriális eredetű betegségekhez (208) viszonyítottunk.

Az abszolút és relatív gyakoriságok ábrázolására sok lehetőség kínálkozik. Ezek közül mutatunk be kettőt az 1. a. és b. ábrán.



1. b. ábra. Torta diagram.
Relatív gyakoriságok a vírusos eredetű fertőző betegségek (kategóriák) megoszlásáról.

Ellenőrző kérdés:

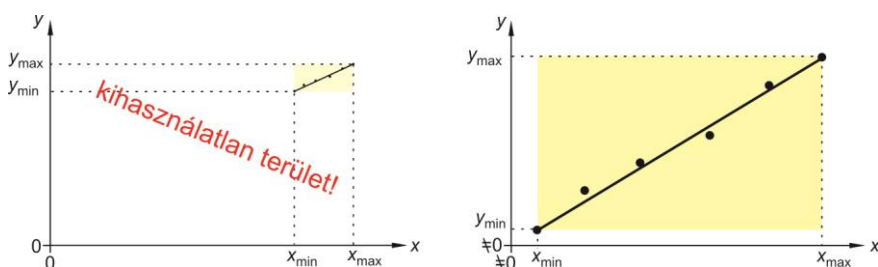
Nincs-e ellentmondás a következő állításokban: az én csoportomban a lányok relatív gyakorisága kisebb, mint a barátomban, de nálunk mégis több lány van, mint náluk?

AZ ADATOK KÖZÖTTI KAPCSOLAT SZEMLÉLETES BEMUTATÁSA, GRAFIKUS ÁBRÁZOLÁS

Kísérleteink, méréseink során gyakran előfordul, hogy **többféle jellemzőt** is meghatározunk, illetve, hogy **egy jellemzőt egy előre meghatározott paraméter**, mondjuk az idő **függvényében vizsgálunk**. (Ez utóbbira példa a rendszeres lázmérés a kórházakban.) Ilyen esetekben a különböző adatsorok közötti kapcsolatra, az esetleges összefüggésekre vagyunk kíváncsiak. Ezért a jobb áttekinthetőség érdekében célszerű az eredményeket **grafikusan** ábrázolni. Fontos megjegyeznünk, hogy a „**kapcsolat**” szó a legáltalánosabb értelemben használandó, tehát **nem jelent ok-okozati összefüggést**.

Grafikus ábrázolásnál először két fontos dologban kell döntenünk. Az egyik a **tengelyek beosztása** (skalázás), a másik a **kezdőértékek megválasztása**. Nem szükségszerű ugyanis grafikonunkon az origót minden esetben feltüntetni. Ha például koncentráció meghatározás érdekében vizes fehérjeoldatok törésmutatóját mérjük, előre tudjuk, hogy a desztillált víz törésmutatójánál (1,333) kisebb értéket nem fogunk kapni.

A fenti kérdések eldöntésében általános irányadó szempont az lehet, hogy grafikonunk töltse ki a rendelkezésre álló terület nagy részét (lásd 2. ábra).



2. ábra. A grafikon helytelen és helyes elhelyezése a milliméterpapíron.

Szükségtelen azonban a skálabeosztásokat olyan mértékben megnövelni, hogy a grafikon kiértékelése nagyobb pontossággal történhessék, mint amilyenre a mérőeszköz maga egyáltalán lehetőséget nyújt. A pontosság ilyen látszólagos növelése megtévesztő lehet.

Méréseinknél bizonyos hibával mindig számolnunk kell. Ezért a görbe „kihúzásánál” ne ragaszkodjunk szigorúan a mérési pontokhoz. A görbe alá és fölé kerülő mérési pontok száma, illetve a görbétől való eltérésük mértéke körülbelül

egyezzen meg. Ily módon ritka kivételektől eltekintve sima lefutású, folytonos görbét illeszthetünk adatainkhoz (lásd lineáris regresszió).

STATISZTIKAI KÖVETKEZTETÉS ÉS A VALÓSZÍNŰSÉGSZÁMÍTÁS

A statisztikai módszerek végső célja a következtetés. A statisztikai következtetések sémája nagyon hasonlít a logikai következtetések sémájához. A logikában a szillogizmus a következtetés egyik fajtája, amelyben bizonyos dolgok megállapításából **szükségszerűen következik** valami más. (A klasszikus példa szerint: *minden ember halandó, Levente ember, tehát Levente halandó.*)

A statisztikai következtetés ettől annyiban tér el (ami persze nem jelentéktelen különbség), hogy míg a **logikai következtetést teljes** (100%-os) **bizonyossággal** állíthatjuk, addig a **statisztikait** csak **adott**, (100%-nál mindig kisebb) **bizonyossággal**. Ebből következik tehát, hogy a statisztikai következtetéseknel **tévedhetünk**. Ha például valamit 95% bizonyossággal állítunk, akkor az azt jelenti, hogy átlagosan minden 100 eset közül 5 esetben tévedünk; a bizonytalanság 5%-os. Állításaink biztonságát tehát számszerűen is kifejezhetjük.

A következtetés újrafogalmazásával a bizonytalanságot általában tetszés szerint csökkenthetjük, ez azonban többnyire azzal a veszteséggel jár, hogy állításunk, azaz a következtetés egyre semmitmondóbbá válik. Erre is lássunk egy példát! Egy bankrablást követően az eseményekről beszámoló rendőri jelentés így ír: „... a szemtanúk látták, amint az elkövetők gépkocsiba szállnak; a rendőrség megállapította, hogy a tettesek vagy saját gépkocsijukon hagyták el a helyszínt, vagy taxival távoztak, vagy lopott, esetleg bérelt autót használtak”. **Ha tehát következtetésünk bizonytalanságát csökkentjük, akkor annak általában az ára, hogy a következtetés értéke, használhatósága is csökken.** Következtetési módszereinket tehát e két ellentétes tendencia szabályozza.

A bizonytalanság oka az, hogy a statisztikai következtetéseknel (a logikaival ellentétben) nem tudunk minden körülményt számba venni. A feldobott pénzérmét nem a „vak véletlen” vezérli, amikor egyik vagy másik oldalára esik. egyszerűen arról van szó, hogy nem ismerjük kellő pontossággal azokat az adatokat, amelyek egyértelműen meghatároznák a pénzérme végső állapotát, nevezetesen azt, hogy a dobás „eredménye” fej vagy írás. Mivel nem tudunk minden körülményt figyelembe venni, ezért nem tudunk egyértelmű választ sem adni, tehát csak azt mondhatjuk, hogy **a jelenség véletlenszerű, ahol a „véletlen” szó csak ismereteink hiányát fejezi ki.**

A szerencsejátékokkal kapcsolatban már régen megfigyelték, hogy **a véletlen tömegjelenségek** is bizonyos **törvényszerűségeket követnek**. A feldobott pénzérme példájánál maradva, ha sokszor megismételjük a „kísérletet”, akkor az érme körülbelül az esetek felében esik fejjre, felében írásra. Bár bizonyítani nem tudjuk, tapasztalatból állíthatjuk, hogy nagy számú (független) kísérlet esetén a fejek és írások **relatív gyakorisága**, azaz a $[(\text{fejek száma})/(\text{fejek} + \text{írások száma})]$ illetve az $[(\text{írások száma})/(\text{fejek} + \text{írások száma})]$ **stabilitást mutat**, mindkettő egyaránt $\frac{1}{2}$ közelében lesz. (Ez a nagy számok törvénye.) Ennek alapján azt is mondhatjuk, hogy annak a **valószínűsége**, hogy egyetlen pénzfeldobáskor fejet kapjunk, éppen $\frac{1}{2}$. (Az írás valószínűsége természetesen ugyanekkora.)

A **valószínűségszámítás matematikai leírást ad** az anyagi világban tömegméretekben lejátszódó **olyan jelenségek törvényszerűségeire**, melyek lefolyását **a számba vehető körülmények nem határozzák meg egyértelműen**. kísérleteink, megfigyeléseink, mérési eredményeink éppen ebbe a kategóriába esnek, így a statisztika a valószínűségszámítás eredményeire építhet.

POPULÁCIÓ, VÁLTOZÓ, MINTA

Kísérleteket, méréseket, megfigyeléseket mindannyian végzünk, ezt illusztrálja néhány példával a 3. táblázat, (ahol az utolsó oszlopban található, zárójelben lévő számok a hallgatói mérések jegyzetbeli sorszámát mutatják).

KI MIT MÉR?		
FIZIKUS	ORVOS	ORVOSTANHALLGATÓ AZ ORVOSI FIZIKA GYAKORLATOKON
hosszúság	testmagasság	vörösvérsejt átmérő (3.)
frekvencia	pulzusszám	impulzus gyakoriság (9., 20.)
hőmérséklet	testhőmérséklet	–
koncentráció	vércukor-szint	vérplazma fehérje- koncentráció (5.)
feszültség	EKG-jel	EKG-jel (24.)
teljesítménysűrűség	hallásküszöb	hallásküszöb (22.)
nyomás	vérnyomás	–
impedancia	bőrimpedancia, (bőrellenállás)	bőrimpedancia (21.)

3. táblázat. Mit mér a fizikus, az orvos és az orvostanhallgató?

Ismét hangsúlyozzuk, hogy méréseink célja valaminek a megismerése, valamilyen kérdésnek a megválaszolása.

Válasszuk ki táblázatunkból példaképpen a pulzusszám mérést, amely könnyen elvégezhető az orvosi fizika gyakorlaton is. A pulzusszám a szívverés frekvenciája, (elvileg) folytonos jellemző és csak az egyszerűség kedvéért, illetve megszokásból használunk diszkrét (egész) értékeket. Mértékegysége az 1/perc. Ezek figyelembevételével a továbbiakban csak a mérőszámokkal dolgozunk, de nem szabad elfelejtenünk, hogy a végső eredményeket mindig mértékegységgel együtt kell megadnunk. Sokféle kérdést tehetünk fel, amire ettől a méréstől várjuk a választ. Például:

1. *MEKKORA* X.Y. hallgató pulzusszáma?
2. *MEKKORA* a normális pulzusszám?
3. *VÁLTOZIK-E* a pulzusszám egyperces lélegzet-visszatartás után?
4. *VAN-E KÜLÖNBESÉG* a lányok és a fiúk pulzusszáma között? , stb.

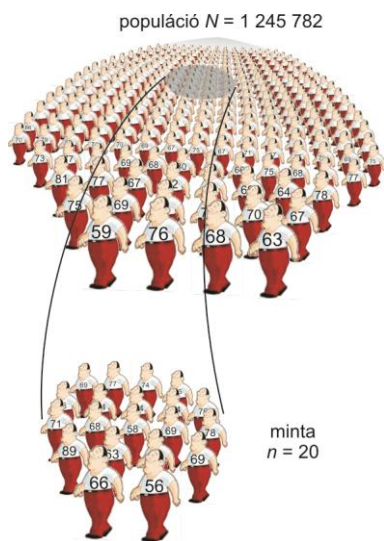
Vegyük sorra ezeket. **Statisztikai ismeretektől mentesen** bárki azt gondolhatná, hogy az első kérdésre igazán könnyű válaszolni, nevezetesen megmértem X.Y. pulzusszámát, **kijön egy eredmény és kész**. Ha azonban valaki már egy kicsit statisztikával „fertőzött”, akkor **kétkedése felébred** és egyértelmű válasz helyett, újabb kérdések merülnek fel benne: biztos, hogy ez a jó eredmény, nem hibáztam valahol? Ha még azt is tudja az illető, hogy a mérés eredménye, a „kísérlet” lefolyása a „véletlentől” is függ, tehát a legnagyobb igyekezettel sem tudunk „hibátlanul” mérni, akkor arra az elhatározásra jut, hogy **újra mér**.

A többszöri mérés végzésekor hallgatólagosan mindig feltételezzük, hogy **ugyanazt mérjük** még egyszer, ezért várhatóan **ugyanakkora** lesz az **eredmény** is. Másképpen mondva X.Y. pulzusszáma az ismételt mérések során **tartósan egyik irányban sem változik**, ennek ellenére a számba nem vehető, véletlen ingadozások folytán **mégsem kapunk azonos eredményeket**. Azt is mondhatjuk, hogy a kísérlet eredményének két része van, az egyik, — nevezzük fő résznek — determinisztikus (meghatározott), a másik pedig, — nevezzük ezt mellék résznek — sztochasztikus (véletlen). (E két részt természetesen nem tudjuk eleve különválasztani.)

A többszöri mérést úgy is tekinthetjük, hogy létezik egy halmaz, más néven **alapsokaság**, vagy **populáció**, és minden mérés során ennek a halmaznak választjuk ki egy elemét. (Ebben az esetben a halmaz elvileg végtelen számú elemet tartalmaz, de ez nem alapfeltétel.) A halmaz általános elemét **változónak** nevezzük és szokásos módon például *x*-szel jelöljük. Ez a változó különböző



 minta
 sample
 Stichprobe



3. ábra. A populáció, a változó és a minta szemléltetése.

értékeket vehet fel, hogy éppen melyiket, azt mondja meg az adott mérési eredmény.

Egyetlen mérés alapján a többi feltett kérdésre sem tudunk választ adni. A második kérdés esetében úgy gondolhatjuk, hogy létezik egy determinisztikus normális pulzusszám, és az egyének pulzusszáma e körül ingadozik véletlenszerűen. Itt a populációt például úgy lehet elképzelni (3. ábra), hogy egy adott pillanatban ismerjük sok embernek (mondjuk $N = 1245782$ egyednek) a pulzusszámát, és ezek az emberek a pulzusszámukkal együtt alkotják az alapsokaságot. (Megjegyezzük, hogy ebben az esetben az alapsokaság véges sok elemű.)

Az első esetben a méréseket ugyanazon a személyen ismételjük meg (akárhányszor), a másodikban különböző személyeken mérünk pulzusszámot. A változó tehát mindkét esetben igen hasonló, mégis két különböző populációról, alapsokaságról van szó. A harmadik, illetve negyedik kérdésre később visszatérünk.

Bár a feltett kérdések valójában a populációra vonatkoznak, a legtöbb esetben nem áll módunkban annak teljes megismerése. Emiatt az N elemű sokaságból (N végtelen is lehet) **mintát** veszünk. **A mintavétel a sokaság n számú elemének véletlenszerű kiválasztásából áll.** (Ezt hajtjuk végre például a többszöri mérés alkalmával.) Az egésznek természetesen csak akkor van értelme, ha n lényegesen kisebb lehet, mint N (lásd 3. ábra).

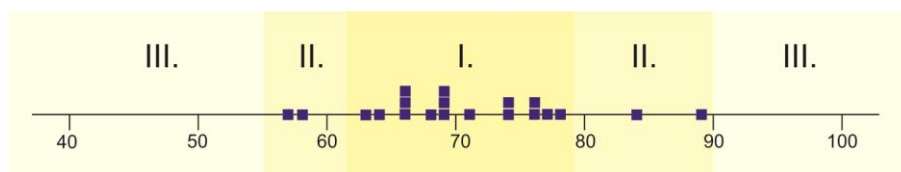
A MINTA ELOSZLÁSA, GYAKORISÁGI ELOSZLÁS, HISZTOGRAM

Ezt a fogalmat egy példa kapcsán vezetjük be. Mivel végül választ szeretnénk adni az előző részben feltett kérdésekre, most válasszuk ki a másodikat, (**MEKKORA** a normális pulzusszám?) és ennek érdekében mérjük meg egy tanulócsoporthallgatóinak pulzusszámát. A tanulócsoporthallgatóinak pulzusszám adatokkal együtt a feltett kérdéssel kapcsolatos alapsokaságból vett mintának tekinthető ($n = 20$) (lásd 3. ábra). Mérési adatainkat — melyeket általánosan x_i -vel jelölünk (most $i = 1, 2, 3, \dots, 20$) — az alábbi 4. táblázatban tüntettük fel.

66	56	89	63	66	69	71	68	58	69
78	66	64	84	74	76	69	77	74	76

4. táblázat. A hallgatói csoport megmért pulzusszám adatai (minta).




Bár egy ilyen táblázatos forma talán jobban mutat, mint az adatok egyszerű felsorolása, nem nyújt kellő áttekintést az adatok egymáshoz való viszonyáról. Ha adatainkat egy számegegyenes mentén ábrázoljuk (4. ábra), jobban megfigyelhető a „normális” érték körüli ingadozás:



4. ábra. A minta elemei egy számegegyenes mentén ábrázolva.

Itt önkényesen ugyan, de háromféle tartományt különböztethetünk meg: I. sok adat, II. kevés adat, III. nincs adat.

Ha ezt a képet finomítjuk, akkor eljuthatunk a **gyakorisági eloszlás** fogalmához, amit adataink további osztályozása útján nyerhetünk. E feladat elvégzésének érdekében osszuk fel a számegegyenest egyenlő részekre (intervallumokra), és számoljuk meg, hogy az így kapott **osztályokban** hány adat található. Ily módon meghatározhatjuk az egyes osztályokban a gyakoriságokat, illetve a relatív gyakoriságokat. (Megjegyezzük, hogy az osztályokat nem kell feltétlenül egyenlőnek választani, de célszerű.)

 gyakorisági eloszlás
 frequency distribution
 Häufigkeitsverteilung

 hisztogram
 histogram
 Histogramm

Természetesen, mivel az osztályhatárok megválasztása önkényes, ugyanazokból az adatokból többféle gyakorisági eloszlást is készíthetünk. Az alábbi 5. táblázat egyet mutat be a lehetséges esetek közül.

OSZTÁLYHATÁROK	GYAKORISÁG	RELATÍV GYAKORISÁG
$55 \leq x_i < 60$	2	0,10
$60 \leq x_i < 65$	2	0,10
$65 \leq x_i < 70$	7	0,35
$70 \leq x_i < 75$	3	0,15
$75 \leq x_i < 80$	4	0,20
$80 \leq x_i < 85$	1	0,05
$85 \leq x_i < 90$	1	0,05
összesen:	$n = 20$	1,00

5. táblázat. Egy, a mintából képzett gyakorisági eloszlás.

A kapott gyakoriságokat, illetve relatív gyakoriságokat a jobb áttekinthetőség kedvéért oszlop diagrammal szemléltethetjük. Változó osztályszélesség esetén is célszerű a következő ábrázolásmód: **minden osztály fölé olyan téglalapot („oszlopot”) rajzolni, melynek területe arányos az osztályba eső adatok gyakoriságával.** Az így kapott ábrát **hisztogramnak** nevezzük. (Az azonos osztályszélesség választás épp azért előnyös, mert ilyen esetben a téglalap területe és a magassága arányos egymással.)

Ilyen hisztogramokat láthatunk az 5. ábrán. Az első két esetben az osztályszélességek azonosak, csak az osztályhatárok különböznek, a második két esetben az osztályszélességek is különböznek. A hisztogram készítésére tehát nincs túl szigorú előírás, esztétikai szempontjaink azért lehetnek (lásd 1. megjegyzés).

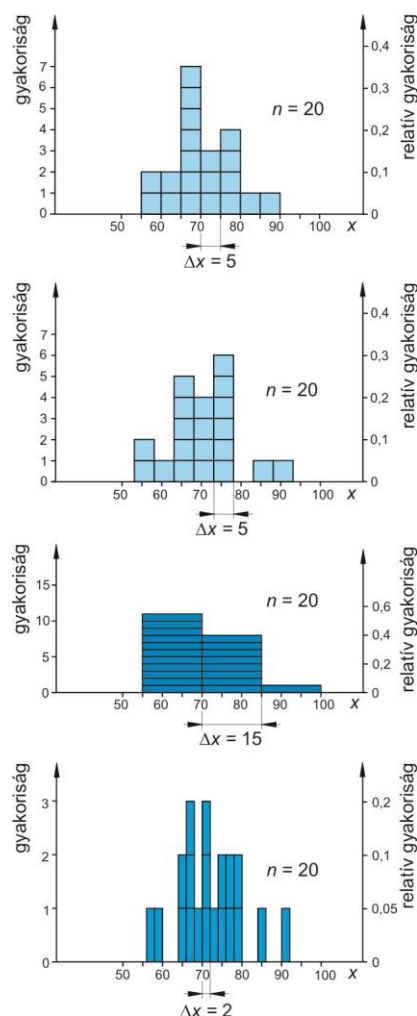
A grafikonok vízszintes tengelyén tehát a változó értéktartományai, a függőlegesen pedig az abszolút, illetve a relatív gyakoriságok vannak feltüntetve. Minden négyzet illetve téglalap egyetlen adatnak felel meg, ezért ezek száma az összes adat számával egyezik meg ($n = 20$). Azt is mondhatjuk, hogy ez éppen az oszlopokat burkoló görbe alatti terület, ami a relatív gyakoriságok szerint (az n elemszámmal való osztás miatt) éppen 1, azaz 100%.

Bár az egyes hisztogramok konkrét alakja jelentős eltérést mutat, mégis találunk közös vonásokat. Megfigyelhetjük például, hogy mindegyik „kipúposodik” a közepe táján, azaz nagyobb az értéknél, de ezen túlmenően még a „szélességük” is hasonló mértékű. Ha az adatok számát növelnénk, az osztályok szélességét pedig csökkentenénk (és ezt akármeddig folytathatnánk), akkor hisztogramjaink durva lépcsős burkoló görbéi egyre jobban kisimulnának és egyetlen folytonos, sima görbébe mennének át (lásd 6. ábra).

A SOKASÁG ELOSZLÁSA, ELMÉLETI ELOSZLÁS

Vizsgáljuk meg közelebbről a 6. ábrán bemutatott tendenciát. Véges elemű sokaság esetén, amennyiben a minta elemszámát (n -t), növeljük akkor előbb-utóbb a sokaság „összes” eleme kiválasztásra kerül ($n = N$). Ilyenkor a **N elemű „minta” eloszlása** az osztályhatárok bizonytalanságától eltekintve **„ugyanolyan” lesz, mint a sokaság eloszlása.** Végtelen elemű, vagy annak tekinthető sokaságnál csak azt mondhatjuk, hogy a minta elemszámának növelése a minta eloszlását a sokaság eloszlásához egyre közelebb viszi. Ilyenkor a populáció eloszlása egy **elméleti eloszlással** írható le.

