



# Alapfogalmak áttekintése



Pszichológiai statisztika, 1. alkalom

# Hipotézisek

Milyen a jó null hipotézis?

? H0: Léteznek kitűnő tanuló diszlexiások.

? H1: Nem léteznek.

Sokkal inkább:

H0: Nincs diszlexiás kitűnő tanuló általános iskolában Mo-on.

H1: Van.

Fontos definiálni, hogy kit tekintünk diszlexiásnak, kitűnő tanulónak és mely populációt vizsgáljuk.

?H0: A taktaharkányi kerti törpe gyárban nincsenek a nők bér tekintetében hátrányosan megkülönböztetve.

Sokkal inkább:

H0: A taktaharkányi kerti törpe gyárban a nők ugyanannyit keresnek, mint a férfiak.

H1: A taktaharkányi kerti törpe gyárban a nők keresete eltér a férfiakétól.

Ha minden más tekintetben párhuzamba állítható női és férfi csoportot lehet elkülöníteni, az eredeti kérdés is vizsgálható.

A jó null hipotézis egyértelműen vizsgálható és tagadása bizonyítható. A null hipotézis állítja, hogy nincs kapcsolat, nincs változás, a vizsgált jelenségek függetlenek, ugyanakkorák.

# Hipotézisek

A jó alternatív/ellenhipotézis következik a  $H_0$  tagadásából, alternatívák nincsenek, éppen ezért egyes kutatók (Harris, 1997 in Vargha, 2000) cáfolják az egyoldali ellenhipotézis létjogosultságát.

$H_0$ : X antidepresszáns nincs hatással a depressziós betegek állapotára (BDI kérdőíven elért pontszám kezelés után nem változik).

$H_1$ : X antidepresszáns kedvező hatással van a depressziós betegek állapotára

!!! X antidepresszáns rontja a depressziós betegek állapotát.

$H_0$  tagadása (ha  $H_0$  megfelelően valószínűtlen):

$H_1$

$H_0$  megtartása:

Nem  $H_1$  tagadása, hanem  $H_1$  tagadása nem bizonyított.

Diszlexiás példa:

Diszlexiás kitűnő tanulók (ált. isk., Mo.) léte a vizsgált minta alapján nem bizonyított.

Kerti törpe gyár példa:

A gyárban dolgozó nők és férfiak bére nem tér el szignifikáns mértékben.

Antidepresszáns példa:

A gyógyszer 95%-os valószínűséggel nem befolyásolja a depressziós betegek állapotát.

# Változók

---

Statisztikai változó:

A megfigyelési egységek jellemzői

Itt is fontos az egyértelműség!

Egy változó lehet adott kérdésre adott válasz. Pl.

?Ön hazudik-e?

a. Igen b. Nem c. Ritkán

Inkább:

a. Gyakran b. Ritkán c. Soha

?Ön szerint a globális felmelegedés oka...

a. a dohányzás b. a gépkocsiforgalom

Inkább

a. a dohányzás b. a gépkocsiforgalom c. Egyéb, éspedig.....

Vagy

Ön szerint melyik felelős leginkább az alább felsoroltak közül a globális felmelegedésért?

# Változók

Skála típusai:

Nominális

Milyen cigarettát szív?

Ordinális (értékei sorba rendezhetők)

Dohányzik?

Gyakran, ritkán, soha.

Intervallum (értékei sorba rendezhetők, fokozatai azonos mértékű különbséget jelentenek)

Dohányzási szokásokat mérő kérdőív alapján a függőség.

Arány (értékei sorba rendezhetők, fokozatai azonos mértékű különbséget jelentenek, van abszolút nulla értéke, értelmezhető az értékek más számokkal való osztása, szorzása)

Naponta elszívott cigaretták száma

Kvalitatív: Nominális, ordinális. Kvantitatív: Intervallum, arány.

Értékek száma alapján: diszkrét (nominális és majdmind ordinális), folytonos (összes többi)

# Populáció, minta

Fontos tisztázni a megfigyelés egységét.