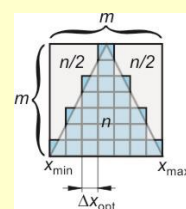
A **populáció eloszlása a változó minden jellemzőjét meghatározza**, azaz megadja, hogy a változó a lehetséges értékeit mekkora valószínűséggel veszi fel. (Ennél többet nem is mondhatunk a változóról.) Ha például kijelölünk valahol a számegegyenesen egy (a, b) intervallumot, akkor az intervallumra eső görbe alatti terület azzal a valószínűséggel egyenlő, hogy egy véletlenül kiválasztott érték éppen az adott (a, b) intervallumba essen. Ha az (a, b) intervallumot olyan helyen választjuk, ahol a görbe kis értékeket vesz fel, akkor ott a görbe alatti terület kicsinyisége miatt a változó lehetséges értékeinek előfordulási valószínűsége is kicsi lesz (7/1. ábra). Ha azonban az (a, b) intervallum a görbe nagyobb értékeinél van kijelölve, akkor ott a terület, és emiatt a valószínűség is nagyobb lesz



5. ábra. A minta alapján készített különböző osztályszélességű hisztogramok. (A legfelső hisztogram készült az 5. táblázat alapján.) Minden téglalap egyetlen adatnak felel meg.

1. megjegyzés:

A hisztogram akkor „esztétikus”, ha nem hézagos, de azért van szerkezete, azaz nincs minden adat egy-két osztályba beszüfölve. Ha egy négyzet alakú diagramban akarjuk az „optimális” hisztogramot megrajzolni, akkor az intervallumok száma körülbelül megegyezik az egy intervallumba eső elemek maximális számával, mindkettőt jelöljük m -mel.



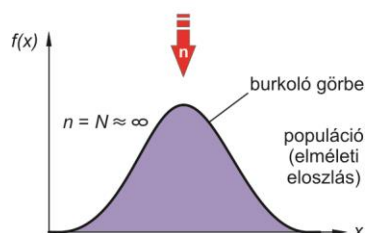
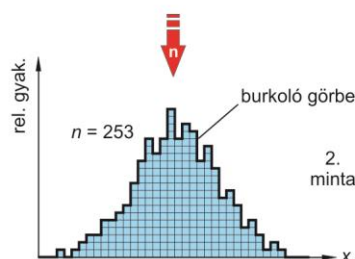
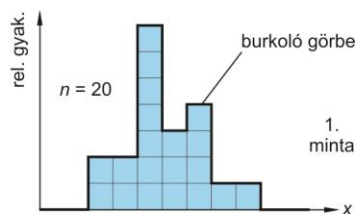
Ekkor egy elem egy négyzet alakú területet foglal el (a hosszúság téglalap helyett). Az ábra alapján az optimális intervallumok száma:

$$m = \sqrt{2n}.$$

Az intervallumok optimális méretét (Δx_{opt}) megkaphatjuk, ha a legnagyobb (x_{\max}) és a legkisebb (x_{\min}) adatok különbségét elosztjuk az optimális intervallumszámmal:

$$\Delta x_{\text{opt}} = \frac{x_{\max} - x_{\min}}{m}.$$

Az 5. ábrán a felső két hisztogram ilyen szempontok szerint készült.



6. ábra. A minta elemszámának növelése és az osztályszélesség csökkentése kisimítja a hisztogramot burkoló görbét.

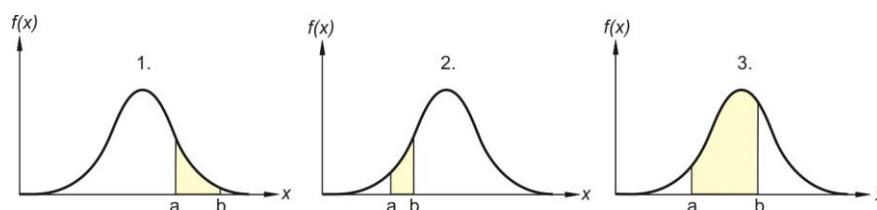
normális eloszlás, Gauss-eloszlás
 normal distribution,
 Gaussian distribution
 Normalverteilung, Gauss Verteilung

2. megjegyzés:

A mintának az alapsokaságra nézve reprezentatívnak kell lennie, vagyis alapfeltétel, hogy a vizsgált jellemző eloszlása a mintában véletlen hibától eltekintve ugyanaz legyen, mint az egész sokaságban. Méréseink, kísérleteink tervezésekor erre is gondolnunk kell.

Ha például egy felmérés során a pajzsmirigy betegségek gyakoriságát akarjuk megállapítani hazánkban, akkor a népesség területi eloszlását figyelembe véve az egész ország területéről kell adatokat gyűjtenünk. Egy régió túlréprezentált jelenléte a mintában hamis eredményre vezethet, mivel például északi megyéink területén az ivóvíz jó-d-szegénysége miatt az ilyen megbetegedések előfordulása sokkal gyakoribb, mint délen.

(7/2. ábra). Ha egy ilyen helyen az intervallum nagyságát növeljük, akkor nyilvánvalóan a görbe alatti terület még nagyobb lesz, ami azt jelenti, hogy nagyobb intervallumba nagyobb valószínűséggel fog beleesni a véletlenül kiválasztott érték (7/3. ábra). Ebből az is következik, hogy ugyanúgy, mint a hisztogramoknál, **a teljes görbe alatti területnek mindig 1-t kell adnia**, hiszen, ha az (a, b) "intervallum" olyan nagy, hogy $-\infty$ -tól $+\infty$ -ig terjed, akkor abban bármilyen tetszőlegesen kiválasztott érték biztosan benne van. (Lásd korábban: a következtetés bizonytalansága és értéke (használhatósága) közötti kapcsolatot.)



7. ábra. Az elméleti eloszlás görbe alatti területének szemléletes jelentése.

Egy fontos dolgot vegyünk azonban észre: nem véletlen, hogy az előbbiek során mindig intervallumról beszéltünk, ugyanis egyetlen kiragadott érték fölött nincs terület (olyan "területről" van szó, amelynek nulla a szélessége). Így folytonos változók esetén nulla annak a valószínűsége is, hogy egy véletlenül kiválasztott érték pontosan egy előre megadott értékkel legyen egyenlő. Ez az oka annak, hogy ilyenkor elvileg az összes mérési eredményünk különbözik egymástól. Az, hogy ez a gyakorlatban mégis így, azzal magyarázható, hogy valójában minden mérési adat már a leolvasáskor egy intervallumot takar, ami annak az egyszerű ténynek a következménye, hogy a gyakorlatban használt számaink mindig véges tizedes törtek. Így a leírt utolsó számjegy mindig kerekített érték. (Lásd korábban: folytonos és diszkrét jellemzők.)

Az elméleti eloszlás tehát az összes lehetséges adat, vagyis a populáció jellemzésére szolgál, a hisztogram pedig csak egy ebből vett minta elemeire, a konkrét mérési adatokra vonatkozik.

A STATISZTIKA ALAPTÉTELE

Most gondoljuk meg még egyszer, hogy hogyan nyertük az elméleti eloszlást: úgy, hogy a minta elemszámát, azaz méréseink, adataink számát növeltük. A matematikai statisztika alaptételének szemléletes tartalma éppen az, hogy **nagy minták esetén a tapasztalati eloszlásfüggvény** (azaz a hisztogram burkolója) **nagyon jól megközelíti az elméleti eloszlásfüggvényt**. Azt reméljük tehát, hogy **minél gyakoribb egy értékcsoport előfordulása a mintában, annál valószínűbb a megjelenése az alapsokaságban is**.

A matematikai statisztika segítségével egy populáció vagy alapsokaság valamely jellemzőjét úgy határozzuk meg, hogy a sokaságnak csak bizonyos számú (lehetőleg kevés) elemét vizsgáljuk meg. A mintavétel feladata a megvizsgálásra szánt elemek (a minta) kijelölése úgy, hogy belőlük az egész sokaságra megbízható következtetéseket vonhassunk le. Ezt általában úgy érhetjük el, hogy **a mintaelemeket véletlenszerűen választjuk ki** (lásd 2. megjegyzés). (A mintavételnél sok esetben orvosi szempontokra is tekintettel kell lennünk.)

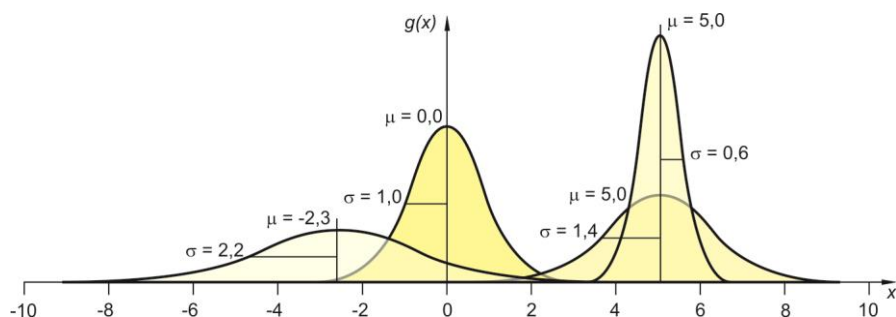
A NORMÁLIS VAGY GAUSS-ELOSZLÁS

Az elméleti eloszlás a vizsgált változótól függően **különböző alakú lehet**, de az esetek többségében **egyetlen csúccsal rendelkező, szimmetrikus, harang alakú görbe** — aminek okára még visszatérünk —, és amit normális vagy Gauss-eloszlásnak nevezünk. (Ilyen eloszlást tüntettünk fel már a 6. és 7. ábrán is.) Az eloszlást leíró függvény:

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (1)$$

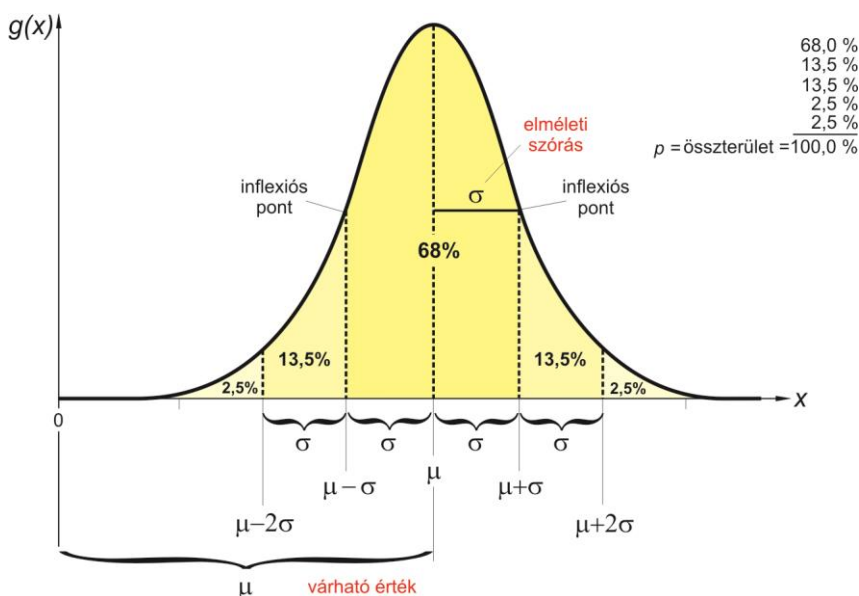
Ez a kifejezés elég bonyolultnak tűnik, de valójában nem más, mint az $f(x) = e^{-x^2}$

függvény állandókkal megtűzdtelt változata. Normális vagy Gauss-eloszláson tehát nem egyetlen eloszlást értünk, hanem épp a konstansok miatt egy egész eloszláshalmazt (lásd 8. ábra): az egyes eloszlások tengely menti elhelyezkedése, szélessége és magassága más és más, csak az alakjuk hasonló.



8. ábra. Néhány eltérő helyzetű (μ), szélességű (σ), és magasságú Gauss-eloszlás.

A normális eloszlásra jellemző haranggörbe (vagy Gauss-görbe) szárai mindkét irányban a végtelenbe nyúlnak, az x -tengelyhez azonban annyira közel, hogy külön nem is rajzolhatók. A görbe alatti terület szükségképpen 1. (lásd korábban: elméleti eloszlás.)



9. ábra. A Gauss-eloszlás haranggörbéje és paraméterei.

A fenti képletben szereplő μ és σ állandók az eloszlás paraméterei. Ezek mutatják meg, hogy a végtelen sok lehetséges normális eloszlásból éppen melyikről van szó. A μ paraméter a várható érték, ami megadja az eloszlás maximumának helyét az x -tengelyen. A σ paraméter az elméleti szórás, ami az eloszlás szélességét jellemzi: a haranggörbe szélessége a magassága felénél mérve körülbelül 2σ -val egyenlő (pontosabban a görbe úgynevezett inflexiós pontjai éppen σ távolságra esnek a μ értéktől) (9. ábra). Mindezek alapján a normális eloszlás szokásos jelölése $N(\mu, \sigma)$.

A haranggörbe alakja és a paraméterek közötti kapcsolatról ennél többet is tudunk mondani: a $\mu - \sigma$ és $\mu + \sigma$ értékek között van a görbe alatti terület mintegy kétharmada (kb. 68%-a), $\mu - 2\sigma$ és $\mu + 2\sigma$ között kb. 95 %-a, a $\mu - 3\sigma$, $\mu + 3\sigma$ intervallumon kívül pedig mindössze csak 2 ezrelék jut. A görbe gyakorlatilag egy 6σ hosszúságú szakaszon helyezkedik el a várható érték körül. Csak megjegyezzük, hogy a görbék magassága valójában lényegtelen paraméter, σ -val fordítottan arányos, ami a rögzített (egységnyi) görbe alatti terület következménye.

A könnyebb kezelhetőség kedvéért ki szokás emelni egy speciális eloszlást a végtelen sok normális eloszlás közül: ez a $\mu = 0$, $\sigma = 1$ paraméterekkel jellemzett



Carl Friedrich Gauss (1777-1855)
német matematikus.

	várható érték
	expected value
	Erwartungswert
	elméleti szórás
	theoretical standard deviation, SD
	theoretische Streuung

standard normális eloszlás, a bevezetett jelölés szerint $N(0, 1)$. (a 8. ábrán balról a második görbe).

A normális eloszlás kiemelkedő jelentőségére a valószínűségszámítás egyik nevezetes tétele, a **centrális határeloszlás tétel** mutat rá. Eszerint a sok apró, egymástól független hatás eredményeképpen kialakult értékek normális eloszlást követnek. Ezzel magyarázható, hogy a természetben előforduló változók jelentős része normális eloszlású, ezért a továbbiakban főként csak ezzel az eloszlástípussal foglalkozunk.

„Orvosi” példaként megemlíthetjük, hogy Gauss eloszlású a **testmagasság**, vagy a **vérnyomás** is. Magyarországon a felnőtt férfiak testmagasságának eloszlása napjainkban (cm-ben kifejezve) nagyjából az $N(171, 7)$ eloszlásnak felel meg. Iskoláskorú fiúk diasztolés (alsó érték) vérnyomása Hgmm-ben az $N(58, 8)$, dohányzó fiatal férfiaké pedig az $N(84, 10)$ eloszlással adható meg.

A testmagasság példájánál maradva, ahol $3\sigma = 21$ (cm) azt mondhatjuk, hogy a felnőtt férfiak nagy többségének (több mint 99%-ának) testmagassága 150 és 192 cm között van. Vannak persze 2 m magas férfiak is, de ez nagyon nem tipikus érték. A nagyon tipikus a 170 cm körüli magasság, de azt is megfigyelhetjük, hogy viszonylag gyakori a 160, illetve 180 cm-es férfi is. Ez is azt mutatja, ami az **élővilág** egyik fontos jellegzetessége, hogy **vannak ugyan tipikus értékek, de a sokféleség**, az egyedek közötti **különbözőség legalább olyan fontos jellemző**.

Ha a másik példát tekintjük, első ránézésre látható, hogy a dohányzó fiatal férfiak vérnyomása nemcsak magasabb, de nagyobb a szórása is ($10 > 8$). Ha azonban kiszámítjuk a **relatív** (elméleti) **szórásokat**, azaz a σ/μ hányadosokat, akkor ebben a paraméterben már fordított a helyzet ($10/84 \approx 0,12 < 8/58 \approx 0,14$). Sok esetben a relatív szórás, amit százalékban is megadhatunk ($(\sigma/\mu) \cdot (100\%)$) többet árul el az abszolút szórásnál. A szóródás tehát lehet kicsi vagy nagy, de talán még fontosabb, hogy mihez képest, így μ és σ meghatározása egyaránt fontos célunk. Könnyen belátható azonban, hogy a paraméterek „pontos” meghatározása igen fáradságos, esetenként — például végtelen elemű alapsokaság esetén — lehetetlen feladat. Így megelégszünk azzal, ha meg tudjuk becsülni őket.

A PARAMÉTEREK BECSLÉSE, A MINTA STATISZTIKAI JELLEMZŐI

Láttuk, hogy a Gauss-görbét két paramétere (μ és σ) egyértelműen jellemzi, ezért célunk az, hogy a paramétereket egy minta alapján történő becsléssel minél jobban megközelítsük.

A **μ várható értéket** leggyakrabban az **átlaggal** (\bar{x}) **közelítjük**, amely az adatokból (a minta elemeiből) képzett **számítási közép**:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}. \quad (2)$$

Ez a **legstabilabb középérték** (normális eloszlás esetén, lásd 3. megjegyzés), amely az összes mintaelem számszerű értékére támaszkodik, és így a minta változására legkevésbé érzékeny. A többi középérték közül kiemelkedő jelentőségűvé még az teszi, hogy ez az a szám, amelytől az adatok eltéréseinek összege éppen 0-val egyenlő, ugyanis a pozitív és negatív irányú eltérések éppen kiejtik egymást:

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = \sum_{i=1}^n x_i - n\bar{x} = 0. \quad (3)$$

Ennek alapján azt is mondhatjuk, hogy az átlag a számegyenesen, az adatok között a „súlypontnak” megfelelő helyen van.

A **σ elméleti szórás** leggyakrabban a **tapasztalati szórással** (s) **közelítjük**, amely az átlagtól való átlagos eltérésekkel kapcsolatos definíció:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}. \quad (4)$$

3. megjegyzés:

A μ várható érték becslésére kínálkozó további lehetőségek:




1. A **modus** az az érték, amelyikből a „legtöbb” van a mintában, vagyis a gyakorisági eloszlásból készített hisztogram maximumának megfelelő érték. Mivel — mint láthattuk — a gyakorisági eloszlás nem egyértelmű (függ az osztályok megválasztásától), ezért a modus sem az. Különösen kevés adat esetén nem igazán jó jellemző. (A 4. táblázatban megadott minta modusa az 5. legfelső ábra alapján: 65 és 70 között van.)




2. A **medián** a nagyság szerint sorba állított adatok közül a középső, vagy a középső kettő átlaga. Vegyük észre, hogy ennek az értékét a szélső adatok egyáltalán nem befolyásolják. Éppen ezért amikor a szélsőséges adatok például mérés technikai okokból nagyon megbízhatatlanok, akkor ez a legjobb becslése a várható értéknek. (A 4. táblázatban megadott minta mediánja a 4. ábráról leolvasható: 69.)

(Gauss-eloszlás esetén, nagy elemszámú mintákra igaz, hogy az




átlag \approx modus \approx medián.

Vannak azonban olyan, például nem szimmetrikus eloszlások is, amelyekre ez nem teljesül.)

 átlag
 mean, average
 Durchschnitt

 tapasztalati szórás
 empirical standard deviation, SD
 empirische Streuung

 variancia
 variance
 Varianz

 szabadságfok
 degree of freedom
 Freiheitsgrad

(Használatos még az s_x jelölés is, ahol a félreértések elkerülése végett a változót indexként feltüntetjük.) Ennek négyzete, a tapasztalati **szórásnégyzet**, más néven **variancia**:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}. \quad (5)$$

Mivel a (4) illetve (5) képlet számlálójában szereplő négyzetes (idegen szóval kvadrátikus = „Quadrátikus”) kifejezés, vagy hozzá nagyon hasonló, még később is fog szerepelni, ezért külön jelölést (Q) vezetünk be rá, továbbá kis átalakítással a kiszámítására alkalmasabb alakra hozzuk:

$$Q_x = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}) = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}. \quad (6)$$

A (5) képlet nevezője (az $(n-1)$ kifejezés) a **szabadságfok**. Ez a paraméterek becslésével kapcsolatos matematikai statisztikai fogalom szoros összefüggésben van az adatok számával, de láthatóan nem mindig egyenlő vele. A számítás kezdetekor n adatból álló minta szabadságfokainak száma n . Ha egy mintából valamely paramétert úgy kell becsülnünk, hogy ahhoz ugyanebből a mintából már előzetesen meghatározott paramétereket fel kell használnunk, akkor annyit kell levonnunk az eredetileg n szabadságfokból, ahány korábbi paramétert a becslés közben felhasználtunk. Mivel a tapasztalati szórás becslésénél az n adaton kívül az ugyanabból a mintából már meghatározott átlagot is fel kell használnunk, ezért a tapasztalati szórás szabadságfokainak száma $n-1$. Bonyolultabb esetekben a szabadságfokot külön képlet segítségével kell kiszámítanunk. (A 4. megjegyzésben a 4. táblázatban megadott minta néhány fontos statisztikai jellemzőjét foglaltuk össze.)

A paramétereknek a mintából kiszámított, becsült értéke és a „valódi” értéke között általában több-kevesebb eltérés mutatkozik. Ez az eltérés a becsült paraméter hibája, amire még visszatérünk. Ezt a hibát **pontatlanság** és **torzítás** okozhatja. **Pontatlanságnak** tekintjük a hibát, ha az **eltérés** a valódi értéktől **pozitív és negatív irányban véletlenszerű**. **Torzításnak** tekintjük a hibát, ha a **paraméter becsült értéke szisztematikusan kisebb vagy nagyobb** a „valódi” értéknél. Míg a pontatlanság számszerűen mérhető, addig a **torzítás többnyire nem mérhető** becslési hiba.

Az átlag és a tapasztalati szórás fenti definícióira igaz a következő állítás (amennyiben a torzításoktól eltekintünk): Ha a minta elemszáma végtelenhez tart (\rightarrow), akkor az átlag a várható értéket, a tapasztalati szórás pedig az elméleti szórást közelíti egyre nagyobb pontossággal, azaz

$$n \rightarrow \infty \text{ esetén } \bar{x} \rightarrow \mu \text{ és } s \rightarrow \sigma. \quad (7)$$

Az s **tapasztalati szórás azt mutatja meg, hogy az adatok átlagosan mennyire térnek el az átlagtól**, azaz — a Gauss-eloszlásnál (9. ábra) ismertettekhez hasonlóan — a **minta elemeinek kb. 68%-a található az $\bar{x} \pm s$ intervallumban, kb. 95%-a az $\bar{x} \pm 2s$ és több mint 99%-a az $\bar{x} \pm 3s$ intervallumban.**

A nagy elemszámú ($n \approx 1000$) mintából számolt $\bar{x} \pm k s$ tartományt, amelyben a minta elemeinek pontosan 95%-a található ($k \approx 2$), **referencia tartománynak** vagy **normál tartománynak** is szokás nevezni, amit elsősorban a laboratóriumi diagnosztikában alkalmazunk. (Egyes orvosi alkalmazások során a normál tartomány ettől eltérő értelmezése is előfordul). Ezt a gyakorlatban úgy lehet felhasználni, hogy amennyiben egy laboratóriumi adat a normál tartományon belüli, akkor arról 95%-os bizonyossággal azt mondhatjuk, hogy nem utal kóros elváltozásra (lásd 5. megjegyzés).

Ilyen esetben a torzítás nem jelent problémát, mert amennyiben minden adat szisztematikusan el van tolódva, akkor ugyanennyivel a referencia tartomány is eltolódik. Ezt esetenként meg is figyelhetjük, ha különböző laboratóriumokban

4. megjegyzés:

Még egyszer az adatok, $n = 20$:

	x_i
x_1	66
x_2	56
x_3	89
x_4	63
x_5	66
x_6	69
x_7	71
x_8	68
x_9	58
x_{10}	69
x_{11}	78
x_{12}	66
x_{13}	64
x_{14}	84
x_{15}	74
x_{16}	76
x_{17}	69
x_{18}	77
x_{19}	74
x_{20}	76
$\sum x_i = 1413$	

Az **átlagos** pulzusszám:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1413}{20} \approx 71 \text{ (1/perc)}$$

(egészre kerekítve).

A **kvadrátikus összeg**:

$$Q_x = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 101075 - \frac{1413^2}{20} = 1246,55$$

A **variancia**:

$$s^2 = \frac{Q_x}{n-1} = \frac{1246,55}{19} \approx 66 \text{ (1/perc)}^2$$

(egészre kerekítve).

A **szórás**:

$$s = \sqrt{\frac{Q_x}{n-1}} = \sqrt{\frac{1246,55}{19}} \approx 8 \text{ (1/perc)}$$

(egészre kerekítve).

A **szabadságfok**: 19.

5. megjegyzés:



Egy ilyen állítás értékére talán jobban rávilágíthatunk a következő példával. Annak a valószínűsége, hogy egy dobókockával 6-ost dobunk 1/6, kb. 17%, tehát annak a valószínűsége, hogy ne dobunk 6-ost 5/6, ami ($5/6 \approx 0,83$) kb. 83%-nak felel meg.



Ha szabályos dobókocka helyett egy ikozaédert (20 lapú szabályos testet) használnánk, aminek minden lapja meg van számozva, akkor annak a valószínűsége, hogy 20-ast dobunk $1/20 = 0,05 = 5\%$. A nem 20-as dobás valószínűsége pedig $19/20 = 0,95 = 95\%$.

Eszerint az az állítás, hogy „egy laboratóriumi adat a normál tartományon belül” nagyjából azzal ekvivalens bizonyosságú, hogy az említett ikozaéddal elvégzett dobás eredménye nem 20.

elvégzett vizsgálatok eredményét hasonlítjuk össze. A referencia tartományok ugyanarra a változóra nézve kissé eltérhetnek egymástól. Ennek oka például az lehet, hogy az alkalmazott mérési módszerek, illetve a mérőberendezések különbözőek.

KONFIDENCIA INTERVALLUMOK, A BECSÜLT PARAMÉTER VÉLETLEN HIBÁJA, PONTATLANSÁGA

Előljáróban még egyszer hangsúlyozzuk, hogy a becsült paraméter hibája lehet torzítás is, ami általában nem mérhető, így az alábbiakban a hiba csak a pontatlanságot jelenti, azaz a **véletlen hibát**. Az átlag az előzőek szerint ugyan n növelésével egyre jobban közelíti a „meghatározni” kívánt értéket (lásd (7) kifejezés), de arra a kérdésre, hogy egy adott n elemszámú minta esetén mennyire tér el a populációra jellemző várható értéktől, azaz, hogy **mekkora az átlag hibája**, még nem kaptunk választ.

6. megjegyzés:

A pulzusszáma vonatkozó mintából (4. táblázat) már kiszámítottuk az átlagot (71 (1/perc)) és a szórást (8 (1/perc)). A standard hiba

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{8}{\sqrt{20}} \approx 2 \text{ (1/perc);}$$

a **hibakorlát** pedig (kb. 95% konfidencia szintnél):

$$\bar{x} \pm 2s_{\bar{x}} = 71 \pm 4 \text{ (1/perc).}$$

(mindenhol egészre kerekítve)

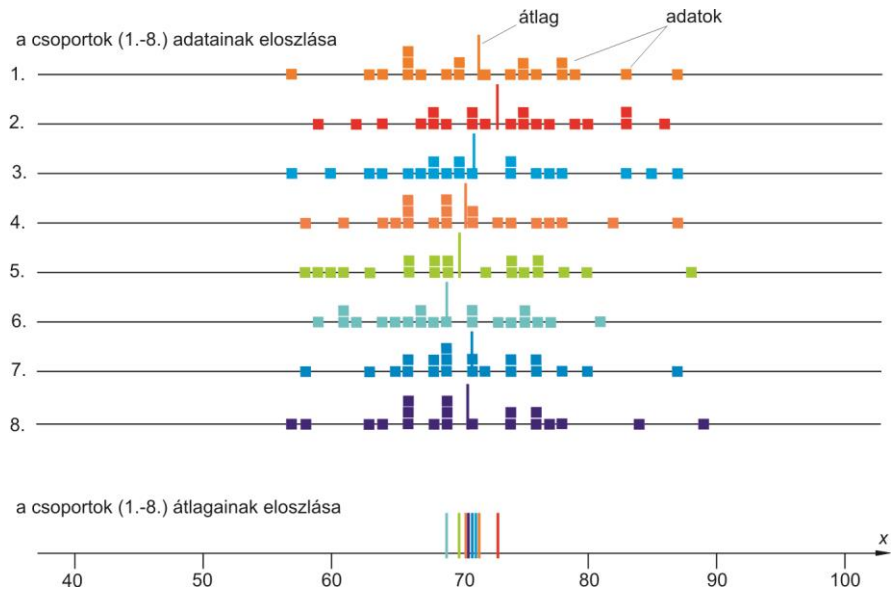
A mérést az alábbiak szerint jellemezhetjük:

„az elvégzett mérések alapján kb. 95 % bizonyossággal állíthatjuk, hogy a populáció várható értéke a 67 - 75 (1/perc) tartományban található” (11. ábra).

Ha ennél nagyobb bizonyossággal, egyúttal nagyobb pontossággal akarjuk határok közé vonni a pulzusszám várható értékét, akkor növelnünk kell a minta elemszámát.

Meddig növeljük a minta elemszámát?

Általános szabály nincs, de a konkrét esetről a következőket mondhatjuk. Mivel a pulzusszámot egészre kerekítve szokás megadni ezért a pontosság ± 1 (1/perc)-en túli növelésének nincs sok értelme. A bizonyosság kérdése nem annyira egyértelmű, de ritkán van szükségünk 99%-os, vagy pláne annál nagyobb bizonyosságra. Eszerint (a biztos hibakorlátot választva) addig kell növelnünk n -t, amíg a hiba le nem esik annyira, hogy a $3s_{\bar{x}} \leq 1$ (1/perc) feltétel teljesüljön.



10. ábra. Adatok és átlagaik: nyolc 20 fős tanulócsoport hallgatóinak pulzusszámát és azok átlagait tüntettük fel. Figyeljük meg, hogy az átlagok jóval kevésbé szóródnak, mint az adatok.

Az előző részben azt is említettük, hogy az átlag, mint középérték a minta változására nem nagyon érzékeny, hiszen a számolás az összes mintaelem figyelembe vételével történik, így főleg nagyobb elemszámú minták esetén egy-egy adatnak csak kevés módosító szerep jut. Ennek az az eredménye, hogy a különböző mintákból számolt átlagok csak kevésbé térnek el egymástól (lásd 10. ábra). Azt is mondhatjuk, hogy a minták átlagai ($\bar{x}_1, \bar{x}_2, \dots, \bar{x}_i$) jóval kevésbé „szóródnak” a várható érték körül, mint az adatok. Ezt a „szóródást” fejezi ki a **standard hiba** (más néven az átlag szórása):

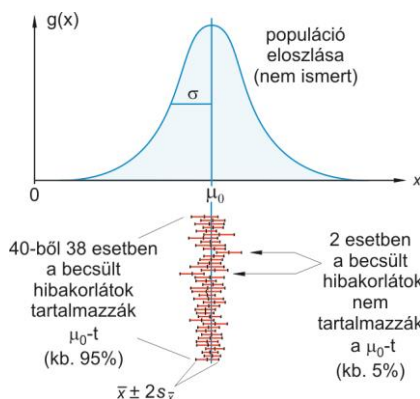
$$s_{\bar{x}} = \frac{s}{\sqrt{n}}. \quad (8)$$

Ennek segítségével a végső eredményt a következőképpen szokás megadni:

$$\bar{x} \pm k \cdot s_{\bar{x}}. \quad (9)$$

Végeredményként tehát mindig egy intervallumot adunk meg, azaz pozitív és negatív irányban egy-egy határoló értéket, azzal a céllal, hogy **ezek a keresett várható értéket közrezárják**. Problémánk most már csak az, hogy hol vonjuk meg a két határt, azaz mekkora legyen k értéke? A határok megvonása ugyanis nem egyértelmű. Ha nagyon tág határokat adunk meg (k nagy), akkor várható ugyan, hogy ezek a valódi értéket közrefogják, és az a következtetésünk, hogy a valódi érték a két határ között fekszik, feltehetően nem lesz téves (lásd korábban: statisztikai következtetés). A nagyon tág határoknak azonban többnyire kevés a gyakorlati haszna. Ha ezzel szemben szűkítjük a határokat, fokozottan növekszik annak a kockázata, hogy a határokat tévesen adjuk meg, azok már nem fogják

standard hiba
standard error
Standardfehler



11. ábra. A mintából számolt hibakorlát 95% bizonyossággal tartalmazza a populáció várható értékét.

közre a várható értéket, következtetésünk biztonsága, bizonyossága ezért csökken. Így tehát egyrészt a bizonyosság fokozása, illetve a téves következtetés kockázatának csökkentése tág határokat, másrészt a szakmai értelmezhetőség szűk határokat igényel.

A statisztikai módszerek lehetővé teszik a bizonyosság mértékének rögzítését. Ez azt jelenti, hogy két olyan határoló értéket állapítunk meg, melyek a valódi értéket meghatározott bizonyossággal (konfidencia) fogják közre. Az így megállapított határoló értékeket **konfidencia határoknak**, a két konfidencia határ által közrefogott szakaszt **konfidencia intervallumnak**, a **bizonyosság mértékét** pedig **konfidencia szintnek** nevezzük.

Bár k értéke a minta elemszámától (szabadságfoktól) is függ, **nagy minták esetére** a következőket mondhatjuk: ha $k = 1$, a konfidencia szint kb. 0,68, ha $k = 2$, akkor kb. 0,95, ha $k = 3$, akkor nagyobb, mint 0,99 (lásd 6. táblázat).

konfidencia szint (kb.)	68%	95%	99%
konfidencia intervallum	$\bar{x} \pm s_{\bar{x}}$	$\bar{x} \pm 2s_{\bar{x}}$	$\bar{x} \pm 3s_{\bar{x}}$
		hibakorlát	biztos hibakorlát

6. táblázat. Konfidencia szintek és intervallumok.

Ezek ismeretében a „MEKKORA?” kérdésre a választ, (illetve a mérés végeredményét) a (9) szerint kell megadnunk (lásd 6. megjegyzés, 11. ábra.).

Az átlag szórásának definíciójából (8) kitűnik, hogy a hiba az adatok számának növekedtével egyre csökken:

$$n \rightarrow \infty \quad \text{esetén} \quad s_{\bar{x}} \rightarrow 0, \quad (10)$$

ami rámutat a sokszori mérés értelmére. Így az adatok számának növelésével elérhető az, hogy rögzített konfidencia szint mellett a konfidencia intervallum tetszőleges mértékben csökkenjen (lásd 6. megjegyzés).

GRAFIKUS ADATFELDOLGOZÁS

Az adatok közötti kapcsolat grafikus bemutatásáról már ejtettünk néhány szót. Itt a grafikonok különböző transzformációit mutatjuk be.

Az egyenes egyszerűségéből adódó könnyebbségek (például az, hogy a mérési pontok egyenes vonalzóval „összeköthetők”) olyan nagymértékűek, hogy még ott is egyenessel szeretnénk dolgozni, ahol a változók közötti kapcsolat nyilvánvalóan nem lineáris. Ilyen esetekben olyan transzformációkat hajtunk végre adatainkon, hogy a transzformált adatok végül egyenest határozzanak meg. Így pl. az

$$y = a \cdot e^{bx} \quad (11)$$

exponenciális függvény **logaritmikus transzformáció** után

$$\lg y = (b \cdot \lg e) \cdot x + (\lg a) \quad (12)$$




alakú lesz, ahol a **lg y** és az **x** közötti összefüggés már lineáris. Az egyenes meredeksége ($b \lg e$), tengelymetszete pedig ($\lg a$) lesz. Ha egy speciális, **lin - log** koordinátarendszerű **milliméterpapírt** használunk, az y értékek logaritmalását nem kell elvégeznünk, hiszen a logaritmikus beosztású y tengely ezt „automatikusan” megvalósítja (lásd 12. felső ábrák). Vagy az




$$y = a \cdot x^b \quad (13)$$

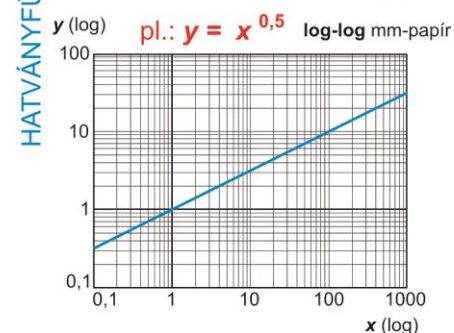
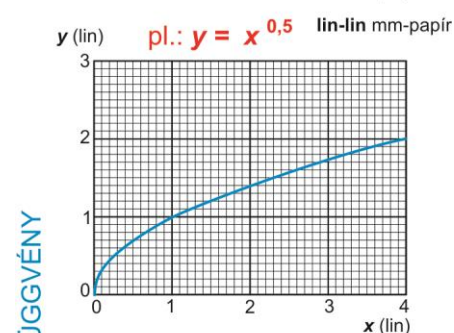
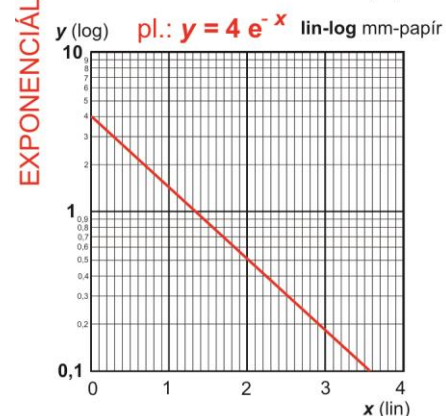
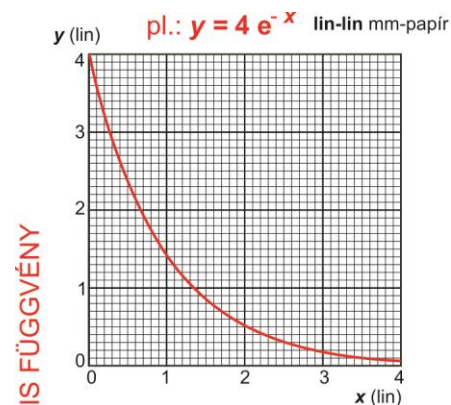
hatványfüggvény, szintén logaritmikus transzformáció után,

$$\lg y = (b) \cdot \lg x + (\lg a) \quad (14)$$

alakú lesz, ahol a **lg y** és a **lg x** között van lineáris összefüggés. Az egyenes meredeksége (b), tengelymetszete pedig ($\lg a$) lesz. A logaritmalás elvégzését itt is elkerülhetjük, egy másik, speciális, **log - log milliméterpapír** használata esetén.

	konfidencia intervallum
	confidence interval
	Konfidenzintervall

	konfidencia szint
	confidence level
	Konfidenzniveau



12. ábra. Az exponenciális, ill. a hatványfüggvény „kiegyenesítése” lin-log, ill. log-log mm-papír segítségével.

Ilyenkor az x illetve az y értékeket logaritmálás nélkül, a tengelyek logaritmikus értékbeosztásai szerint ábrázoljuk (lásd 12. alsó ábrák).

(Megjegyzés. A számítógépek világában az adatok ábrázolása, illetve azok transzformációja egyszerűen elvégezhető. Egyetlen parancs kiadásával a tengelyek átskálázhatók vagy szükség esetén logaritmikus beosztásúvá alakíthatók.)

LINEÁRIS REGRESSZIÓ

A legegyszerűbb görbéhez, az egyeneshez papíron, ceruzával, vonalzóval szubjektív módon könnyen eljuthatunk, de a mérési pontok hibájából származó bizonytalanságainkat, kételyeinket a görbekihúzás korábbi kvalitatív elveivel nem tudjuk mindig eloszlatni. Ezért tovább él bennünk az a kérdés, hogy **hogyan találjuk meg a pontjainkra valóban legjobban illeszkedő egyenest.**

Tudjuk, hogy az egyenes egyenlete

$$y = a \cdot x + b, \quad (15)$$

amelyben a az egyenes meredeksége, b pedig a tengelymetszet, azaz y -nak azon értéke, amelynél az egyenes az y tengelyt az $x = 0$ helyen metszi. E **két paraméter az egyenest egyértelműen jellemzi**. A feladat tehát az, hogy határozzuk meg a mérési pontjainkra legjobban illeszkedő egyenes a^* és b^* -gal jelölt paramétereit. Első lépésként azt kell megmondanunk, hogy **mit jelent** az, hogy egy egyenes a **legjobban illeszkedő**.

Tegyük fel, hogy az x és y mennyiségekre vonatkozóan rendelkezünk 4 mérési ponttal (lásd 13. a. ábra). Feltesszük továbbá azt is, hogy a mérés során az x_i -k „pontos” (hiba nélküli), előre beállított értékek, tehát (mérési) hiba csak az y_i -ket terheli.

Húzzunk a mérési pontjainkon át egy tetszőleges — a_1, b_1 paraméterekkel adott — egyenest ($y = a_1 \cdot x + b_1$) és határozzuk meg a pontoknak az egyenestől függőlegesen mért távolságait (lásd 13. b ábra, függőleges vonalak).

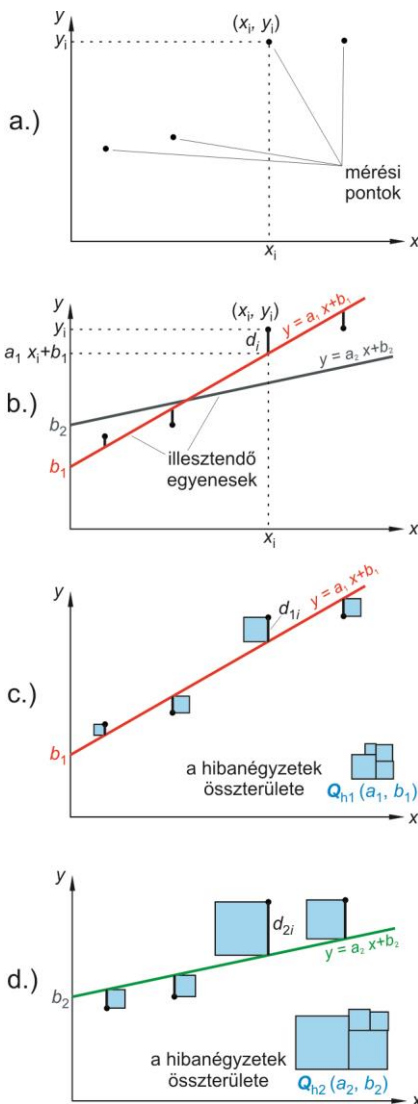
Egy kiszemelt (x_i, y_i) pont x tengelytől mért távolságát az y_i koordináta adja meg, ugyanakkor az x_i koordinátahoz tartozó egyenesbeli pont x tengelytől mért távolságát az egyenes egyenletébe való behelyettesítéssel az $(a_1 \cdot x_i + b_1)$ érték adja meg. A kettő különbsége a **pont és egyenes függőlegesen mért távolsága** $d_{1i} = (y_i - (a_1 \cdot x_i + b_1))$. Ha a pont az egyenes fölött van, akkor ez a kifejezés pozitív, ha alatta, akkor negatív. Az így definiált távolságot a többi pontra hasonlóképpen kiszámíthatjuk, majd vegyük ezek négyzeteit.

A tapasztalati szórás bevezetésénél szereplő négyzetes kifejezés (6) mintájára adjuk össze az összes pontra kiszámított ilyen **távolságnégyzetet** (lásd 13. c. ábra, kék területek) és az összeget jelöljük Q_{h1} -gyel. Ha most behúzzunk egy másik — mondjuk a_2, b_2 paraméterekkel megadott — egyenest, akkor ugyancsak kiszámíthatók ettől a másik egyenestől mért függőleges távolságok is (d_{2i}), és végeredményként egy másik Q_{h2} -vel jelölt négyzetes összeget kapunk (lásd 13. d. ábra). Megfigyelhetjük, hogy a pontok (egyenestől mért) „szétszórtsága” akkor nagyobb, amikor Q_h is nagyobb ($Q_{h2} > Q_{h1}$).

Mivel az (x_i, y_i) pontok változatlanok, ezért a Q_h változását csak az a, b paraméterek változása okozza. A fenti eljárással (hozzárendeléssel) egy olyan függvényt adtunk meg, ahol a két független változó a és b , a függő változó pedig a Q_h :

$$Q_h(a, b) = \sum_{i=1}^n [y_i - (a x_i + b)]^2. \quad (16)$$

A kapott $Q_h(a, b)$ kétváltozós függvény — a távolságnégyzetek miatt — mindkét változójától négyzetesen függ, ami azt jelenti, hogy egy olyan gödörhöz hasonló felülettel reprezentálható, amelynek mindkét tengelyirányú síkmetszete ($a =$ állandó, illetve $b =$ állandó) parabola (14. ábra).