A pszichológiában általában a személy, de lehet az iskola vagy család is.

Populáció:

A számunkra érdekes megfigyelési egységek összessége, melyekről állításokat akarunk megfogalmazni.

Minta:

A kísérleti vagy vizsgálati egységek összessége.

A minta akkor ideális, ha reprezentatív a populációra nézve, ha kellően nagy és ha véletlen, független mintaválasztást alkalmazunk. A minta kiválasztása lehet random mintavétel vagy önkéntes jelentkezés.

Csak random mintavétel esetében következtethetünk a populációra!

Megkülönböztetünk kísérletet (kontrollált körülmények) és megfigyelést.

PI. Belső és külső motiváció és kreativitás (Teresa Amabile vizsgálata)

Ok-okozati összefüggést csak kísérlet esetében lehet levonni, mert lehetnek olyan változók a megfigyelésben, amelyeket nem ismerünk, de a csoportba sorolást és az eredményt is befolyásolják.

PI. A fizikai erő a kerti törpe gyárban hat a teljesítményre és az a fizetésre.

Megfigyelés is hasznos: -ok-okozati összefüggés feltárása nem mindig cél (kínai és európai reakciója nyugtatóra)

-megalapozhat kísérletet (a szoba színe és a teljesítmény)

-ritkán mégiscsak levonható ok-okozati összefüggés (gyermekrablás és PTSD)

# Populáció, minta



# Randomizálás

---

Mindenki azt mondja, hogy véletlen mintavételt alkalmaz, ill. random módon sorol csoportokba azonban szinte senki nem teszi ezt tökéletesen.

## 1. Példa :

A ruhák fekete színének megőrzését vizsgáló mosószerekkel kapcsolatos attitűdvizsgálat során 10 üzletben a mosószert vásárlókat kérdőíves módszerrel vizsgálják.

2 probléma

## 2. Példa:

A cél a debreceni munkahelyek tipikus szervezeti struktúrájának vizsgálata. A mintavétel a Debrecen térkép segítségével történik. A véletlenszerűen kiválasztott koordinátákhoz legközelebb eső munkahelyeket választják be a mintába.

[http://travel.yahoo.com/p-map-485339-map\\_of\\_debrecen-i](http://travel.yahoo.com/p-map-485339-map_of_debrecen-i)

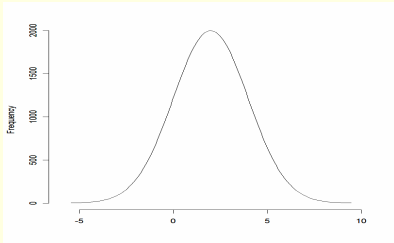
## 3. Példa:

Különböző munkaterápiák hatásának vizsgálata pszichiátrián kezelt betegekre. A folyosón szembe jövő első tíz beteg az udvaron dolgozik, következő tíz iratmegsemmisítést végez, azután következő tíz a konyhán segít.



# Mintavétel-hipotézisvizsgálat

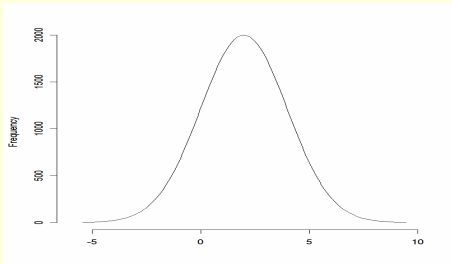
A populáció:  $N(2,2)$ ,  $\mu=2$ ,  $\sigma=2$



$\bar{x}_1 = 1.6$     $\bar{x}_2 = 2.3$     $\bar{x}_3 = 1.8$     $\bar{x}_4 = 2.1$