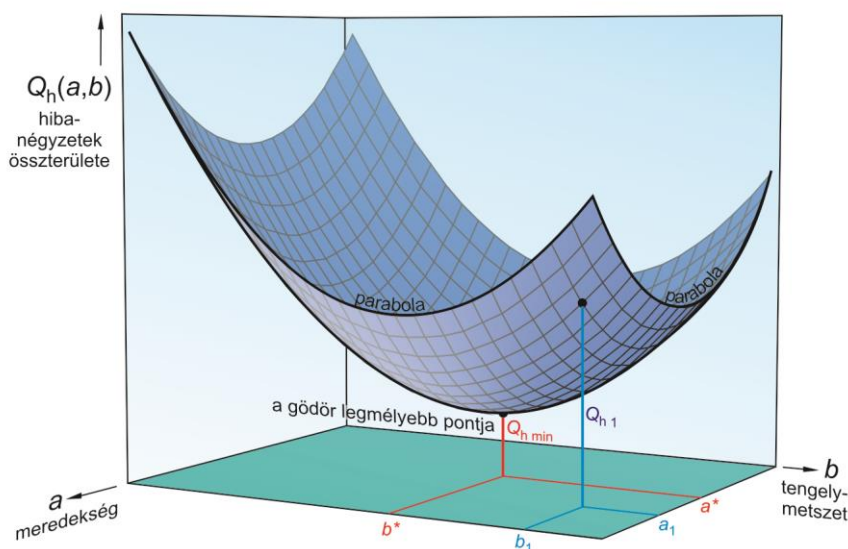
13. ábra. A mérési pontokra legjobban illeszkedő egyenes keresése.

lineáris regresszió
linear regression
lineare Regression

Néhány további feltétel teljesülése esetén (például a mérési hibák az egyes pontokban legyenek egymástól függetlenek) a pontokhoz **legjobban illeszkedő** egyenes az, amelyre nézve az említett **távolság négyzetösszeg**, azaz Q_h **minimális**.

Tehát azt az (a^*, b^*) értékpárt keressük, ahol az imént definiált függvény (16) értéke a legkisebb. Még egyszerűbben: meg kell keresnünk a gödör (14. ábra) legmélyebb pontjának (a^*, b^*) koordinátáit. A fenti módon illesztett egyenest **regressziós egyenesnek**, az illesztési módszert **legkisebb négyzetek módszerének** vagy általánosabban **lineáris regresszióknak** szokás nevezni.

Megjegyezzük, hogy a **regresszió** szó a „visszahatásra”, utal, azaz a mérési pontokból a mennyiségek között fennálló kapcsolatra való következtetést fejezi ki. (A „kapcsolat” általánosan továbbra sem jelent feltétlenül oksági kapcsolatot.)



14. ábra. A hiba függése az a és b paraméterektől. A gödör legmélyebb pontján a hibanégyzetek összege minimális.

A minimum keresés elvégzése után azt kapjuk, hogy

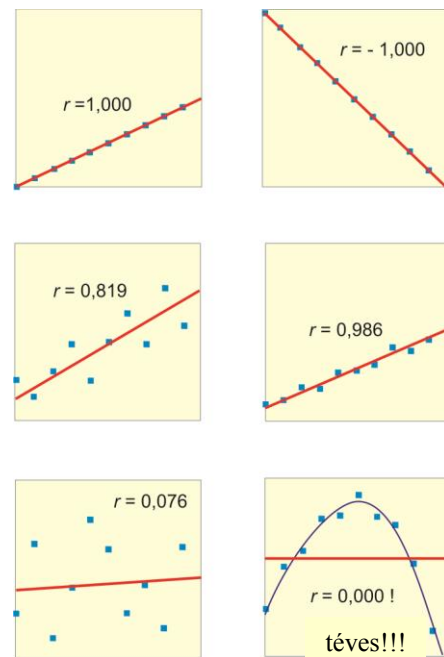
$$a^* = \frac{Q_{xy}}{Q_{xx}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \text{ vagy } a^* = \frac{s_{xy}^2}{s_x^2}, \quad (17)$$

$$b^* = \bar{y} - a^* \bar{x} = \frac{\sum_{i=1}^n y_i}{n} - a^* \frac{\sum_{i=1}^n x_i}{n}, \quad (18)$$

ahol Q_{xx} , Q_{xy} a (6)-ban bevezetett jelölésnek felel meg,

- $s_{xy}^2 = Q_{xy} / (n-1)$ a **kovariancia**,
- s_x^2 az x **variáciája**,
- \bar{x} , \bar{y} pedig a megfelelő **átlagokat** jelöli.

A fenti képletek alapján tetszőleges (x_i, y_i) mért érték párok esetén megadható a legjobban illeszkedő egyenes ($y = a^* x + b^*$), még akkor is, ha a pontok szemmel láthatóan nem egyenesen, hanem valamilyen görbe mentén helyezkednek el.

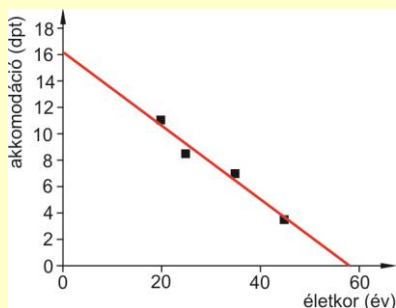


15. ábra. Néhány példa a korrelációs együttható értékére.

7. megjegyzés:

Az alábbi táblázatban a szem akkomodációs képességét tüntettük fel az életkor függvényében

életkor (év)	20	25	35	45
akkomodáció (dpt)	11	8,5	7	3,5



16. ábra. A szem akkomodációja az életkor függvényében. Regressziós egyenes illesztése a mérési pontokra.

Az ábrázolás és a lineáris regresszió elvégzése után az illesztett egyenes paraméterei:

$$a^* = -0,28; \quad b^* = 16,2.$$

A korrelációs együttható:

$$r = -0,98.$$

Mivel nem ismerünk olyan matematikai modellt, törvényszerűséget, ami a két mennyiséget összekapcsolja, ezért a paramétereket csak interpolációs becslésre használhatjuk fel. Például megkaphatjuk az akkomodáció becslött értékét 40 éves korban:

$$-0,28 \cdot 40 + 16,2 \approx 5 \text{ (dpt)}$$

ami a táblázatból egyébként hiányzik. (Tudjuk azonban, hogy még egészen fiatal korban sincs 16 dpt körüli akkomodáció.)

Mivel szubjektív megítélés alapján nagyon kétséges annak eldöntése, hogy a mérési pontok mennyire jól illeszkednek a regressziós egyenesre, ezért célszerű meghatározni a **korrelációs együtthatót**:

$$r = \frac{Q_{xy}}{\sqrt{Q_{xx} \cdot Q_{yy}}} = \frac{s_{xy}^2}{s_x \cdot s_y}, \quad (19)$$

ahol a használt jelölések ugyanazok, mint a (17) kifejezésben.

A korrelációs együttható a változók közötti **kapcsolat szorosságát jellemzi**, r értéke $+1$ és -1 között változhat. Pozitív értékeihez pozitív, negatív értékeihez negatív meredekségű egyenes tartozik. Ha a mérési pontok jól megközelítik a regressziós egyenest, az $|r|$ értéke közel lesz az 1-hez (pl.: $r = 0,9860$). Ha az egyenes minden egyes ponton átmegy, akkor $|r| = 1$, egyébként r annál jobban megközelíti a nullát, minél inkább eltérnek a pontok az illesztett egyenestől (lásd a 15. ábrát és egy konkrét példát a 7. megjegyzésben).

Nem beszéltünk még arról, hogy **milyen esetekben és milyen célra** használhatjuk a lineáris regressziót. Itt egy fontos **szempontot** kell figyelembe vennünk. Nevezetesen azt, hogy a változók (x, y) között **van-e** valamilyen „oksági” összefüggést leíró **modell**, amelynek paraméterei fizikai értelemmel bírnak, vagy csak azért illesztünk egyenest, hogy az a szükséges pontossággal reprezentálja mérési adatainkat. (Természetesen amennyiben a modell nem ismeretes, az még nem jelenti azt, hogy az oksági viszony nem állhat fenn.)

Az első esetre többek között példa lehet a folyadékok törésmutatójának koncentrációtól való függése ($n = n_0 + k c$) (lásd [4. REFRAKTOMETRIA](#) című fejezet), és a folyadékok optikai denzitásának koncentrációtól való függése ($\lg(J_0/J) = \varepsilon c x$) (lásd [6. FÉNYABSZORPCIÓ](#) című fejezet). Ilyenkor a lineáris regresszió paraméterei **extrapolációra** (a mérési tartományon kívüli becslésre) is használhatók. Az összefüggések (törvények) érvényességi tartományait azonban itt is figyelembe kell vennünk.

A második esetben a paramétereket csak **interpolációra** (a mérési tartományon belüli becslésre) használhatjuk és még 1-hez közel eső korrelációs együttható sem ad alapot arra, hogy **oksági kapcsolatról beszéljünk** (lásd 7. megjegyzés).

STATISZTIKAI DÖNTÉS, HIPOTÉZISVIZSGÁLAT

Az eddigi számítások arra vonatkoztak, hogy a **mintából** a változóra jellemző eloszlás megfelelő paramétereit minél jobban megközelítsük. Az ehhez hasonló számszerű, **kvantitatív következtetések** a **becslés** témakörébe tartoznak.

Gyakran kerül sor azonban a következtetések másik fajtájának, a **kvalitatív következtetéseknek** az alkalmazására is. Ez azt jelenti, hogy valamilyen **eldöntendő kérdésre kell igen-nem típusú választ adnunk**. Ilyen típusú kérdéseket már feltettünk az első részben a pulzusszámmal kapcsolatban. Azok a kérdések úgy kezdődtek, hogy: „**VÁLTOZIK-E ...?**”, „**VAN-E KÜLÖNBESÉG ...?**”.

A döntést minden esetben a **minta alapján** kell meghoznunk. Mivel a kérdés eldöntése mindig egy kiinduló feltételezés, hipotézis cáfolatát vagy megerősítését jelenti, ezeket a következtetéseket **hipotézisvizsgálatoknak** nevezzük (lásd 8. megjegyzés).

A módszer lényege, illetve főbb lépései egy példán keresztül mutathatók meg. Azt mondhatjuk, hogy a **büntető ügyekben a bírósági tárgyalás során „hipotézisvizsgálat” történik**. (Bár tudjuk, hogy ez a megállapítás túlzottan leegyszerűsített, mégis segíthet abban, hogy a hipotézisvizsgálat legfontosabb elemeit megragadjuk.) Esküdtbíráskodás esetén ugyanis az esküdtszék dönt arról, hogy a vádlott bűnös vagy nem bűnös. (A bíró csak az ítélet mennyiségi kérdéseit mérlegelheti.) Tehát van egy eldöntendő **kérdés**, amire csak igen-nem válasz adható: **bűnös-e a vádlott?** Számos jogrendben azonban alkalmazzák az **ártatlanság vélelmét**, ami annyit tesz, hogy mindaddig, amíg a vádlottról be nem bizonyosodik, hogy bűnös, ártatlannak kell tekinteni. (Tehát bizonyítani a bűnösséget és nem az ártatlanságot kell.) Ez a „**nem bűnös**” állítás az **alaphelyzet**, vagy, ha tetszik, **kiindulási hipotézis**.

Az ügyésznek (vagy a vád képviselőjének) a feladata a **bizonyítékok** felvonultatása annak igazolására, hogy a vád megalapozott. Az ügyvéd (vagy a védelem képviselője) a felhozott bizonyítékok hitelét, megbízhatóságát próbálja gyengíteni. Az esküdtszék végül **értékeli, mérlegeli a bizonyítékok „erejét”, és dönt**. A döntés maga az **alaphelyzet** vagy **kiindulási hipotézis**, nevezetesen a „**ártatlan**” állítás **elfogadását**, vagy **elvetését** jelenti. Bárhogyan is dönt az esküdtszék, döntése lehet igazságos vagy téves, így összesen négyféle kimenetel valósulhat meg.

Igazságos (helyes) döntés két esetben lehetséges:

- ha a bíróság **elfogadja az „ártatlan” hipotézist**, miközben a vádlott valójában **ártatlan**, vagy akkor is,
- ha a bíróság **elveti az „ártatlan” hipotézist** (azaz kimondja a bűnösséget), miközben a vádlott valójában **bűnös**.

Téves (helytelen) döntés is két esetben lehetséges:

- ha a bíróság **elfogadja az „ártatlan” hipotézist**, annak ellenére, hogy a vádlott valójában **bűnös**, vagy akkor is,
- ha a bíróság **elveti az „ártatlan” hipotézist** (azaz kimondja a bűnösséget), de a vádlott valójában **ártatlan** (lásd 9. megjegyzés).

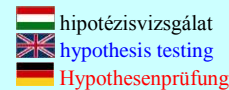
A valódi **hipotézisvizsgálat** abban különbözik ettől (az előbb említett leegyszerűsítésektől eltekintve), hogy a **mérlegelést számszerű eredmények alapján végezhetjük**, így a döntést szubjektív elemek kevésbé befolyásolják.

A mellékelt példa (lásd 1. feladat) megoldása során a fenti lépések szerint fogunk haladni.

A megválaszolandó **kérdés**: *Megtiltható-e a további forgalmazás az eltérő hatóanyag tartalom miatt, vagy még egyszerűbben, eltér-e a hatóanyag tartalom a megadottól?* („bűnös”-e?)

A vizsgálandó állítás, **alaphelyzet** vagy **kiindulási hipotézis**: *A további forgalmazás nem tiltható meg, mert a hatóanyag tartalom nem tér el a megadottól* („ártatlanság vélelme”, nem bűnös).

A **bizonyítékok**: *A mg-ban megadott adatok együttese.*



8. megjegyzés:

A leggyakrabban vizsgált hipotézisek a következők:

1. Egy eloszlás valamely paraméterére vonatkozó hipotézis. Például tudjuk, hogy egy vizsgált változó normális eloszlású és ellenőrizni akarjuk, helyes-e az a hipotézis, hogy várható értéke egyenlő valamely μ_0 számmal. Ilyen fajta vizsgálat szükséges például a változások eldöntéséhez is.

2. Két (vagy több) eloszlás paraméterére vonatkozó hipotézis. Például tudjuk, hogy két független változó normális eloszlású és ellenőrizni akarjuk azt a hipotézist, hogy várható értékeik egyenlőek. Így dönthető el például az a kérdés, hogy mondjuk egy adott populációban a férfiak vagy a nők élnek tovább, (vagy másképpen mondva, van-e különbség a várható élettartamuk között).

3. Függetlenségvizsgálat. Azt a hipotézist ellenőrizzük, hogy két vagy több változó független-e egymástól, (illetve van-e kapcsolat közöttük).

4. Homogenitásvizsgálat. Azt kérdezzük, hogy két (vagy több) változó eloszlása megegyezik-e.

5. Illeszkedésvizsgálat. Egy változó eloszlására vonatkozó hipotézis. Például azt vizsgáljuk, hogy egy változó normális eloszlású-e.

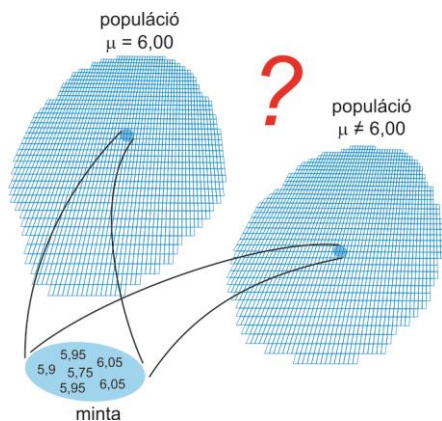
9. megjegyzés:

Az ítélet lehetséges kimenetelei:

valójában	az ítélet	
	felmentés	elítélés
ártatlan	igazságos	téves
bűnös	téves	igazságos

1. feladat:

Egy gyógyszer rendszeres forgalmazásának egyik feltétele a 6 mg hatóanyag tartalom. Egy ellenőrzés alkalmával a tárlóból kivett néhány tableta adatai a következők voltak (mg-ban): 6,05; 5,95; 5,75; 5,9; 5,95; 6,05. Megtiltható-e a további forgalmazás az eltérő hatóanyag tartalom miatt?



17. ábra. A kiválasztott minta származásának lehetőségei. (A feltüntetett értékek mg-ban értendők.)

10. megjegyzés:

Transzformációk:

A változók transzformációja voltaképpen azt jelenti, hogy az értékeket más skálán fejezzük ki; az eredeti értékekhez egy másik skálán új értékeket rendelünk hozzá.

Transzformáció az is, amikor egy fizikai mennyiséget az egyik mértérendszerből a másikba alakítunk át, azaz transzformálunk. Például: energia [eV] $\cdot 1,6 \cdot 10^{-19}$ = energia [J].

A transzformáció célja és értelme az, hogy a transzformált változókkal olyan statisztikai eljárásokat is elvégezhessünk, amelyeket az eredeti változóval nem, de a következtetéseink az eredeti változóra nézve is érvényesek lesznek.

Normalizáló transzformáció segítségével nem normális eloszlású változókat normális eloszlásúvá "tehetünk".

Kategorizáló transzformációkkal folytonos eloszlású változókat ordinális vagy nominális változókká alakíthatunk át. Akkor hasznosak többek között az ilyen eljárások, ha valamely, általunk vizsgált jelenség egy folytonos mennyiségi jellemző változásával minőségileg is megváltozik.

Ilyen változó például az életkor, ahol a következő transzformációval, folytonos változóból kétértékű nominális, ún. bináris változóhoz juthatunk:

életkor < 18 év \rightarrow gyermek (0),
életkor \geq 18 év \rightarrow felnőtt (1).

Rangsor transzformációt számszerű vagy ordinális változó esetén végezhetünk. Ilyenkor a minta elemeit nagyság szerint sorba rendezzük, majd a sorszámokat (rangsorszámokat) használjuk az eredeti értékek helyett. Rangsor transzformáción alapul számos, ún. **nem-paraméteres statisztikai eljárás** (lásd a NÉHÁNY PÉLDA AZ ORVOSI STATISZTIKA KÖRÉBŐL című részt).

A következő lépéseket (**értékelés, mérlegelés, döntés**) egy kicsit részletesebben kell megvizsgálunk. **Amennyiben minden ismeret a birtokunkban lenne**, azaz ismernénk a populációt (azaz az összes (!) tabletta hatóanyag tartalmát), vagy ami azzal ekvivalens, annak eloszlását, akkor nem kellene mást tennünk, mint a populációra jellemző eloszlás várható értékét (μ -t) össze kellene hasonlítanunk 6 mg-mal. Ebben a helyzetben nincs szükségünk mérlegelésre. Egyszerűen azt mondhatjuk, hogy ha $\mu = 6$, akkor elfogadjuk, ha $\mu \neq 6$ akkor elvetjük a kiindulási hipotézist. Azt is mondhatjuk, hogy egy ilyen lépéssel nem tettünk mást, mint számszerűsítettük a kiindulási hipotézist, nevezetesen a $\mu = 6$ állítás esetünkben ekvivalens azzal, hogy „a további forgalmazás **nem tiltható meg**, mert a hatóanyag tartalom **nem tér el a megadottól**”. (Meg kell jegyeznünk azért, hogy a gyakorlatban az is fontos feltétel, hogy a tabletták azonos mennyiségű hatóanyagot tartalmazzanak, tehát a szórást is vizsgálnunk kellene.)

Mivel ez az ideális eset a valóságban gyakorlatilag sohasem áll fenn, ezért a helyzet ennél bonyolultabb. Tudjuk, hogy egy minta alapján meghatározott, jól megválasztott konfidencia intervallum a várható értéket adott bizonyossággal közreárja. Ezért úgy járhatunk el, hogy **először megválasztjuk** a „szükséges” bizonyosságot, azaz azt a **konfidencia szintet** amelyet megfelelőnek tartunk, **majd megnézzük, hogy az ehhez tartozó konfidencia intervallum közreárja-e a 6-ot**. Ha **igen**, akkor **elfogadjuk**, ha **nem**, akkor **elvetjük** a kiindulási hipotézist. Tehát a döntést egy **minta alapján**, egy előre megválasztott és **lerögzített konfidencia szint mellett** hozzuk meg.

Egy másik szempontból úgy is tekinthetjük ezt a problémát, hogy létezik egy populáció, aminek ismert a várható értéke ($\mu = 6$) és a kérdés az, hogy az adott minta ebből a populációból lett véletlenül kiválasztva, vagy egy másiktól, olyanból, amelynek a várható értéke, $\mu \neq 6$ (lásd 17. ábra).

Válasszuk a konfidencia szintet 95%-osra. Ez azt jelenti, hogy 100 hasonló mintavétel esetén csak 5-ször fordulhat elő, hogy az ennek megfelelő konfidencia intervallum a „véletlen” folytán nem zárja közre a 6-ot. Tudjuk, hogy ez az intervallum az $\bar{x} \pm k s_{\bar{x}}$ képlet segítségével meghatározható, és azt is tudjuk, hogy nagy minták esetén $k \approx 2$. A mi esetünkben (kis elemszámú mintáról van szó) k egyelőre nem ismert, de ezt a hiányosságot hamarosan pótoljuk. Amennyiben $k = 2$ -vel számolunk, az adatok alapján a konfidencia intervallum: 5,85 és 6,03 között húzódik. Mivel ez közreárja 6-ot, ezért **elfogadjuk a kiindulási hipotézist**.

Ez azt jelenti, hogy a vizsgálat eredménye nem szolgáltatott elegendő bizonyítékot ahhoz, hogy a forgalmazást betiltsák, ezért „a további forgalmazás (ha csak ez volt a kritérium) **nem tiltható meg**”, (hiszen a hatóanyag tartalom nem tér el a megadottól, pontosabban nem esik a konfidencia intervallumon kívülre).

A fenti gondolatmenethez hasonlókkal találkozunk majd az ún. statisztikai próbák megbeszélése során, melyre hamarosan rátérünk, de a jobb érthetőség kedvéért előbb egy fontos matematikai segédeszközt ismertetünk.

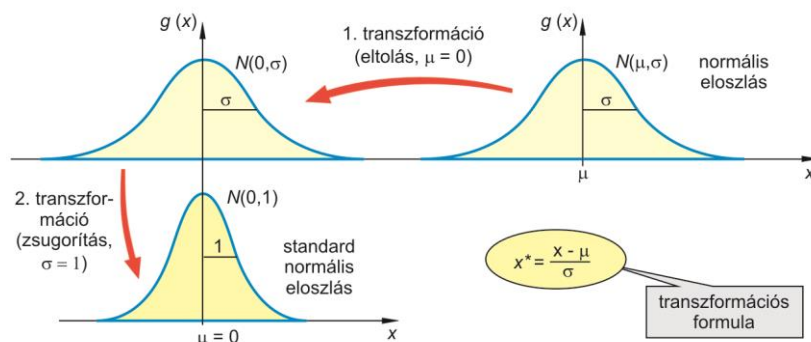
A VÁLTOZÓK TRANSZFORMÁCIÓJA ÉS AZ ÚJ VÁLTOZÓK ELOSZLÁSA

Az adatok transzformációjáról már szót ejtettünk a GRAFIKUS ADATFELDOLGOZÁS című részben, amikor az volt a célunk, hogy mérési pontjainkat egyenessel „összekössük”. Itt egy kicsit részletesebben megvizsgáljuk a transzformáció kérdését.

Amikor egy, vagy több adatból matematikai műveletek elvégzése után megkapjuk az eredményt, azt mindig egy **új változónak** is tekinthetjük, hiszen más adatokból kiindulva ugyanaz a művelet sor általában más eredményre vezet. Az ilyen átalakításokat általánosan **transzformációnak** nevezzük. A legegyszerűbb transzformációk között említhetjük egy állandó hozzáadását, vagy egy állandóval való szorzást, de az átlag, vagy a tapasztalati szórás meghatározása is transzformációnak tekinthető (lásd még a 10. megjegyzést).

Vegyük például az $N(\mu, \sigma)$ normális eloszlású populáció x általános elemét (változót), és hajtsuk végre rajta a $x^* = (x - \mu) / \sigma$ transzformációt. (Azt is mondhatjuk, hogy a populáció minden elemén végrehajtjuk ezt a művelet sor.)

A kérdés az, hogy milyen az új x^* változó eloszlása. Az első lépésben tehát az összes elemet eltoljuk egy várható értéknyivel, ami az $N(0, \sigma)$ normális eloszlást eredményezi, majd a 0-tól való eltéréseket σ -val osztva az $N(0, 1)$ standard normális eloszlást kapjuk (lásd 18. ábra). Egy ilyen transzformációnak fontos gyakorlati haszna van, ugyanis bármilyen normális eloszlású változó standard normális eloszlásává transzformálható, ami egységes adatfeldolgozást tesz lehetővé.



18. ábra. Általános helyzetű és szélességű normális eloszlás transzformációja standard normális eloszlássá ($N(\mu, \sigma) \rightarrow N(0, 1)$).

Tegyük most fel, hogy „ugyaneyt” a transzformációt szeretnénk végrehajtani, de σ -t nem ismerjük, így annak csak becslőt értékét, egy n elemű mintából meghatározott tapasztalati szórást (s) használjuk fel a célra. Az így kapott változó természetesen nem lehet $N(0, 1)$ eloszlású, csak ahhoz hasonló. Az új eloszlás az $n - 1$ szabadságfokú **Student-, vagy t-eloszlás**: $t = (x - \mu)/s$ (lásd 19. ábra). (Azon nem csodálkozhatunk, hogy ez az eloszlás n -től is függ, hiszen a transzformációban szerepel egy n -től függő paraméter (s). Az eloszlás további tulajdonságairól a következő részben ejtünk néhány szót.)

Más transzformációkkal más eloszlásokhoz juthatunk. Ezek közül még egyet itt bemutatunk: ha van n darab $N(0, 1)$ eloszlású változónk és ezeknek a négyzetösszegét vesszük, akkor az új változó eloszlása n szabadságfokú **χ^2 -eloszlás** lesz (lásd 20. ábra). (Egy ilyen változó negatív értékeket értelemszerűen nem vehet föl. Ez a transzformáció figyelhető majd meg a (24) összefüggésben is.)

STATISZTIKAI PRÓBÁK

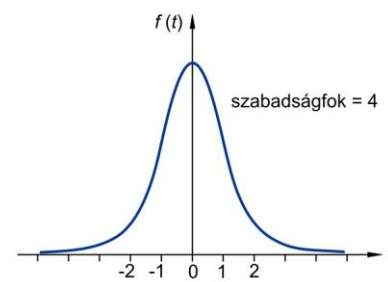
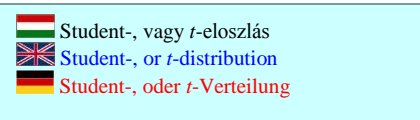
Bár a hipotézisvizsgálatot a már bemutatott módon is el lehet végezni (k pontos megadásával ugyan egyelőre adósak maradtunk), az egyszerűbb kezelhetőség érdekében inkább **statisztikai próbákat** használunk.

A statisztikai próbák igen **sokfélék** aszerint, hogy **mi az ellenőrizendő hipotézis**, mik az **alkalmazhatóság feltételei**, és mi a **végrehajtás módja**; valamennyinek **közös** azonban a **gondolatmenete**.

Abból indulunk ki, hogy a mintából számolt **becsült paramétereknek is van valamilyen eloszlásuk**, hiszen, ha másik mintát választunk, a becsült paraméterek is mások lesznek. Az eloszlás konkrét alakja

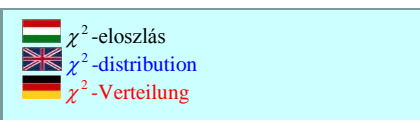
- egyrészt az eredeti változó eloszlásától függ,
- másrészt attól, hogy melyik becsült paraméterről, vagy általánosabban mondva, melyik statisztikai jellemzőről (ami lehet például az r korrelációs együttható is) van szó,
- harmadrészt pedig a minta elemszámától, pontosabban az azzal szorosan összefüggő szabadságfoktól.

Az egyszerűsítés kedvéért sok-sok eloszlás helyett azoknak csak viszonylag kevés számú standardizált változatát fogjuk használni. Ennek érdekében azonban a vizsgált becsült paramétert, illetve statisztikai jellemzőt mindig — az adott standardizált eloszláshoz illeszkedően — a kívánt alakra kell transzformálnunk (lásd az előző részt, mint például $N(\mu, \sigma) \rightarrow N(0, 1)$).

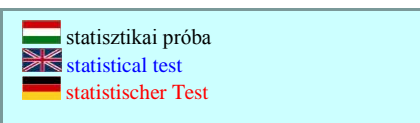


19. ábra. A t-eloszlás 4-es szabadságfokú változata.

A görbe az $N(0, 1)$ eloszláshoz hasonlít, de a szabadságfoktól függően, többé-kevésbé eltér attól.









20. ábra. A χ^2 -eloszlás 6-os szabadságfokú változata.









 t -próba
 t -test
 t -Test

 χ^2 -próba
 χ^2 -test
 χ^2 -Test

 nullhipotézis
 null hypothesis
 Nullhypothese

 alternatív hipotézis
 alternative hypothesis
 Alternativhypothese

 kétoldali próba
 two-tailed test
 zweiseitiger Test

 egyoldali próba
 one-tailed test
 einseitiger Test

A legegyszerűbb és elég gyakran használt statisztikai próbák a **t -próbák** és a **χ^2 -próbák**. Mindkét típusú próbához tartozik egy-egy standardizált elméleti eloszlás, a t -eloszlás, illetve a χ^2 -eloszlás. Ezek valójában nem egy-egy konkrét eloszlást jelentenek, hanem eloszlás sorozatokat, hiszen bennük a szabadságfok — mint szabad paraméter — befolyásolja az eloszlás konkrét alakját.

Térjünk vissza az előbbi példához (1. feladat) és nézzük meg, mit is jelent a próbák „nyelvén” a hipotézisvizsgálat. A megválaszolendő **kérdés** változatlan: **Megtiltható-e a további forgalmazás az eltérő hatóanyag tartalom miatt**, vagy még egyszerűbben, **eltér-e a hatóanyag tartalom a megadottól?** („bűnös”-e?)

A vizsgálandó állítást, az alaphelyzetet vagy kiindulási hipotézist itt **nullhipotézisnek** nevezzük (ennek okára még visszatérünk): **A további forgalmazás nem tiltható meg, mert a hatóanyag tartalom nem tér el a megadottól** („ártatlanság vélelme”, nem bűnös).

Ezt egy kicsit átfogalmazzuk a számszerűsíthetőség kedvéért: az adott átlaggal ($\bar{x} = 5,94$) jellemzett **minta** a $\mu_0 = 6$ **várható értékű populációból lett kiválasztva**, és nem egy másik $\mu' \neq 6$ várható értékű populációból. Azt is mondhatjuk, hogy mivel a **populáció várható értékét, μ -t nem ismerjük**, csak a minta átlagát, tulajdonképpen $5,94 \approx 6$ és az eltérés „nem valódi”, csak a véletlen eredményezi. (Tehát valójában $\mu = \mu_0$, illetve $\bar{x} \approx \mu_0$, vagy másképpen $\mu - \mu_0 = 0$, illetve $\bar{x} - \mu_0 \approx 0$). A **nullhipotézis**, amit **H_0 -al** szokás jelölni, tehát úgy is megfogalmazható, hogy $\mu - \mu_0 = 0$, de ez direkt módon nem ellenőrizhető, így csak $\bar{x} - \mu_0 \approx 0$ marad, noha erről meg tudjuk, hogy **nem teljesülését a véletlen is okozhatja**.

A nullhipotézis megfogalmazásakor egy **alternatív hipotézist, H_1 -et** is fel kell állítanunk, ami akkor lép életbe, ha a nullhipotézist elvetjük. Első gondolatunk az lehet, hogy ez teljesen felesleges, hiszen az ellentétes állítás megfogalmazása egyértelmű. Így amennyiben $H_0 : (\mu - \mu_0 = 0)$, akkor $H_1 : (\mu - \mu_0 \neq 0)$. Sok esetben ez igaz, tehát a **H_0 -t akkor is elvetjük, ha $\mu - \mu_0 < 0$, meg akkor is, ha $\mu - \mu_0 > 0$** . Ekkor beszélünk **kétoldali próbáról**.

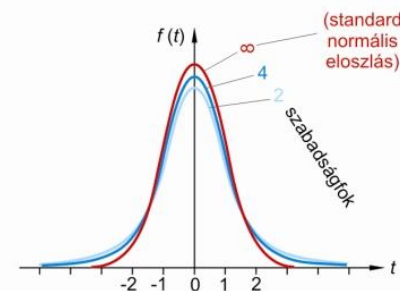
Vannak azonban olyan esetek is, amikor csak az egyik irányú eltérés felel meg elvárásainknak. Ha például egy gyógyszer hatását vizsgáljuk, legyen az mondjuk egy vérnyomáscsökkentő vagy lázcsillapító, akkor a lehetséges két irányú változás közül csak az egyik, nevezetesen a csökkenés jelenti a gyógyszer hatásosságát. Ilyenkor eleve feltesszük, hogy vérnyomás-, illetve hőmérsékletnövekedés csakis véletlenül következhet be. Ilyen esetekben a nullhipotézis ugyanaz marad, $H_0 : (\mu - \mu_0 = 0)$, de az alternatív hipotézis $H_1 : (\mu - \mu_0 < 0)$, tehát **H_0 -t csak akkor vetjük el, ha $\mu - \mu_0 < 0$** . Ekkor beszélünk **egyoldali próbáról**.

Folytatva a megkezdett gondolatsort, azt mondhatjuk, hogy **ha a hiba** (véletlen hiba) **elég nagy**, akkor az az érvéltetés, hogy a nullhipotézis $\bar{x} - \mu_0 \approx 0$ alakja **igaz, csak nem pontosan teljesül**, elfogadható. Ha azonban ez a **hiba kicsi**, akkor egyáltalán nem lehetünk annyira biztosak a dolgunkban, tehát **az eltérés akár „valódi” is lehet**. Emiatt a **döntéshez az $\bar{x} - \mu_0$ mérhető eltérést és a standard hibát** (ami éppen a véletlen eltérések jellemzője), **más néven az átlag szórását** ($s_{\bar{x}}$ -ot) **kell összehasonlítani**. Az előző részben láthattuk, hogy az $x \sim N(\mu, \sigma)$ eloszlású változó a $t = (x - \mu)/s$ összefüggés segítségével $n-1$ szabadságfokú t -eloszlásúvá transzformálható. Ebből arra következtethetünk, hogyha a változó (x) és annak szórása (s) helyett az átlagot (\bar{x}) és annak szórását ($s_{\bar{x}}$) írjuk a transzformációs képletbe, szintén ugyanarra az eredményre jutunk, tehát

$$t = (\bar{x} - \mu_0) / s_{\bar{x}}.$$

Így esetünkben a még ismeretlen t -próba használata tűnik a legkézenfekvőbbnek.

A 21. ábrán ismét a t -eloszlás látható, de most több (2, 4 és ∞) szabadságfok esetében. Megfigyelhetjük, hogy az eloszlás várható értéke mindig 0 ($t=0$), lefutása pedig, mint már az előző részben említettük, hasonlít a standard normális eloszláshoz ($N(0,1)$), sőt, ha a szabadságfok végtelen, akkor egzaktul megegyezik vele. Láthatjuk azt is, hogy az imént megfogalmazott **nullhipotézisnek** ($\bar{x} - \mu_0 \equiv 0$) a $t = 0$ elvi érték felel meg. Az adott **mintából kiszámított** t_m pedig arra jellemző, hogy az adatok inkább a nullhipotézis elfogadása vagy elvetése mellett szólnak, így t_m **együttal az adatok** (bizonyítékok) „**erejét**” is méri. (Az eredeti hasonlathoz visszatérve a vád célja t növelése, a védelem pedig annak csökkentése.)



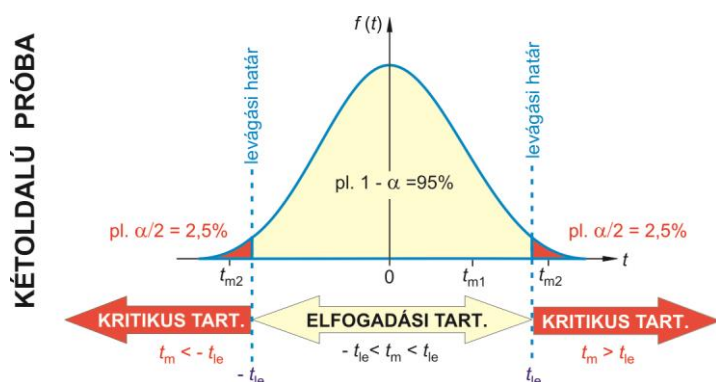
21. ábra. Különböző szabadságfokú t -eloszlások.

Egy sor statisztikai jellemző eloszlását bizonyos átalakításokkal a t -eloszlásra tudjuk visszavezetni. A változóban feltételezett **megváltozást, eltérést, különbséget, kapcsolatot** minden esetben valamilyen paraméter méri. Ezt a paramétert fogjuk a mintából számolt statisztikai jellemzővel közelíteni, és ennek standardizált alakja mindig egy t_m értéket ad.

Az így kiszámított t_m értéke **elvileg nulla**, ha **nincs megváltozás, nincs eltérés, nincs különbség**, vagy **nincs kapcsolat**. Ez az oka annak, hogy az eredetileg feltett eldöntendő kérdésre tagadó választ megfogalmazó állítást nullhipotézisnek nevezik. A nullhipotézisnek kitüntetett szerepe van a hipotézisvizsgálatban. Függetlenül attól, hogy melyik állítás megerősítését várjuk a vizsgálattól, az eljárás során mindig a nullhipotézis helyességét tesszük föl. Végül is ezt megerősítve vagy cáfolva fogadjuk el a tagadó, illetve az igenlő választ az eredeti kérdésre.

Igen könnyű megtalálni ennek az eljárásnak az indokát. Az **igenlő** válasz esetén a paraméter értéke végtelen sokféle lehet, a **tagadó** állításhoz ellenben csak egyetlen paraméterérték, a nulla tartozik. A paraméter rögzítése egyértelművé teszi a kiszámításra kerülő statisztikai jellemző eloszlását, a nullhipotézisnek tehát egyetlen lehetséges eloszlás (a t -eloszlás) felel meg, az ellenkező állításnak pedig végtelen sok. Hangsúlyozzuk azonban, hogy a kiszámított t_m érték csak elméletileg lehet 0. A valóságban a számítások elvégzése után legtöbbször egy 0-tól különböző t_m értéket kapunk. Az így kapott t_m érték és a hipotézisben feltételezett t -eloszlás egymáshoz való viszonyából kell döntésünket meghoznunk.

A döntéshozatal szempontjából könnyű lenne a dolgunk, ha a t -eloszlás eleve csak egy adott tartományra, mondjuk ($t_{\text{kezdő}} - t_{\text{vég}}$) terjedne ki. Ilyenkor csak azt kellene megnéznünk, hogy a t_m érték ezen a tartományon belül, vagy kívül helyezkedik-e el. A tartományon belül a nullhipotézist elfogadnánk, a tartományon kívül pedig elvetnénk. Tudjuk azonban, hogy a t -eloszlás — ugyanúgy, mint a Gauss-eloszlás — $-\infty$ -tól $+\infty$ -ig terjed, így olyan tartomány, ami alapján egyértelmű döntést hozhatnánk, nem létezik.



22. ábra. Elfogadási és kritikus tartományok kétoldali t -próba esetén.

Mivel a döntéshez mindenképpen valamilyen tartományra van szükségünk, vágjuk le a t -eloszlás „nullától távoli” értékeit. (Hogy pontosan hogyan, illetve, hogy milyen szabályok szerint, arra természetesen még visszatérünk.) Ezután már

szignifikáns
significant
signifikant

kritikus tartomány
critical region
kritischer Bereich (Ablehnungsbereich)

elfogadási tartomány
acceptance region
Annahmehereich

11. megjegyzés:

A „szignifikáns eltérés” sohasem jelenthet „biztosat” és a nem szignifikáns még kevésbé, hogy biztosan nincs eltérés. Nyilvánvaló ugyanis, hogy lehetnek olyan valódi, de kis különbségek, melyek jelenlegi kísérleti, mérési stb. módszereink és műszereink hibahatárain belül vannak. Meg kell jegyeznünk továbbá azt is, hogy a szignifikancia vizsgálat sohasem felelhet arra a kérdésre, hogy mi az ok.

szignifikancia szint
significance level
Signifikanzniveau

elsőfajú hiba
type I error
Fehler 1. Art

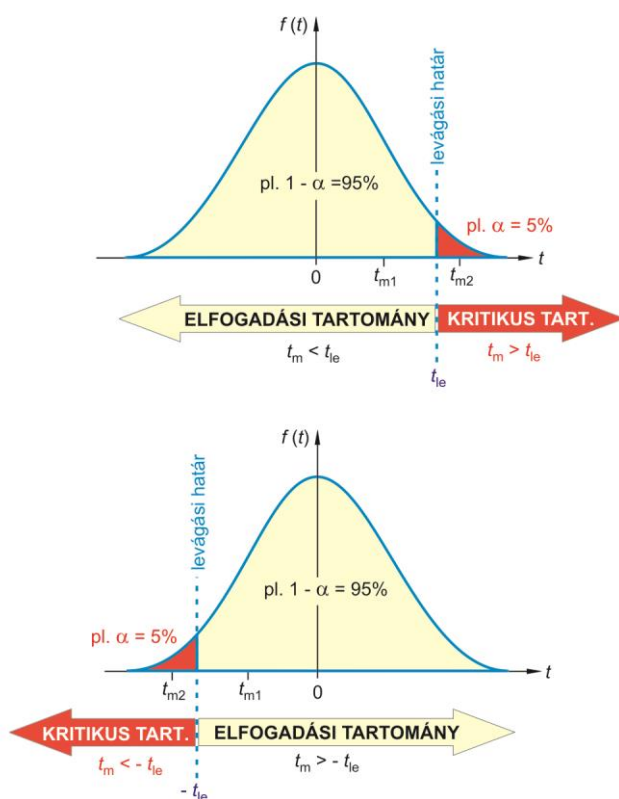
másodfajú hiba
type II error
Fehler 2. Art

megkérdezhetjük, hogy a kiszámított t_m érték a levágási határokon (t_{le}) belül vagy kívül helyezkedik-e el (lásd 22. és 23. ábra). Ha belül van (például t_{m1}), elfogadjuk a nullhipotézist, ha kívül van (például t_{m2}), akkor elvetjük azt, és azt mondjuk, hogy a kiszámított t_m értéke az adott levágás mellett a nullától szignifikánsan eltér, vagy röviden csak így: **szignifikáns** (lásd 11. megjegyzés). A levágott részeket **kritikus tartománynak**, a megmaradó részt **elfogadási tartománynak** nevezzük.

Bárhogyan is döntünk, tehát akár elvetjük, akár elfogadjuk a nullhipotézist, döntésünkben — mint minden döntésben — benne van a tévedés lehetősége.

Azt a fajta tévedést, hibát, hogy egy valójában **igaz nullhipotézist mégis elvetünk, elsőfajú hibának** nevezzük. (Az eredeti hasonlatban ez annak felel meg, amikor az ártatlant tévesen bűnösnek mondja ki az esküdtszék.) Ezt a hibát akkor követjük el, ha a kiszámított t_m -érték igazából a (0 körüli) t -eloszláshoz „tartozik”, de a levágás miatt a kritikus tartományba esik, és ezért elvetjük a nullhipotézist. Pontosan **meg tudjuk adni e hiba elkövetésének valószínűségét**: a levágással „kívülre” kerülő értékek előfordulásának valószínűsége éppen a **levágott terület nagyságával mérhető** (lásd a 22. és 23. ábrákat, valamint a korábbi 9. ábrát is).