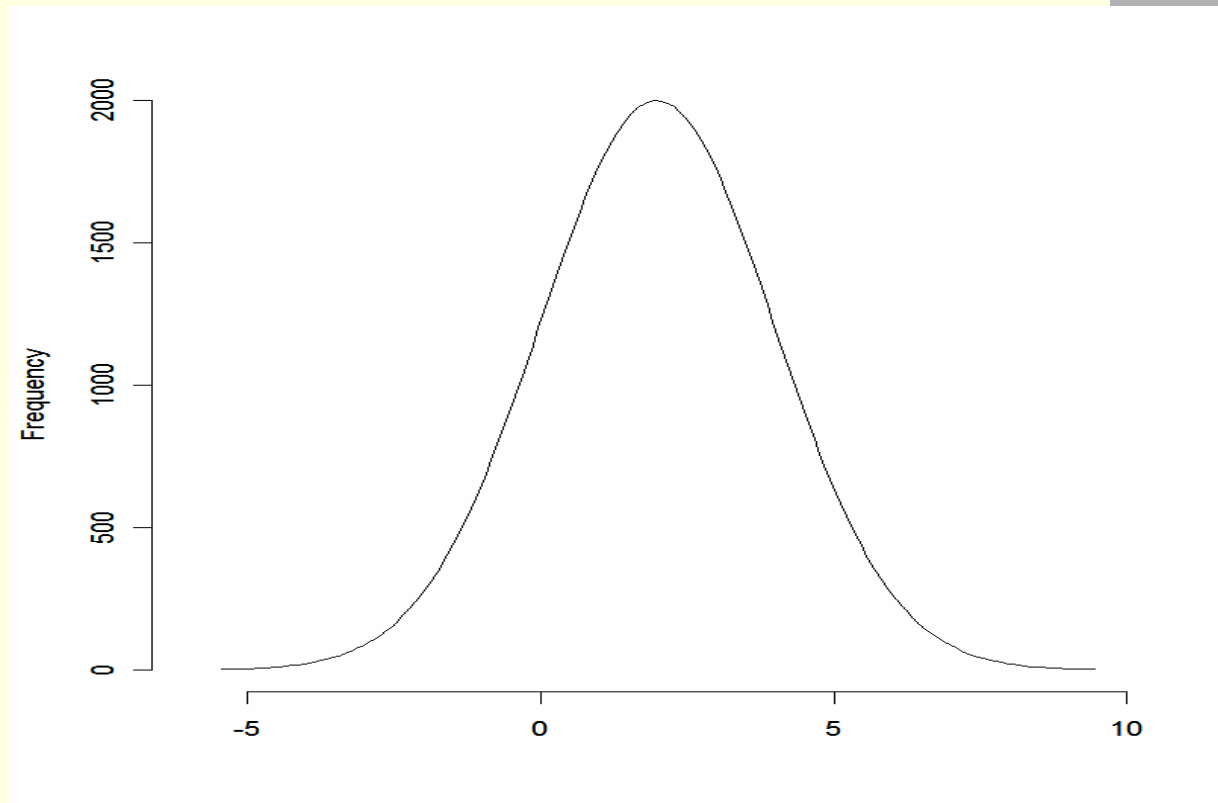
Számos mintavétel, a mintaátlagokat kiszámoljuk

A mintaátlag a legritkább esetben egyezik a populációátlaggal, viszont jobb híján belőle következtetünk a pop. átlagra. Minél nagyobb a minta annál jobb a közelítés.



Centrális határeloszlás elve: bármely eloszlásból kellően nagy elemszámú számos mintát véve a paraméterek eloszlása normális lesz. Ez teszi kiemelkedővé a normális eloszlást.