EGYOLDALÚ PRÓBÁK

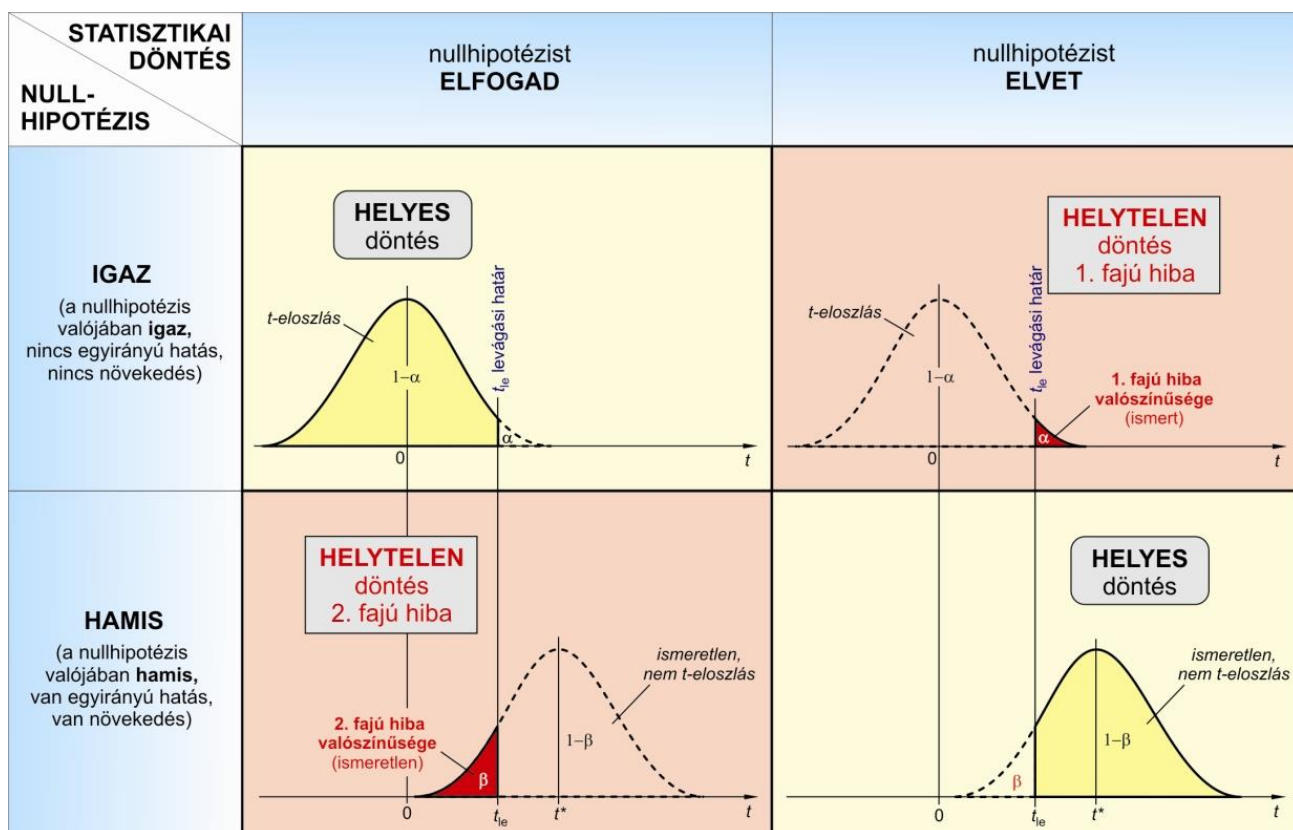


23. ábra. Elfogadási és kritikus tartományok egyoldali t -próba esetén.

Az elsőfajú hiba valószínűségét (probabilitás; p) α -val jelölik és a **statisztikai döntés valószínűségi szintjének**, ritkábban a tévedés valószínűségének nevezik. A gyakorlatban úgy járunk el, hogy nem a nullától való távolságot választjuk meg szabadon, hanem az elhagyott értékeknek (a levágott területtel jellemezhető) α valószínűségét. Ezt a valószínűséget **szignifikancia szintnek** is nevezik, mert egyértelműen meghatározza, hogy mely t értékek tekinthetők szignifikánsnak és melyek nem.

Abban az esetben is hibát követünk el, tehát tévedünk, ha **egy hamis nullhipotézist** — amiről persze nem tudjuk, hogy valójában hamis — **elfogadjuk**. Ezt a hibafajtát **másodfajú hibának** nevezzük. (Az eredeti hasonlatban ez annak felel meg, amikor a bűnöst tévesen felmenti a bíróság.) Ez a helyzet akkor áll elő, ha a kiszámított t_m érték valamilyen $t^* \neq 0$ várható értékű eloszláshoz „tartozik”, tehát nem a (0 körüli) t -eloszláshoz, de mi tévesen mégis ahhoz „tartozónak”

képzeljük. Az ilyen hiba elkövetésének ($p = \beta$) valószínűségét most a t^* várható értékű eloszlás megfelelő görbe alatti területével mérhetnénk, de mivel t^* -ot nem ismerjük, ezért a levágott terület is ismeretlen, és emiatt **a másodfajú hiba valószínűségét nem lehet meghatározni** (lásd a 24. ábrát is). (Csak megjegyezzük, hogy azért vannak módszerek, amelyekkel ez a hiba is becsülhető.)



24. ábra. Az első- és másodfajú hiba szemléltetése.

Fontos hangsúlyoznunk, hogy α -nak csak a hipotézis elvetésekor, β -nak pedig csak a hipotézis elfogadásakor **van értelme**. Ugyanakkor nyilvánvaló, hogy az egyik hiba valószínűségének csökkentésével a másik hiba valószínűsége növekszik. Mivel csak az elsőfajú hiba valószínűsége (α) rögzíthető, ezért annak megválasztásakor figyelembe kell vennünk ezt a körülményt is.

A statisztikai döntésnek az orvosi, biológiai területeken szokásos valószínűségi szintje ($p = \alpha = 0,05$ (azaz 5%). Ez azt jelenti, hogy átlagosan 100 olyan eset közül, amikor nullhipotézisünk igaz, 5 esetben fogjuk azt tévesen elvetni. Számos esetben azonban az $\alpha = 0,05$ tévedés nem engedhető meg. Ilyenkor a valószínűségi szint $\alpha = 0,01$ -ra, $\alpha = 0,001$ -re, vagy tetszés szerint tovább csökkenthető. Ezzel szemben természetesen növekszik a másodfajú hiba valószínűsége, nevezetesen az, hogy a helytelen nullhipotézist elfogadjuk (amennyiben tényleg elfogadjuk).

Nézzük meg ezek után azt, hogy a gyakorlatban hogyan végezzük el t -próbával a hipotézisvizsgálatot. Ehhez tudnunk kell azt, hogy bár a t -eloszlás bonyolult képlet formájában is megadható lenne, egyszerűbb, ha — ugyanúgy, mint pl. a szögfüggvények esetében — táblázatos formában adjuk meg. Lényeges felépítésbeli különbség van azonban a t -eloszlás táblázata és mondjuk a szinusz függvény táblázata között. Míg a szinusz függvény esetében az adott x értékekhez az $f(x) = \sin x$ függvényértékeket adjuk meg, addig a t -eloszlás táblázatának a jobb használhatóság kedvéért különleges szerkezete van (lásd 25. ábra és 7. táblázat).

Először is a **t táblázat** nem egy, hanem **sok t -eloszlást tartalmaz**, amelyeket a szabadságfok különböztet meg egymástól. A bal oldali **oszlopban** találjuk a **szabadságfokokat**: ennek megfelelően minden sorban egy-egy eloszlás szerepel, ez azonban nem függvényértékeket jelent, hanem t értékeket, azaz a független változó speciális értékeit.

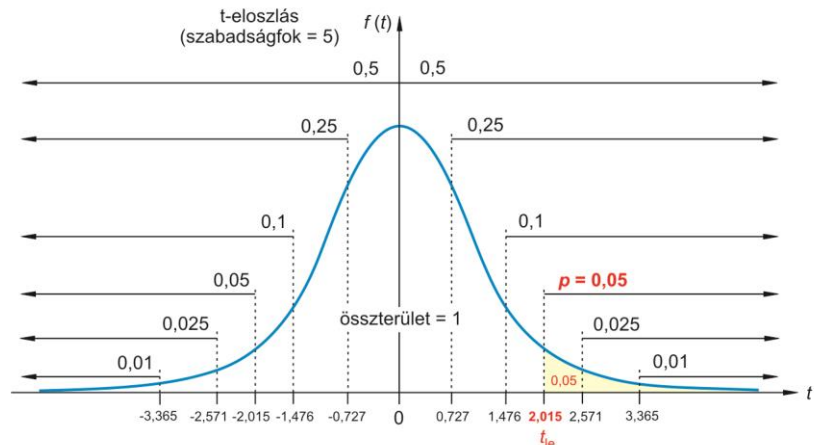
A táblázat jelentésének jobb megértése érdekében hasonlítsuk össze az 5-ös szabadságfokú t -eloszlás ábráját a táblázat ötödik sorában álló eloszlással. A táblázat két **fejlécében a p valószínűségek** állnak (egy, illetve kétoldali próba esetére): az ábrán ezeknek görbe alatti területek felelnek meg. A táblázat annak a t -nek az abszolút értékét adja meg, amelynél az eloszlás egyik vagy mindkét végét (szimmetrikusan) levágva, a lehasított területek együttesen a fejlécen álló valószínűséggel egyenlők. Másképpen mondva tehát a t táblázat legfelső soraiban állnak a szabadon megválasztható szignifikancia szintek, az ezekhez tartozó (levágási) t_{le} értékeket pedig a t táblázat adott szabadságfokú sorából olvashatjuk ki.

Az egyértelmű döntéshozatal szempontjából az imént ismertetett eljárás egyes lépéseinek sorrendje is nagyon fontos. **Először kell megválasztanunk azt a szignifikancia szintet**, amelyen majd a későbbi döntésünket meghozzuk, és csak utána végezzük el az összehasonlítást (lásd a 12. megjegyzést is).

12. megjegyzés:

A komputerok világában a t táblázatot egyre kevésbé használjuk, hiszen a számítógép tetszőleges levágási értékhez meg tudja határozni a valószínűséget. (Illetve legtöbbször ki sem számolja t értékét.) Így a próbák eredménye egy p valószínűség, és a mérlegelést ennek alapján végezzük el. Ha p elég kicsi (kisebb, mint az általunk előre megválasztott α), akkor elvetjük H_0 -t (és emiatt H_1 -lép életbe), hiszen kicsi a valószínűsége annak, hogy egy igaz hipotézist vetettünk el. (Azt is mondhatjuk, hogy kicsi a valószínűsége annak, hogy csupán az adatok véletlen eltérése okoz 0-tól ennyire eltérő t értéket.) Ha p nagy, akkor hasonló érvelés miatt ezt nem tehetjük meg, tehát megtartjuk H_0 -t.

(p kicsiségét vagy nagyságát viszont nekünk kell eldöntenünk.)



25. ábra. t -eloszlás. Az egyoldali t -próba-hoz tartozó levágási határok és valószínűségek.

p (egyoldali próba)									
	0,45	0,35	0,25	0,15	0,10	0,05	0,025	0,01	0,005
p (kétoldali próba)									
szab. fok	0,90	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
1	0,158	0,510	1,000	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,445	0,816	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,424	0,765	1,250	1,638	2,35	3,182	4,541	5,841
4	0,134	0,414	0,741	1,190	1,533	2,132	2,776	3,74	4,604
5	0,132	0,408	0,727	1,156	1,476	2,015	2,571	3,365	4,032

7. táblázat. A t -eloszlás levágási határai és p valószínűségei egy-, és kétoldali próba esetén.

A következő lépésként összehasonlítjuk a táblázatból kapott t_{le} értéket a mintából meghatározott t_m értékkel. **Ha $t_m \leq t_{le}$, akkor a nullhipotézist elfogadjuk, ha $t_m > t_{le}$, a nullhipotézist elvetjük**, és azt mondjuk, hogy az adott szignifikancia szint mellett az eltérés szignifikáns.

A t-PRÓBA ALKALMAZÁSA AZ 1. FELADATRA

Az 1. feladattal odáig jutottunk, hogy megfogalmaztuk a nullhipotézist többféleképpen is:

1. A további forgalmazás **nem tiltható meg**, mert a hatóanyag tartalom **nem tér el** a megadottól.
2. $H_0: (\bar{x} - \mu_0 \equiv 0)$.

Ezután megfogalmazzuk az alternatív hipotézist: $H_1: \mu - \mu_0 \neq 0$. Ebből tudjuk, hogy kétoldali próbát kell alkalmaznunk. A további lépések:

- Kiszámítjuk a mintából t_m értékét: $t_m = 1,28$.
- Megválasztjuk a szignifikancia szintet: $\alpha = 0,05$.
- Megállapítjuk a szabadságfokot: $n - 1 = 5$.
- A táblázatból kikeressük az értékpárnak megfelelő t_{le} értéket: $t_{le} = 2,571$.
- Mivel $1,28 < 2,571$ (azaz $t_m < t_{le}$), ezért a nullhipotézist elfogadjuk.
- Tehát: A további forgalmazás **nem tiltható meg**, mert a hatóanyag tartalom **nem tér el** a megadottól.

Ez még pontosabban azt jelenti, hogy **adataink nem szolgáltatnak elegendő bizonyítékot** (t_m , a bizonyítékok „erejét” mérő paraméter nem elég nagy) **ahhoz, hogy a nullhipotézist elvessük**, tehát, hogy ennek alapján a **forgalmazást megtiltsuk**.

További megjegyzések: Mivel a nullhipotézist elfogadtuk, ezért bennünket nem α , hanem β érdekelne jobban, tehát annak a valószínűsége, hogy **esetleg mégis hamis hipotézist fogadtunk el**. Mint már említettük, β -t viszont nem lehet számszerűen meghatározni. Tudjuk azonban, hogy α növelése β csökkenésével jár együtt. Így azt tehetjük, hogy α -t egészen addig növeljük, (ezzel együtt t_{le} -et addig csökkentjük), amíg $t_{le} = t_m$ nem teljesül. Ha ennél csak egy kicsit nagyobb levágási értéket (t_{le}) választunk, a nullhipotézist ugyanúgy elfogadjuk, mint az előbb, viszont az ehhez tartozó α lényegesen nagyobb, mint 0,05: $p = 0,26$. Bár ennek, mivel elfogadtuk a nullhipotézist, önmagában nincs sok értelme, áttételesen azt jelenti, hogy β , tehát a másodfajú hiba valószínűsége csökkent. Így ilyen esetben célszerű ezt az α értéket megadni.

A t-próba alkalmazhatóságának feltételei a következők:

1. a változó eloszlása legyen normális eloszlás,
2. a minta vagy minták elemei legyenek függetlenek egymástól,
3. két minta esetén a szórások legyenek „elégge” egyformák.

A három feltétel közül az 1.-t és a 3.-at nem kell nagyon szigorúan venni, de a 2. - at igen. A t-próbát általában középértékek, illetve középértékek különbségének vizsgálatához használjuk, de például a korrelációs együttható is vizsgálható vele.

Ezek után lássuk a t-próba legismertebb formáit és a t_m értékét megadó képleteket.




EGYMINTÁS t-PRÓBA




Az ide vonatkozó **korábbi kérdés**: **VÁLTOZIK-E** a pulzusszám egyperces lélegzet-visszatartás után (lásd 13. megjegyzés)?




Általánosabban megfogalmazva: **megváltozik-e a populáció eloszlásának várható értéke** valamilyen beavatkozás következtében, más szóval van-e hatása a beavatkozásnak. Vagy **különbözik-e a populáció eloszlásának várható értéke egy előre megadott értéktől?**

$$t_m \text{ kiszámítása: } t_m = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{s} \sqrt{n}, \quad (20)$$

t_m szabadságfoka: $n-1$.

 egymintás t-próba
 t-test for a single sample
 Einstichproben-t-Test

 kétmintás t-próba
 t-test for two samples
 Zweistichproben-t-Test

 korrelációs t-próba
 t-test for correlation
 t-Test zur Korrelationsanalyse

13. megjegyzés (részletesen kidolgozott megoldás):

- H_0 : a pulzusszám nem változik, azaz $\mu - \mu_0 = 0$, ahol $\mu_0 = 0$, vagy $\bar{x} \equiv 0$, ahol \bar{x} a változások átlagát jelenti.
- H_1 : $\mu - \mu_0 \neq 0$, (kétoldali próba). Az első két oszlop a felvett adatokat tartalmazza, egyperces lélegzet-visszatartás előtt (x_e), és után (x_u), $n = 6$ kísérleti személy esetén. A harmadik és a negyedik oszlop az adatok különbségeit (x'_i), illetve azok négyzeteit tartalmazza (ez utóbbiak az átlag és a szórás kiszámításához szükségesek).

x_e	x_u	$x'_i = x_u - x_e$	x'^2_i
69	71	2	4
60	63	3	9
68	70	2	4
75	76	1	1
71	70	-1	1
66	69	3	9
		$\Sigma x'_i = 10$	$\Sigma x'^2_i = 28$

Az adatok különbségeinek átlaga:

$$\bar{x}' = \frac{\Sigma(x_u - x_e)}{n} = \frac{10}{6} = 1,67.$$

Az adatok különbségeinek szórása (4), illetve (6) szerint:

$$s' = \sqrt{\frac{\Sigma x'^2_i - \frac{(\Sigma x'_i)^2}{n}}{n-1}} = \sqrt{\frac{28 - \frac{10^2}{6}}{6-1}} = 1,51.$$

A mintából számolt t_m értéke (20) szerint ($\mu = 0$):

$$t_m = \frac{\bar{x}'}{s_{\bar{x}'}} = \frac{\bar{x}'}{s'} \cdot \sqrt{n} = \frac{1,67}{1,51} \sqrt{6} = 2,72.$$

Legyen a szignifikancia szint: $\alpha = 0,05 \rightarrow 5\%$.

A szabadságfok: $(n-1) = (6-1) = 5$.

A 7. táblázat kétoldali próbákhoz tartozó $p = 0,05$ oszlopának és az 5-ös szabadságfoknak megfelelő 5. sorának metszete megadja a levágási határ értékét:

$$t_{le} = 2,571.$$

Mivel: $t_m = 2,72 > t_{le} = 2,571$,

így t_m a kritikus tartományba esik, tehát a **nullhipotézist elvetjük**. Az egyperces lélegzet-visszatartás tehát **szignifikáns** pulzusváltozást okozott (5%-os szignifikancia szint mellett).

KÉTMINTÁS *t*-PRÓBA

Az ide vonatkozó **korábbi kérdés**: *VAN-E KÜLÖNBESÉG* a lányok és a fiúk pulzusszáma között (lásd 14. megjegyzés)?

Általánosabban megfogalmazva: két populáció eloszlásának várható értéke különbözik-e?

$$t_m \text{ kiszámítása: } t_m = \frac{\bar{x}_1 - \bar{x}_2}{s^*} \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}}, \text{ ahol } s^* = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2}}, \quad (21)$$

t_m szabadságfoka: $n_1 + n_2 - 2$,

ahol Q az (6)-ban bevezetett jelöléssel ekvivalens az 1. illetve 2. mintára vonatkoztatva. (Megfigyelhetjük, hogy ez a kifejezés formailag nagyon hasonlít a (20) egyenlőséghez.)

KORRELÁCIÓS *t*-PRÓBA

Az ide vonatkozó **korábbi kérdés**: *VAN-E KAPCSOLAT* az életkor és a szem akkomodációs képessége között?

Általánosabban megfogalmazva: a korrelációs együttható és a minta elemszámának figyelembevételével **mondhatjuk-e két mennyiségről** (a változásaik alapján), **hogy kapcsolat van közöttük?**

$$t_m \text{ kiszámítása: } t_m = r \cdot \sqrt{\frac{n-2}{1-r^2}}, \quad (22)$$

t_m szabadságfoka: $n-2$,

ahol n a mérési pontok (összetartozó x és y értékpárok) száma, r pedig a (19) képlettel bevezetett korrelációs együttható. (Mivel az eljárás további része gyakorlatilag megegyezik az előző két t -próbánál bemutatott lépésekkel (lásd a 13. és a 14. megjegyzéseket), ezért itt erre külön nem térünk ki.)

A 8. táblázatban összefoglaltuk a t -próbák itt megbeszélt változatait. A konkrét esetben alkalmazandó statisztikai próba típusát elsősorban a feltett kérdés jellege dönti el. Ennek egyértelmű megfogalmazását segíti az elvégzendő mérés, illetve a rendelkezésre álló adatok ismerete.

14. megjegyzés (kidolgozott megoldás):

- H_0 : a lányok és a fiúk pulzusszáma között nincs különbség, azaz $\mu_{\text{lányok}} - \mu_{\text{fiúk}} = 0$.
 - H_1 : $\mu_{\text{lányok}} - \mu_{\text{fiúk}} \neq 0$, (kétoldali próba).
- Az alábbi táblázat 6 lány és 9 fiú pulzusszámadatait tartalmazza.

$x_{\text{lányok}}$	$x_{\text{fiúk}}$
74	71
87	63
62	70
79	74
71	71
77	69
	82
	56
	78
$\bar{x}_{\text{lányok}} = 75$	$\bar{x}_{\text{fiúk}} = 70$

Az átlagok kiszámításából úgy tűnhet, hogy a lányoknak magasabb a pulzusszáma. Vajon szignifikáns-e ez az eltérés (vagy csak a véletlen okozhatja)?

Feltesszük, hogy a t -próba elvégzéséhez minden feltétel teljesül, (még az is, hogy a két minta szórása is egyforma, amit külön próbával ellenőrizhetnénk).

Számítógép segítségével kiszámítjuk az adatainkhoz tartozó p valószínűséget: $p = 0,296$. (A t -próba függvény paramétereinek beállításánál figyelembe vesszük a kétoldali próbát és az azonos szórásokat.)

Mivel p elég nagy, sokkal nagyobb, mint 0,05, a szokásos szignifikancia szint, ezért a nullhipotézist nem vethetjük el, hiszen ellenkező esetben majdnem 30%-os lenne az elsőfajú hiba, ami annak a valószínűségét méri, hogy igaz hipotézist vetettünk el. Így H_0 -t megtartjuk, tehát azt mondhatjuk, hogy a mintáinkban megfigyelhető különbség nem szignifikáns.

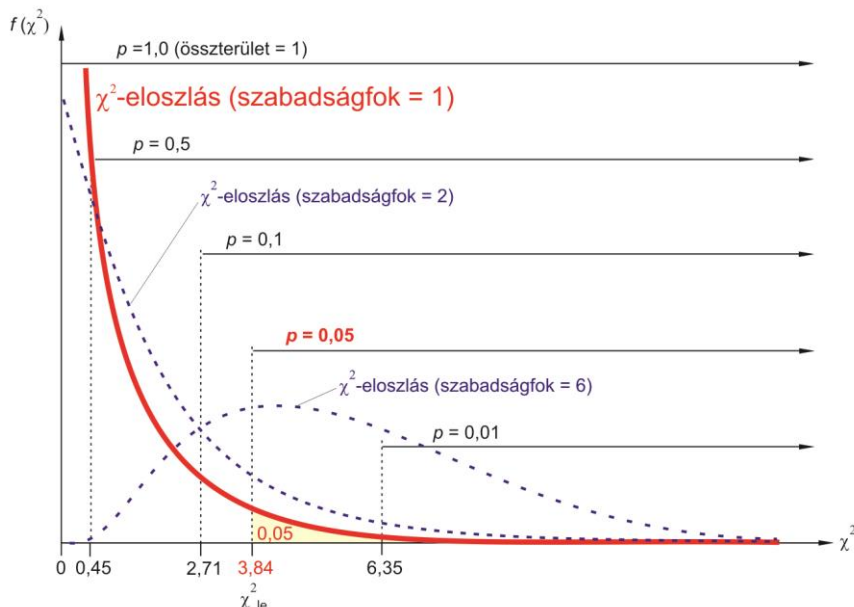
	egymintás <i>t</i> -próba	kétmintás <i>t</i> -próba	korrelációs <i>t</i> -próba
Egy tipikus kérdés az orvostudomány területéről	Hatásos-e a kezelés? (Van-e változás a várt irányban?)	Van-e különbség két kezelés hatása között?	Van-e a kapcsolat két mennyiség között?
az ehhez tartozó nullhipotézis	a kezelés hatástalan	a két kezelés ugyanolyan hatású	nincs kapcsolat
a próbával eldöntendő kérdés pontosabb megfogalmazása	származhat-e a minta egy μ_0 várható értékű eloszlásból?	származhat-e a két minta ugyanabból az eloszlásból?	kis r esetén mondhatjuk-e mégis, hogy van korreláció két (folytonos) változó között?
a nullhipotézis pontosabban	$\mu - \mu_0 = 0$	$\mu_1 - \mu_2 = 0$	valójában nincs korreláció a két változó között
mérés	egy mennyiséget mérünk egy mintán	ugyanazt a mennyiséget mérjük két mintán	két mennyiséget mérünk ugyanazon a mintán
t	$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$	lásd (21)	$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$
szabadságfok	$n - 1$	$n_1 + n_2 - 2$	$n - 2$

8. táblázat. Összefoglaló a leggyakrabban használt t -próbákról.

χ^2 -PRÓBÁK

A t -próbákat természetesen csak számszerű (folytonos) jellemzőkre alkalmazhatjuk. Mit tegyünk azonban olyan esetekben, amikor jellemzőink **kategoriális** típusúak? Ilyenkor is tudunk számszerű összehasonlításokat tenni a gyakoriságok alapján. Többek között erre használhatók a χ^2 -próbák.

Kérdés: különbözik-e egy megállapítható tulajdonság (tünet) előfordulásának gyakorisága két populációban? Például: különbözik-e a tüdőrák előfordulásának a gyakorisága a tüdőbetegek körében a dohányzók és a nemdohányzók között (lásd 15. megjegyzés)?



26. ábra. Különböző szabadságfokú χ^2 -eloszlások, valamint az 1-es szabadságfokú χ^2 -eloszlás néhány levágási határa.

Az alábbi táblázatban a két populációt az A illetve B csoport jelzi. A vizsgálat összesen n főre terjedt ki. A vizsgált tulajdonsággal **rendelkezők** száma a illetve c , a **nem rendelkezők** száma b illetve d fő.

A próbák elvégzésének egyetlen feltétele, hogy egyik cellában se legyen túl kevés adat: a két legkisebb részösszeg szorzata legyen nagyobb, mint $5n$. (Csak megjegyezzük, hogy amennyiben ez nem teljesül, akkor az ún. Fisher-féle egzakt teszt használható, de annak ismertetésére itt nem térünk ki.)

	a vizsgált tulajdonság		összesen
	megvan	nincs meg	
A csoport	a	b	$a+b$
B csoport	c	d	$c+d$
összesen	$a+c$	$b+d$	n

9. táblázat

Nullhipotézis: nincs különbség a vizsgált tulajdonság előfordulási gyakoriságában a két csoport között, azaz a szóban forgó gyakoriság a két mintában azonos, tehát a (23) összefüggés szerint kiszámolt χ_m^2 -érték is zérus, illetve nem különbözik szignifikánsan zérustól.

$$\chi_m^2 \text{ kiszámítása: } \chi_m^2 = \frac{n \cdot (ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}. \quad (23)$$

χ_m^2 szabadságfoka: 1.

15. megjegyzés (kidolgozott megoldás):

- H_0 : a tüdőrák előfordulásának a gyakorisága a dohányzók és a nemdohányzók között megegyezik, azaz $\chi^2 = 0$.
- H_1 : a két gyakoriság különböző, tehát $\chi^2 \neq 0$.

Az alábbi táblázat egy tüdőgondozóból vett minta alapján a két vizsgált csoportban megállapított gyakoriságokat szemlélteti (a sor- és oszlop-összegekkel együtt, $n = 61$).

	tüdőrák	nincs	
dohányzó	14	13	27
nemdohányzó	9	25	34
	23	38	61

Mivel $23 \cdot 27 = 621 > 5 \cdot 61 = 305$, ezért a próba elvégezhető.

A (23) összefüggés szerint:

$$\chi_m^2 = \frac{61 \cdot (14 \cdot 25 - 9 \cdot 13)^2}{23 \cdot 38 \cdot 34 \cdot 27} = 4,13.$$

Látható, hogy $\chi_m^2 \neq 0$, de vajon szignifikáns-e ez az eltérés (vagy csak a véletlen okozza)?

Legyen a szignifikancia szint: $\alpha = 0,05 \rightarrow 5\%$.

A szabadságfok (2×2-es táblázat esetén): 1.

A χ^2 táblázat $p = 0,05$ oszlopának és az 1 szabadságfoknak megfelelő 1. sorának metszete megadja a levágási határ értékét (lásd a 26. ábrát is):

$$\chi_{le}^2 = 3,84$$

$$\text{Mivel: } \chi_m^2 = 4,13 > \chi_{le}^2 = 3,84$$

így χ_m^2 a kritikus tartományba esik, tehát a **nullhipotézist elvetjük**. Ezek szerint a tüdőrák előfordulásának a gyakoriságában tapasztalt eltérés a dohányzók és a nemdohányzók között szignifikáns (5%-os szignifikancia szint mellett).

16. megjegyzés: (az előző kidolgozott megoldás az általános módszer szerint)

Megismételjük a hipotéziseket:

- H_0 : a tüdőrák előfordulásának a gyakorisága a dohányzók és a nemdohányzók között megegyezik, azaz $\chi^2 = 0$
- H_1 : a két gyakoriság különböző, tehát $\chi^2 \neq 0$.

Számítógép segítségével igen egyszerűen eljuthatunk az eredményhez. A **megfigyelt** adatok (gyakoriságok) kontingencia táblázatát már az előző részben elkészítettük:

	tüdőrák	nincs	
dohányzó	14	13	27
nemdohányzó	9	25	34
	23	38	61

Kontingencia táblázat

Ezután készítsük el a **várt** gyakoriságokat tartalmazó segéd-kontingencia táblázatot is, amelyben (változatlan oszlop- és sorösszegek mellett) a relatív gyakoriságok megegyeznek a két csoportban, tehát ahol

$$\frac{a}{b} = \frac{c}{d}.$$

	tüdőrák	nincs	
dohányzó	10,18	16,82	27
nemdohányzó	12,82	21,18	34
	23	38	61

Segéd-kontingencia táblázat

Az adott statisztikai programcsomagból kiválasztjuk a *khi*-próba (vagy ehhez hasonló) függvényt, amelynek argumentumában csak a két tömböt kell kijelölnünk. A számítógép kiszámítja az adatainkhoz tartozó p valószínűséget, esetünkben: $p = 0,042$.

Mivel p elég kicsi, kisebb, mint a szokásos szignifikancia szint (0,05), ezért a nullhipotézist elvetjük. Így H_1 lép életbe, ami azt jelenti, hogy a két gyakoriság különböző (5%-osnál kisebb szignifikancia szint mellett).

A vizsgált tulajdonságnak azonos arányú gyakoriságából, azaz

$$\frac{a}{b} = \frac{c}{d} \text{-ből}$$

következik, hogy: $ad - bc = 0$, tehát χ^2 -nek zérus értéke elégíti ki a nullhipotézist. A kiszámolt χ^2 -érték szignifikáns – nem szignifikáns voltáról az 1-es szabadságfokú χ^2 -eloszlás alapján döntünk.

χ^2 -PRÓBÁK ÁLTALÁNOSABBAN

Az előbbieken csak a χ^2 -próbák legegyszerűbb, de gyakran alkalmazható változatát ismertük meg (2×2-es táblázat). Mind a populációk száma, mind a megvizsgált tulajdonságok száma nagyobb is lehet.

Általános: ($l \times m$)-es táblák (l sor m oszlop) (kontingencia táblázat). **Függőségek vizsgálatára** alkalmazható. Függelenség esetén azonosak a relatív gyakoriságok.

H_0 : függetlenség, ekkor χ^2 várható értéke 0.

χ^2 kiszámítása:

$$\chi^2 = \sum \left[\frac{(O - E)^2}{E} \right], \quad (24)$$

ahol O a **megfigyelt** (Observed) gyakoriság. E a **várt** (Expected) gyakoriság, a szabadságfok pedig $(l-1)(m-1)$.

A várt gyakoriságok kiszámítása. A megfigyelt gyakoriságokat tartalmazó kontingencia táblázat mellé elkészítendő a H_0 fennállása esetén várt gyakoriságokat tartalmazó segéd-kontingencia táblázat. Az oszlop és sorösszegek, mindkét kontingencia táblázatban azonosak. A segéd-kontingencia táblázat egy adott cellájában a várható gyakoriság a H_0 fennállása esetén megegyezik a saját oszlopösszeg és a saját sorösszeg szorzatának és a minta elemszámának hányadosával (sorösszeg · oszlopösszeg / ($l \cdot m$)) (lásd a 16. megjegyzést).

NÉHÁNY PÉLDA AZ ORVOSI STATISZTIKA KÖRÉBŐL

Az orvosi gyakorlatban sokszor előfordul, hogy egy változó ugyan folytonos mennyiség, de mondjuk a t -próbák alkalmazhatóságának egyes feltételeit (lásd a 27. oldalon, például azt, hogy a változó legyen normális eloszlású) nem elégíti ki. Ha ilyenkor mégis t -próbát alkalmazunk, ezt általában megtehetjük, de számolnunk kell azzal, hogy következtetéseink kevésbé lesznek érvényesek.

Ilyen esetekre vonatkozó egyik megoldási lehetőségről, nevezetesen a változó megfelelő transzformációjáról már beszéltünk. Ha ez sem segíthet, akkor használhatók az ún. **nem-paraméteres módszerek**. (Az elnevezés arra utal, hogyha például **ismeretlen az eloszlás típusa**, akkor természetesen a paramétereiről sem beszélhetünk.) Ezek leggyakrabban használt változata ún. rangokat rendel az egyes értékekhez, valamilyen szabály szerint és az eredeti értékek helyett ezeket dolgozza fel. A **rangok** általában egész számok és az eredeti értékeknek egy adott sorban elfoglalt helyét jelölik. Folytonos mennyiségek esetében is előfordulhat (a mérések, vizsgálatok véges pontossága miatt), hogy két vagy több azonos érték szerepel az eredeti értékek sorában. Mivel ilyenkor teljesen önkényes lenne, hogy melyiket helyezzük előre, ezért az azonos értékek mindegyike a megfelelő rangok átlagát kapja meg. (Például, ha egy sorban az első két érték azonos, akkor az 1; 2; 3; ... rangok helyett az 1,5; 1,5; 3; ... rangokat használjuk.)

A nem-paraméteres próbák nem támasztanak különösebb feltételeket (mint pl. a t -próbnál a változó normális eloszlása) **a változókkal szemben**, mondjuk csak annyit, hogy legyen a változó ordinális. Felmerülhet a kérdés, hogy akkor miért nem használjuk ezeket minden esetben. Kimutatható, hogy azonos minta (minták) esetében a t -próbák használatával, ha alkalmazhatóak, kisebb szignifikancia szint mellett hozható döntés. Ezért kell minden esetben megvizsgálni, vajon alkalmazható-e a t -próba, vagy sem.

A következőkben egy olyan példát mutatunk be, amelyik a kétmintás t -próba helyett használható.

MANN-WHITNEY U-TEST

Vizsgáljuk meg azt a kérdést, hogy **hatásos-e az aszpirin, mint fejfájás csillapító?** (Fontos megjegyeznünk, hogy egy ilyen kérdés csak igen sokrétű klinikai vizsgálat, illetve statisztikai elemzés végén válaszolható meg, tehát itt ennek a folyamatnak csak egy elemét mutatjuk be.)

H_0 : Az aszpirin nem hatásos fejfájás csillapító.

Mivel a fejfájás mértéke nem mérhető mennyiség, ezért például a következő kísérletet végezhetjük el. Válasszunk ki két szokványos fejfájással küszködő csoportot, az egyik (1) csoport (n_1 fő) aszpirint, a másik (2) kontroll csoport (n_2 fő) ún. placebót, azaz hatóanyag nélküli tablettát kap. Természetesen a kísérleti alanyok nem tudják, hogy mit kapnak (vak teszt). Egy adott idő eltelte után megkérjük őket, hogy mondják egy 0-tól 10-ig terjedő skálán, minősítsék tapasztalatukat. (Ezen a skálán 0 jelenti a változatlan, vagy erősebb fejfájást, tehát a teljesen hatástalan szert; 10 pedig a teljesen elmúlt fejfájást, azaz az igazán hatásos szert) Az így szerzett adatok alapján történik a kiértékelés. Eredetileg ez az adathalmaz egy ordinális változó kategóriáinak felelne meg (nem javult, kicsit javult, stb.), amelyeket csak az egyszerűbb kiértékelés kedvéért számszerűsítettünk, így az eloszlás ismeretlen, tehát annak paramétereiről nem is beszélhetünk. (Emiatt használjuk a t -próba helyett ezt a nem-paraméteres tesztet.)

A kapott értékeket nagyság szerint sorba rendezzük függetlenül attól, hogy melyik csoportból származnak. Ha H_0 igaz, a két csoportból származó értékek teljesen véletlenszerűen helyezkednek el. Osszuk ki a rangokat és legyen T_1 az 1., T_2 a 2. csoportban a rangok összege. Korábbi matematikai ismereteinkből tudjuk, hogy 1-től $(n_1 + n_2)$ -ig a két csoportban együttesen a rangok összege $(n_1 + n_2) \cdot (n_1 + n_2 + 1) / 2$, ami nyilvánvalóan $(T_1 + T_2)$ -vel egyenlő. Az „átlagos” rang ennek alapján $(n_1 + n_2 + 1) / 2$. Ha tehát H_0 igaz, akkor $T_1 \approx n_1 \cdot (n_1 + n_2 + 1) / 2$ és $T_2 \approx n_2 \cdot (n_1 + n_2 + 1) / 2$, vagy másképpen például: $T_1 - n_1 \cdot (n_1 + n_2 + 1) / 2 \approx 0$.

Ebből az alábbi transzformációval kapott (z) változó jó közelítéssel $N(0,1)$ eloszlású, (megjegyezve azt, hogy túl kis esetszámok esetén nem használható):

$$z = \frac{T_1 - n_1 \cdot (n_1 + n_2 + 1) / 2}{\sqrt{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1) / 12}}$$

Így, ha H_0 igaz, akkor $z = 0$, az ettől való eltérést pedig vagy a véletlen okozza, vagy az, hogy nem igaz az eredeti feltevés. A döntést úgy hozzuk meg, hogy a kiszámított z -t adott p mellett összehasonlítjuk a standard normális eloszlásból ($N(0,1)$) származó értékkel (lásd 17. megjegyzés).

SPEARMAN-KORRELÁCIÓ

A már ismertetett, a 18. oldalon bevezetett (Pearson-féle) korrelációs együttható (r) leggyakrabban használt „nem paraméteres megfelelője”, amely ordinális jellegű változók kapcsolatának vizsgálatára is alkalmazható.

A Spearman-féle korrelációs együttható (r_s) kiszámításához a vizsgált változókon először rangsor-transzformációt hajtunk végre, majd az összetartozó rangszámok különbségeinek (d_i) négyzetét képezzük, melyeket az alábbi összefüggésbe helyettesítjük:

$$r_s = 1 - \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (25)$$

ahol n az összetartozó értékpárok száma (lásd 18. megjegyzés 32. oldal).

Az így számított együttható is r -hez hasonlóan -1 és 1 közötti értékeket vehet fel. A 0 közeli értékek gyenge (semmilyen), az 1 -hez és -1 -hez közeli értékek erős pozitív illetve negatív irányú korrelációt („együtt változást”) jelentenek.

17. megjegyzés: (kidolgozott megoldás)

- H_0 : az aszpirin mint fejfájás csillapító nem hatásos, $z = 0$
- H_1 : az aszpirin mint fejfájás csillapító hatásos, tehát $z > 0$ (egyoldaltú próba).

1. csoport ($n_1 = 8$):

7,5 8,3 9,1 6,2 5,4 8,3 6,5 8,4

2. csoport ($n_2 = 9$):

3,1 5,6 4,5 6,2 5,1 5,3 5,5 4,1 4,3

Rendezzük sorba őket!

sorszám		rang
1	3,1	1
2	4,1	2
3	4,3	3
4	4,5	4
5	5,1	5
6	5,3	6
7	5,4	7
8	5,5	8
9	5,6	9
10	6,2	10,5
11	6,2	10,5
12	6,5	12
13	7,5	13
14	8,3	14,5
15	8,3	14,5
16	8,4	16
17	9,1	17

$$T_1 = 7 + 10,5 + 12 + 13 + 14,5 + 14,5 + 16 + 17 = 104,5$$

$$z = \frac{104,5 - 8 \cdot (8 + 9 + 1) / 2}{\sqrt{8 \cdot 9 \cdot (8 + 9 + 1) / 12}} = 3,13$$

Legyen most a szignifikancia szint nagyon kicsi: $\alpha = 0,01 \rightarrow 1\%$.

$N(0,1)$ eloszlás helyett használhatjuk a ∞ szabadságfokú t -eloszlást (ami azzal ekvivalens).

Az egyoldaltú t -próbákhoz tartozó táblázat $p = 0,01$ oszlopának és a ∞ szabadságfoknak megfelelő sorának metszete megadja a levágási határ értékét,

$$z_{\alpha} = 2,33.$$

Mivel: $z_m = 3,13 > z_{\alpha} = 2,33$,

így z_m a kritikus tartományba esik, tehát a **nullhipotézist elvetjük** (1%-os szignifikancia szint mellett). (A $3,13$ levágási értékhez tartozó valószínűség $p < 0,0009$, tehát nagyon kicsi annak a valószínűsége, hogy ekkora eltérést a véletlen okozzon.)

A fenti vizsgálat alapján tehát azt mondhatjuk, hogy az aszpirin, mint fejfájás csillapító hatásos.

A 2×2-es TÁBLÁZATOK EGYÉB FELHASZNÁLÁSAI

A χ^2 -próbákon kívül a 2×2-es táblázatok más vizsgálatokban is előfordulhatnak. Gyakori kérdés, hogy vajon egyes rizikófaktorok milyen szerepet játszanak egyes betegségek kialakulásában, előfordulásában. A vizsgálatok során gyakorisági adatokat gyűjtünk két mintából és az adatokat egy 2×2-es táblázatban rendezzük el.

rizikófaktor jelen van	beteg	
	igen	nem
igen	<i>a</i>	<i>b</i>
nem	<i>c</i>	<i>d</i>

10. táblázat.

Az *a*, *b*, *c*, *d* betűk a megfelelő gyakorisági adatokat jelölik. Ebben az esetben is azt várjuk, hogy a gyakorisági értékek arányai a beteg és nem beteg csoportokban azonosak, ha a rizikófaktorától független a betegség kialakulása. A körülményektől függően két alapvető vizsgálati módszer alakult ki.

Az egyik esetben az egészséges embereket két csoportra osztják a rizikófaktor megléte vagy hiánya alapján. Ezeket a csoportokat hosszabb időn keresztül megfigyelik és a vizsgálat végén elkészítik a fenti kontingencia-táblázatot. Ilyenkor beszélünk követéses (cohort) vizsgálatról.

KÖVETÉSEK (PROSPEKTÍV, COHORT) VIZSGÁLAT EGY PÉLDÁBAN

Kérdés: Mekkora kockázatot jelent a dohányzás a szívinfarktus kialakulása szempontjából?

H_0 : nincs összefüggés a dohányzás és a szívinfarktus kialakulása között.

Válogatás szempontja: dohányzik – nem dohányzik.

Később (pl. 10 év múlva) megvizsgáljuk az infarktus előfordulási gyakoriságát a két csoportban. Az adatokat egy 2×2-es táblázatban rendezzük el.

rizikófaktor (dohányzik)	beteg (infarktus)		összesen
	igen	nem	
igen	<i>a</i>	<i>b</i>	<i>a+b</i>
nem	<i>c</i>	<i>d</i>	<i>c+d</i>
összesen	<i>a+c</i>	<i>b+d</i>	<i>n=a+b+c+d</i>

11. táblázat.

Meghatározzuk a kockázat mértékét a két csoportban:

- A **kockázat** mértéke a dohányosok között: $a/(a+b)$.
- A **kockázat** mértéke a nem dohányzók között: $c/(c+d)$

Ezekből kiszámítjuk az ún. **relatív kockázatot** (*RR* – relative risk):

$$RR = \frac{a/(a+b)}{c/(c+d)} = \frac{a \cdot (c+d)}{c \cdot (a+b)}, \quad (26)$$

ami megadja, hogy hányszor gyakoribb az adott betegség kockázata adott rizikófaktor megléte esetében. *RR* várható értéke 1, tehát ha H_0 igaz, akkor $RR = 1$. Következő lépésként kiszámítjuk az $\ln(RR)$ érték standard hibáját (*SE* – standard error):

$$SE(\ln RR) = \sqrt{\frac{1 - a/(a+b)}{a} + \frac{1 - c/(c+d)}{c}}. \quad (27)$$

Meghatározzuk az *RR* érték konfidencia intervallumát (figyelembe véve, hogy a fenti összefüggés az $\ln(RR)$ standard hibáját adja meg, emiatt a konfidencia intervallum nem szimmetrikus az *RR*-re nézve), ami a 95% valószínűségi

18. megjegyzés: (kidolgozott feladat)

Két orvos munka-alkalmassági vizsgálatot végez és a kérdés az, hogy mennyire hasonló a minősítésük? Miután egy hét fős csoportot mindketten megvizsgáltak egy 1-10-ig terjedő skálán fejezték ki az egyének alkalmasságát. (1 jelenti a nem nagyon alkalmas, 10 pedig a nagyon alkalmas minősítést.)