# Normál eloszlás



Egy normál eloszlás sűrűségfüggvénye, Gauss-görbe  
 $N(2,2)$  azaz  $\mu = 2, \sigma^2 = 2$

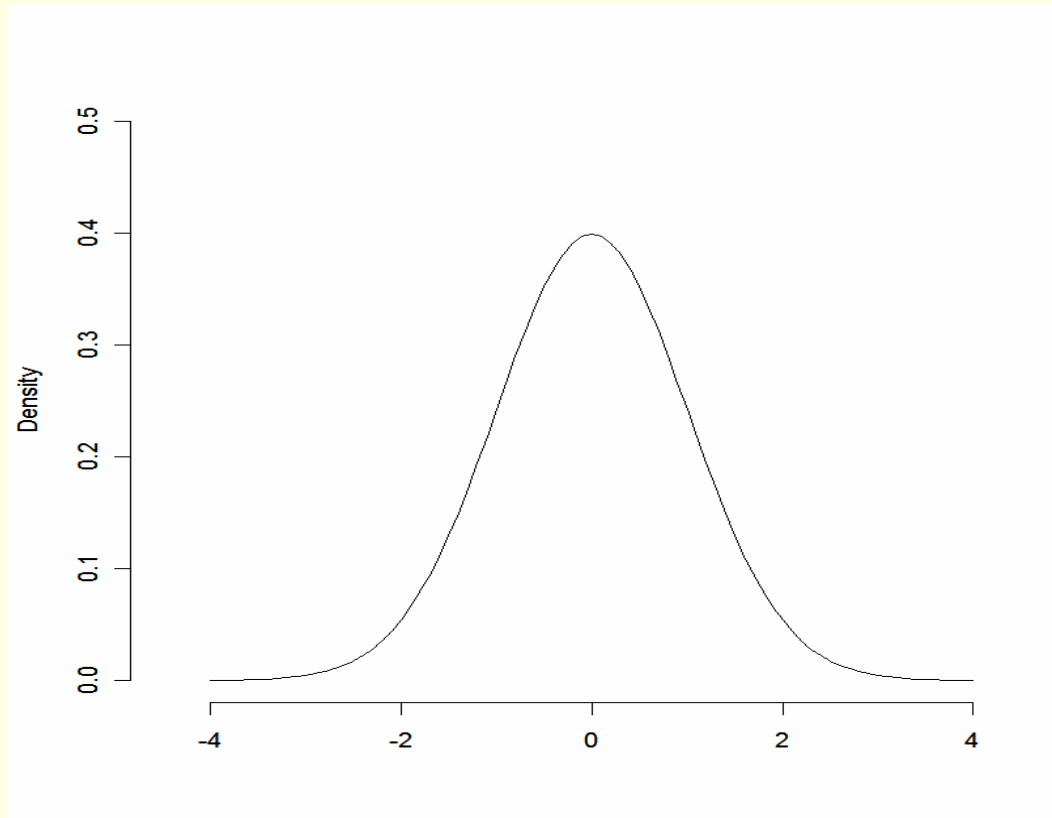
# Normál eloszlás

- Mindig szimmetrikus
- $\gamma_2 = E(Z^4) - 3$  csúcsossági együttható mindig 0
- Értékei az átlag körüli 3 szórásnyi távolságon belül figyelhetőek meg
- Az átlag körüli 2 szórásnyi intervallum tartalmazza az adatok 68%-át  
4 szórásnyi intervallum 95%-át  
6 szórásnyi intervallum 99.8%-át
- Ha a minta elemszám elég nagy, bármely változó értékeiből kiválasztott véletlen minta átlagának eloszlása megközelíti

Centrális határeloszlás elmélet:

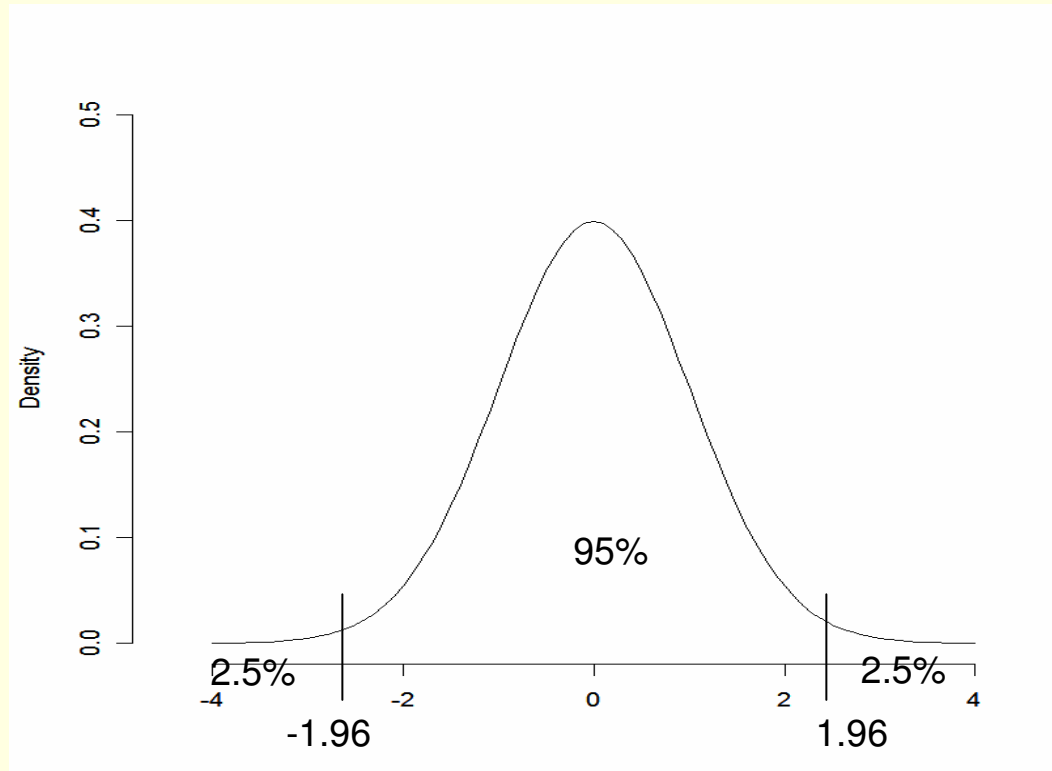
A populációból vett minták átlagának eloszlása normális.