Az eredményeket az alábbi táblázatban foglaltuk össze:

egyen	orvos A	orvos B	rang A	rang B	d^2
A	10	9	7	6,5	0,25
B	7	9	6	6,5	0,25
C	1	3	1	2	1
D	4	6	3,5	3	0,25
E	4	7	3,5	4	0,25
F	3	1	2	1	1
G	5	8	5	5	0
összesen:					3

Amennyiben a Pearson-féle korrelációs együtthatót (*r*-t) határozzuk meg az első két oszlop alapján, akkor a (19) összefüggés szerint $r = 0,78$ -at kapunk.

A Spearman-féle korrelációs együttható a (25) összefüggés szerint:

$$r_s = 1 - (6 \cdot 3) / (7(7^2 - 1)) = 0,95.$$

Ez a két adat éppen azt fejezi ki, hogy bár a számszerű minősítés alapján a két orvos különbözőképpen ítél, de akit az egyik (A) alkalmasabbnak minősít azt a másik (B) is.

Fontos megemlíteni, hogy a rangsor-transzformáció alkalmazásával a kapott változó értékei **kevesebb információt** tartalmaznak, mint amennyit az eredeti változó hordozott. Megfelelően alkalmazva azonban éppen így nyílik lehetőség a változó információ-tartalmából a **lényeg kiemelésére**.

tartomány (kb. 2-szer a standard hiba, pontosabban az 1,96-szorosa) és ezzel jellemezhetjük az érték megbízhatóságát. Ha az intervallum magába foglalja az 1 értéket, akkor megtartjuk a nullhipotézist, ellenkező esetben elvetjük (az adott megbízhatósági szinten).

A másik esetben beteg és nem beteg embereket választanak ki és őket osztják két csoportra, a rizikófaktor megléte vagy hiánya alapján. Ekkor beszélünk visszatekintő (case-control) vizsgálatról.

VISSZATEKINTŐ (RETROSPEKTÍV, CASE-CONTROL) VIZSGÁLAT EGY PÉLDA KAPCSÁN

Kérdés: Lehet-e a betegség kialakulásában szerepet játszó tényező az adott rizikófaktor?

H_0 : nincs összefüggés a rizikófaktor jelenléte és a betegség kialakulása között.

Válogatás szempontja: beteg (eset) – nem beteg (kontroll). Az adatokat egy 2x2-es táblázatban rendezzük el.

rizikófaktor	beteg		összesen
	igen (eset)	nem (kontroll)	
igen	a	b	$a+b$
nem	c	d	$c+d$
összesen	$a+c$	$b+d$	$n=a+b+c+d$

12. táblázat.

Meghatározzuk a betegség esélyét mindkét csoportban:

- A betegség esélye a faktor megléte esetében: a/b .
- A betegség esélye a faktor hiánya esetében: c/d .

Majd kiszámoljuk az ún. **esélyhányadost** (OR – odds ratio):

$$OR = \frac{a/b}{c/d} = \frac{a \cdot d}{b \cdot c} . \quad (28)$$

Az előzőkhöz hasonlóan kiszámítjuk az $\ln(OR)$ érték standard hibáját:

$$SE(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}} . \quad (29)$$

Majd szintén meghatározzuk a megfelelő konfidencia intervallumot. Ha ez tartalmazza az 1 értéket, akkor a betegség esélyét nem növeli az adott rizikófaktor (az adott megbízhatósági szinten).

A két módszer összehasonlítása: A követéses vizsgálat jellegénél fogva néha évekig tart, ezért költséges is. A hosszú idő miatt általában nagyobb elemszámmal kell dolgozni az esetleges kiesések kompenzálására. Viszont abban az esetben, amikor a rizikófaktor előfordulás ritka, gyakorlatilag csak ez a módszer alkalmazható, mert visszatekintéses esetben nehezebb biztosítani a megfelelő nagyságú gyakorisági értékeket.

Ezek az adatok természetesen χ^2 -próbával is feldolgozhatóak. Miért használják mégis ezeket a módszereket? Előnye a fenti eljárásoknak az, hogy a χ^2 -próbával szemben egy számszerű érték is rendelkezhető a döntéshez.

A DIAGNOSZTIKAI TESZTEK JELLEMZÉSÉRE SZOLGÁLÓ STATISZTIKAI MÓDSZEREK

Fontos területe az orvosi statisztikának a diagnosztikai módszerek vizsgálata, illetve megbízhatóságuk jellemzése. Ugyanis attól függetlenül, hogy egy diagnosztikai teszt eredménye pozitív vagy negatív, lehet helyes, de téves is. Egy téves teszt alapján végzett terápia beláthatatlan következményekkel járhat. Ezért igen fontos tudni, hogy mennyire bízhatunk meg a kapott eredményben.

Természetesen az orvos általában nem egy vizsgálat alapján állítja fel a diagnózist, de munkája során tudnia kell, hogy melyik vizsgálat mennyire értékes egy adott betegség kimutatásában.

Az alábbiakban nem térünk ki a terület részletes ismertetésére, csupán bemutatjuk, és röviden jellemezzük az előforduló mennyiségeket.

Kérdés: Mennyire megbízható a teszt eredménye?

Végezzük el a **tesztet**, majd ennek **eredményét hasonlítsuk össze a „valósággal”**. Az alábbi táblázatban a VN , $ÁP$, $ÁN$, VP mennyiségek a lehetséges kimenetek gyakoriságát mutatják.

a valóságban	a teszt eredménye		összesen
	negatív	pozitív	
egészséges	VN (valódi negatív)	$ÁP$ (ál pozitív)	$VN+ÁP$
beteg	$ÁN$ (ál negatív)	VP (valódi pozitív)	$ÁN+VP$
összesen	$VN+ÁN$	$ÁP+VP$	$n=VN+ÁP+ÁN+VP$

13. táblázat.

Egy ilyen táblázatban a „valóság”, a valóságnak megfelelő eredmények helyessége mindig kérdéses lehet. Honnan tudjuk a „valódi” eredményeket? Erre azt mondhatjuk, hogy vannak elég egyértelmű esetek, amikor eleve ismerjük a válaszokat, illetve vannak olyan ún. „arany tesztek”, ezek közismert, nagyon nagy megbízhatóságú tesztek, amelyek segítségével hozzájuthatunk a szükséges információhoz.

Az e területen használt mennyiségek definíciója és jellemzésük:

- **Prevalencia** (előfordulás): a betegek relatív gyakorisága: $(VP+ÁN)/n$.
- **Szenzitivitás (érzékenység)**: a pozitív eredmény relatív gyakorisága az összes olyan esetben, amikor pozitív eredményt kellene kapnunk: $VP/(VP+ÁN)$.
- **Specifitás** (fajlagosság): a negatív eredmény relatív gyakorisága az összes olyan esetben, amikor negatív eredményt kellene kapnunk: $VN/(VN+ÁP)$.

Ezeket a mennyiségeket általában %-os formában adják meg. Ideális esetben az érzékenység és a specifitás is 100%, nagy megbízhatóságú tesztek esetében közelítenek ahhoz.

A gyakorlatban a következő szempontokat veszik figyelembe:

- Ha a diagnózist követő terápia könnyen végrehajtható és nem jelent különösebb kockázatot a paciens számára, akkor a magas érzékenység a cél. (Ilyenkor minden gyanús esetet kiszűrünk).
- Ellenkező esetben a magas specifitás a fontosabb. (Ne kezeljünk nem beteg embereket!).

Bár az érzékenység és a specifitás is fontos jellemzői egy tesztnek, talán még fontosabb az, hogy egy pozitív teszt eredmény alapján milyen mértékben valószínűsíthető a betegség megléte, illetve negatív eredmény alapján annak hiánya.

Ezekről a következő mennyiségek adnak információt:

- **Relevancia** (korrekt pozitivitás, PVP, predictive value positive): a valós pozitív eredmény valószínűsége, ha a teszt pozitív lett: $VP/(VP+ÁP)$.
- **Szegregancia** (korrekt negativitás PVN, predictive value negative): a valós negatív eredmény valószínűsége, ha a teszt negatív lett: $VN/(VN+ÁN)$.

Ezek a mennyiségek (de a többi is) függenek az előfordulás értékétől is. Gyakori előfordulás esetében a korrekt pozitivitás sokkal nagyobb érték, mint ritka betegségek esetében. Fordított a helyzet a korrekt negativitásnál.

A fentiek összefoglalásaként lásd a 14. táblázatot.

HATÁRÉRTÉK		egyéb magyar elnevezések		angol elnevezések	
TESZT EREDMÉNY					
VALÓJÁBAN egészséges beteg		HATÁRÉRTÉK		diszkriminációs szint, levágási pont	
PREVALENCIA ELTERJEDTSÉG (w)		$\frac{\text{ÁN}}{\text{VP}} / \frac{\text{VN}}{\text{ÁP}} = \frac{(\text{VP} + \text{ÁN})}{(\text{VP} + \text{VN} + \text{ÁN} + \text{ÁP})}$ $w = (\text{de-sp}) / (\text{se-sp})$		a betegség gyakorisága a populációban	
SZENZITIVITÁS VALÓDI POZITÍV ARÁNY (se)		$\frac{\text{VP}}{\text{ÁN} + \text{VP}} = \text{se}$		érzékenység, pozitívak a betegek között, mennyire fogja ki a kifogandókat, VPa (valódi pozitív arány)	
SPECIFICITÁS VALÓDI NEGATÍV ARÁNY (sp)		$\frac{\text{VN}}{\text{ÁN} + \text{VN}} = \text{sp}$		fajlagosság, negatívak az egészségesek között, mennyire zárja ki a kizárandókat, VNa (valódi negatív arány)	
ÁLNEGATÍV ARÁNY másodfajú hiba		$\frac{\text{ÁN}}{\text{ÁN} + \text{VP}} = (1 - \text{se})$		negatív a betegek között, ÁNa (álnegatív arány)	
ÁLPOZITÍV ARÁNY elsőfajú hiba		$\frac{\text{ÁP}}{\text{VN} + \text{ÁP}} = (1 - \text{sp})$		pozitív az egészségesek között, ÁPa (álpozitív arány)	
RELEVANCIA KORREKT POZITIVITÁS pozitív prediktív érték (PVP)		$\frac{\text{VP}}{\text{ÁP} + \text{VP}} = \text{PVP}$ $= \text{se} \cdot w / [\text{se} \cdot w + (1 - \text{sp})(1 - w)]$		a betegség jelenlétének valószínűsége, ha a teszt pozitív, betegek a pozitívak között	
SZEGREGANCIA KORREKT NEGATIVITÁS negatív prediktív érték (PVN)		$\frac{\text{VN}}{\text{VN} + \text{ÁN}} = \text{PVN}$ $= \text{sp}(1 - w) / [\text{sp}(1 - w) + (1 - \text{se})w]$		a betegség hiányának valószínűsége, ha a teszt negatív, egészségesek a negatívak között	
TÉVES FIGYELEM- FELKELTŐ ARÁNY		$\frac{\text{ÁP}}{\text{VP} + \text{ÁP}} = (1 - \text{PVP})$		egészségesek a pozitívak között	
TÉVES MEGNYUGTATÁSI ARÁNY		$\frac{\text{ÁN}}{\text{VN} + \text{ÁN}} = (1 - \text{PVN})$		betegek a negatívak között	
EFFEKTIVITÁS DIAGNOSZTIKUS EFFEKTIVITÁS (de)		$\frac{(\text{VP} + \text{VN})}{(\text{VP} + \text{VN} + \text{ÁN} + \text{ÁP})} = \text{de}$ $\text{de} = \text{se} \cdot w + \text{sp}(1 - w)$		Pontosság, besorolási pontosság korrektül klasszifikáltak...	

14. táblázat.

A PRÓBÁK MENETE

19. megjegyzés:

Az itt elmondottak a statisztikának csak igen kis részét fedik le. Részletesebb és alaposabb ismereteket például a felhasznált irodalom valamelyikének áttanulmányozásával nyerhetnek.

A statisztikai próbákhoz alkalmazható statisztikai táblázatok a 30. FÜGGELÉK fejezetben találhatók.

Előfordul, hogy a statisztikai adatgyűjtést és az abból levont következtetéseket meghamisítják, vagy úgymond kozmetikázzák. Ezért mindig szenteljünk figyelmet az adatok és a belőlük készített statisztikák hitelességére, átláthatóságára, ellenőrizhetőségére. Az alábbi ironikus idézet is arra utal, hogy az egykori brit miniszterelnököt valószínűleg többször félretájékoztatták a statisztikák nehézkes ellenőrizhetőségét kihasználva.

„Én csak abban a statisztikában hiszek, melyet magam hamisítottam...”

Sir Winston Churchill

Kiválasztjuk a megfelelő próbát.

Kiindulás: Nullhipotézis (H_0), (általában valamilyen jellemző érték nulla).

Elmélet: a populáció ismeretében a kérdés egyértelműen eldönthető (nulla vagy nem nulla).

Gyakorlat: véges elemszámú minta esetében az érték a véletlen miatt is eltérhet 0 - tól.

feltevések	
H_0 igaz (csak a véletlen okoz eltérést)	H_0 nem igaz

Minta: véletlen minta, tekintetbe véve speciális orvosi követelményeket (pl. adott betegség megléte vagy sem, kizáró okok vannak-e vagy sem.)

Kiválasztunk egy alkalmas, elegendően kicsi (általában $p \leq 0,05$) α -értéket.

A mintából kiszámoljuk a próbának megfelelő alkalmas értéket (t , χ^2 , ...). Az adott érték elméleti eloszlása megadja a véletlen eltérések mértékét.

Kikeressük a megfelelő táblázatból az adott α -értékhez tartozó elméleti értéket.

Döntés: (pozitív egyoldali próba esetén)

1. Elméleti érték \geq számolt érték — H_0 -t megtartjuk, nincs ok elvetni.
2. Elméleti érték $<$ számolt érték — H_0 -t elvetjük, kicsi a valószínűsége annak, hogy az eltérést a véletlen okozza.

A döntés szignifikancia szintje: α , ami egyúttal a 2. esetben a döntés hibás voltának mértéke is.

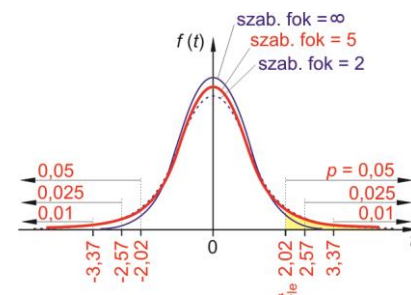
Felhasznált irodalom:

- Dinya E.: Biometria az orvosi gyakorlatban, Medicina, Bp. 2007
- Reiczigel J., Harnos A., Solymosi N.: Biostatisztika, Pars Kft, Nagykovács 2007
- Hajtman B.: Bevezetés a matematikai statisztikába, Akadémiai Kiadó, Bp. 1972
- Farkas M.: Matematikai Kislexikon, Műszaki Könyvkiadó, Bp. 1974
- Sváb J.: Biometriai módszerek a kutatásban, Mezőgazdasági Kiadó, Bp. 1981
- Juvancz I., Paksy A.: Az orvosi biometria alapjai, Medicina, Bp. 1981
- Mendenhall, W.: Introduction to probability and statistics, Duxbury Press, Boston 1987
- Norman T. J. Bailey: Statistical methods in biology, Cambridge university press 1993
- P. Armitage, G. Berry: Statistical methods in medical research, Oxford 1994
- Beth Dawson, Robert G. Trapp: Basic and Clinical Biostatistics, Lange Medical Books/McGraw-Hill 2001

1. STATISZTIKAI TÁBLÁZATOK

t-ELOSZLÁS

szabadsági fok	<i>p</i> (valószínűség, egyoldalú próba)						
	0,4	0,25	0,1	0,05	0,025	0,01	0,005
	<i>p</i> (valószínűség, kétoldalú próba)						
	0,8	0,5	0,2	0,1	0,05	0,02	0,01
1	0,325	1,000	3,078	6,314	12,70	31,82	63,65
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704
60	0,255	0,679	1,296	1,671	2,000	2,390	2,66
120	0,254	0,677	1,289	1,658	1,980	2,358	2,617
∞	0,250	0,674	1,282	1,645	1,960	2,326	2,576



2. FELADATOK

- Négy békavörösvérsejt hossza: 18, 17, 21 és 18 μm . Számítsuk ki az átlagot, a szórást és az átlag szórását! ($\bar{x} = 18,5 \mu\text{m}$, $Q_x = 9 \mu\text{m}^2$, $s_x = 1,73 \mu\text{m}$, $s_{\bar{x}} = 0,87 \mu\text{m}$)
- 25 hallgató testmagasságának átlaga 170 cm, a szórás 8 cm. Becsülje meg a várható értéket 95 %-os konfidencia szinten! ($\bar{x} - 2s_x = 166,8 \text{ cm}$, $\bar{x} + 2s_x = 173,2 \text{ cm}$)
- Két mérési sorozatban (egyik esetben 5, a másikban 20 mérési ponthoz jutottunk) a mérési eredményekből kiszámolt korrelációs együttható egyaránt 0,6-nak adódott. Beszélhetünk-e 5 %-os szignifikancia szinten korrelációról az egyik, illetve a másik esetben? ($r = 0,6$, $n_1 = 5$, $n_2 = 20$, tehát $f_1 = 3$, $f_2 = 18$, és $t_1 = 1,299$, nem igazol korrelációt, ellenben $t_2 = 3,18$ még 1 %-os szignifikanciaszinten is korrelációt mutat.)
- Egy gyógyszerkészítmény hatóanyagtartalma (tömegarányban kifejezve) 20 %. A forgalombahozatal előtt ellenőrizni kell a gyógyszer eltarthatóságát. Egy ilyen ún. stabilitásvizsgálat során a hűtőszekrényben tárolt mintasorozat analízise 20,1; 19,8; 18,9; 19,7; 19,9; 20,2 % hatóanyagtartalmat mutatott. Szignifikáns különbség van-e a talált és a névleges érték között, vagy csupán véletlen szórásról van szó? ($p=27\%$, nem szignifikáns)
- Az előző példában leírt vizsgálattal párhuzamosan szobahőmérsékleten is tettek el egy hat elemből álló mintasorozatot. A kapott értékek ugyanolyan tárolási idő után itt 19,6; 18,9; 19,5; 20,1; 19,3; 19,4 % voltak. Elő kell-e írni, hogy a készítményt hűtőszekrényben kell tárolni? ($p=2,1\%$, tehát elő kell írni)
- Egy kisvárosban a téli influenzaszzezon megelőzően 424 személyt influenza vakcinával oltottak be, 425 fő pedig placebót kapott. A vakcináltak közül 105, a placebósok körében 140 légúti megbetegedés történt; a megbetegedés súlyossága miatt a vakcináltak közül 31-en, a kontrollcsoportból 55-en kényszerültek orvoshoz fordulni. Mit lehet mondani a fentiek alapján az influenza elleni vakcinálás hasznosságáról? (Mind a megbetegedettek mind a súlyosan megbetegedettek aránya szignifikánsan kisebb a vakcináltak körében: $p<1\%$)
- Egy bizonyos szembetegség (nem artériás típusú ischaemiás opticus neuropathia) sikeres műtéti korrekciójáról jelent meg 1989-ben egy közlemény. Minthogy e betegségben korábban semmiféle hatásos kezelési módszer nem volt ismert, ezt a műtétet sok helyen alkalmazni kezdték. Rövidesen azonban eredménytelen beavatkozásokról is megjelentek beszámolók, ezért szembavették 25 klinikai centrum 244 ilyen betegét, akik közül 119 főn elvégezték a műtétet, 125 beteg nem. A felmérés eredménye:

	javult	nem javult	romlott	összesen
műtött	39	52	28	119
nem műtött	53	56	16	125

Vizsgáljuk meg statisztikai módszerekkel:

- valóban nagyobb-e a műtét után javultak száma? (nem, $p>0,1$)
 - valóban nagyobb arányban következett-e be romlás a műtöttek állapotában? (igen, $p<0,05$)
8. Egy orvoscsoport növekedéshormon-hiányos gyermekek osteocalcin (OC) szintjét határozta meg kétéves növekedéshormon-kezelés előtt és után. Kontrollként egészséges gyermekeken is végeztek méréseket. Az eredmények az alábbiak:
- | OC koncentráció ($\mu\text{g/l}$) | | egészséges
gyermekek |
|-------------------------------------|--------------|-------------------------|
| beteg gyermekek
kezelés előtt | kezelés után | |
| 8,5 | 14,5 | 17,3 |
| 7,2 | 10,2 | 16,9 |
| 8,3 | 12,7 | 17,7 |
| 15,5 | 16,5 | 18,6 |
- A vizsgálat annak kiderítésére irányult, hogy valóban
- Alacsonyabb-e a növekedéshormon-hiányos gyermekek kezelés előtti OC szintje, mint az egészségeseké? (igen, szignifikáns)
 - Van-e a kezelésnek hatása az OC szintre? (igen, szignifikáns)

9. A menopausa (klimax) utáni periódusban lévő nők gyakran szenvednek kalciumhiányban. Erre utal az a megfigyelés is, hogy az ebbe a csoportba tartozó nők esetében viszonylag magas a csonttörések gyakorisága. Vajon igaz lehet-e az a feltételezés, hogy az ösztrogénszint menopausa utáni csökkenése összefüggésbe hozható a kalciumhiánnyal? Ennek felderítésére egy orvos három női csoportban meghatározta a csontok ásványi anyag tartalmát g/cm^3 -ben, amely arányos azok kalcium tartalmával. Az első csoportba olyan 25-50 éves nők tartoztak, akik petefészek eltávolításon estek át, s így kimutathatóan ösztrogénhiányban szenvedtek. A második csoportot ugyanilyen korú, egészséges, még a menopausa előtt álló nők alkották. A harmadik csoportba tartozó nők már átestek a változási perióduson. Mindhárom csoportban 14-en voltak. Az eredmények a következők:

	1. csoport	2. csoport	3. csoport
átlag (g/cm^3)	0,93	1,21	0,92
standard hiba (g/cm^3)	0,04	0,03	0,04

A rendelkezésre álló adatok alapján hogyan válaszolna a feltett kérdésre? (Igen, összefüggésbe hozható)