# Standard normál eloszlás



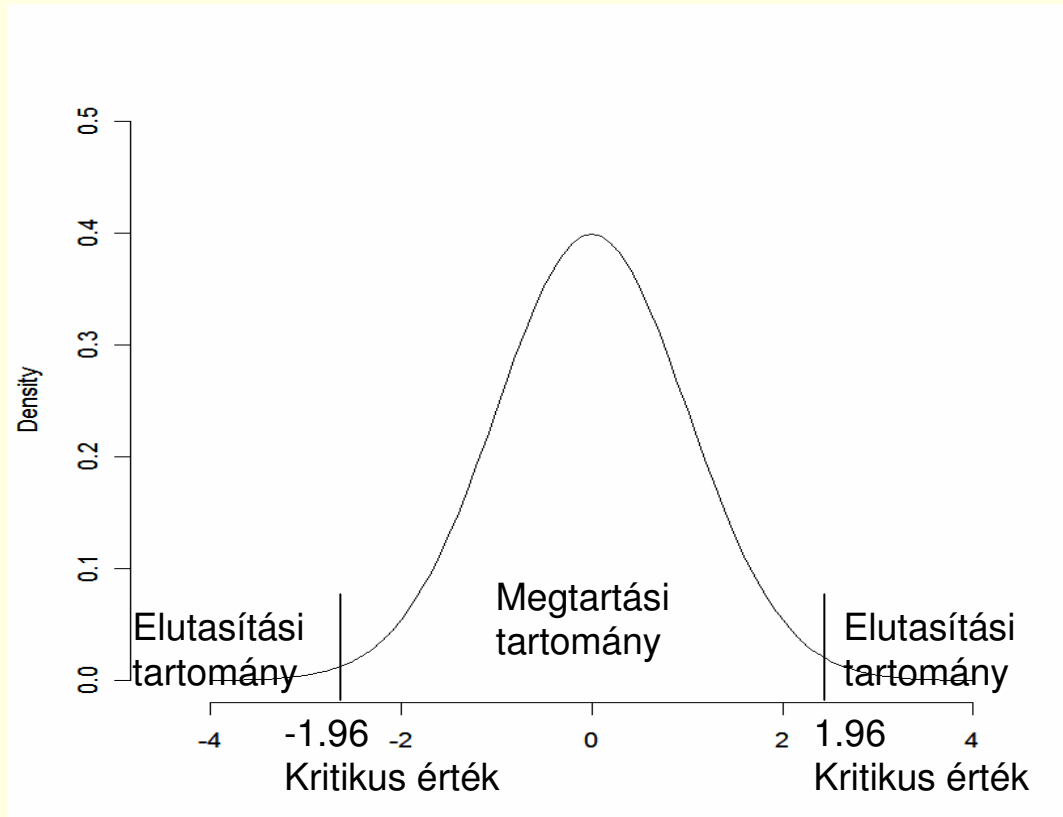
$N(0,1)$

# Standard normál eloszlás



Z eloszlás, a görbe alatti terület adja egy-egy intervallum valószínűségét

# Standard normál eloszlás



# Következtetések

---

A releváns következtetésekhez szükséges (lehet):

- Kontroll a vizsgálatot esetleg torzító tényezők felett
- Kezdeti állapot ismerete

A jó következtetés (Occam's razor) a lehető legegyszerűbb magyarázattal szolgál (letisztult, leborotválja a sallangokat).

**Törpegyár: A vezető szexuális frusztráltságából adódóan hátrányosan megkülönbözteti a nőket.**

Einstein:

A modell legyen olyan egyszerű, amennyire csak lehet, de nem egyszerűbb.

**Antidepresszáns: Hatással van a depressziós betegekre.**

# Következtetések

Döntés	Valóság	
	H0 igaz	H0 nem igaz
H0-t megtartjuk		
H0-t elutasítjuk		



# Következtetések

	Valóság	
Döntés	H0 igaz	H0 nem igaz
H0-t megtartjuk	jogos elfogadás	II. fajú hiba
H0-t elutasítjuk	I. fajú hiba	jogos elutasítás

# Következtetések

---

Az első fajú hiba( $\alpha$ ) meghatározza

a szignifikancia szintet, a téves elutasítás valószínűségét.

Leggyakoribb értékei 0.05, 0.01.

A másodfajú hiba ( $\beta$ ) meghatározza

a statisztikai próba erejét ( $1 - \beta$ ), a kihagyás és indirekt módon a jogos megtartás valószínűségét.

Leggyakrabban 0.2.

Egyoldali ellenhipotézis esetén beszélhetünk ún. harmadik fajú hibáról (Vargha, 2000). Ez annak valószínűsége, hogy H1 és H2 közül rosszul választunk ( $\alpha/2$ )

# Következtetések és mintanagyság

---

Amennyiben meghatározható a legkisebb lényeges eltérés a H0-tól ( $\delta$ ) szakmai szempontok alapján, akkor a szükséges mintanagyság kiszámolható.

$$N \approx \frac{10(1-\beta)s^2}{\delta^2}$$

R-ben: `power.t.test`, `power.anova.test`

A mintanagyság nem változtatható az eredmény fényében!!

Túl nagy minta sem jó.

# Normál eloszlás és következtetések

---

$\alpha = 0.05$  esetén megállapíthatjuk (ha  $H_0$ -t elutasítjuk), hogy az eredmény 5%-os szinten szignifikáns, jelölése ( $p < 0.05$ ).

(ha  $H_0$ -t megtartjuk)  $\alpha = 0.05$ -ös szignifikanciaszint mellett  $H_1$  nem bizonyított.

$\alpha = 0.1$  esetén (ha  $H_0$ -t elutasítjuk) az eredmény 1%-os szinten szignifikáns, jelölése ( $p < 0.01$ ).

(ha  $H_0$ -t megtartjuk)  $\alpha = 0.1$ -ös szignifikanciaszint mellett  $H_1$  nem bizonyított.

$\alpha = 0.01$  tendencia