

## 📖 Mit jelent, ha a korreláció $r = -0.97$ és a p-érték $1.74 \times 10^{-88}$ ?

Tegyük fel, hogy két változó kapcsolatát szeretnénk vizsgálni. A kapott eredmény:

- $r = -0.97$
- $p = 1.74 \times 10^{-88}$
- $n = 150$

Ez a kettő együtt három fontos dolgot mond el:

1. Milyen erős a kapcsolat
2. A kapcsolat „valódi-e” vagy lehet pusztán véletlen
3. Mennyire pontos becslés az  $r = -0.97$

Nézzük sorban.

### 1) 📊 Mit jelent az $r = -0.97$ ?

A korrelációs együttható  $r$  megmutatja:

- milyen erős a kapcsolat (0 = semmilyen, 1 vagy  $-1$  = tökéletes)
- milyen irányú (pozitív vagy negatív)

Az  $r = -0.97$  azt jelenti, hogy:

- **nagyon erős,**
- **negatív,**
- **szinte tökéletes** lineáris kapcsolat van a két változó között.

Ez azt sugallja, hogy egyik változó növekedése szinte teljesen együtt jár a másik csökkenésével — szinte egy egyenesre illeszkednek.

### 2) 🎲 Mit jelent a p-érték: $p = 1.74 \times 10^{-88}$ ?

A p-érték azt fejezi ki:

**Ha a valóságban nincs semmilyen kapcsolat a két változó között, akkor mekkora az esélye, hogy véletlenből mégis ekkora (vagy ennél nagyobb) korrelációt kapunk?**

Tehát a p-érték nem az  $r$  bizonytalanságát méri, és nem azt mondja, hogy „mik a jövőbeli esélyek”.

A te esetében:

- A  $p = 1.74 \times 10^{-88}$  elképesztően kicsi szám.
- Ez azt jelenti:  
**ha valójában nincs kapcsolat, akkor**  
annak esélye, hogy véletlenből  $-0.97$ -hez hasonlóan erős korrelációt kapjunk, gyakorlatilag nulla.

Ezért mondjuk, hogy a kapcsolat **statisztikailag szignifikáns**.

### 3) 🗨️ Miért „kétoldali” ez a p-érték?

A statisztikai programok alapértelmezése a következő hipotézis:

- **Nullhipotézis:** nincs kapcsolat  $\rightarrow \rho = 0$
- **Alternatív hipotézis:** van kapcsolat  $\rightarrow \rho \neq 0$

Ez a „nem egyenlő” jel miatt **kétoldali teszt**.

A kétoldali teszt azt vizsgálja:

Kaphatok-e a véletlenből *bármilyen* nagyon erős korrelációt, akár pozitív (+0.97), akár negatív (-0.97)?

Ezért szerepel a p-érték képletében  $|r|$  (az abszolút érték) — mert csak az számít, mennyire „extrém” az eredmény, nem az iránya.

Ha valaki egyoldali tesztet szeretne (pl. csak negatív kapcsolatot keres), azt külön kell kérni — itt nem ez történt.

### 4) ✖️ Hogyan vetjük el a nullhipotézist?

A klasszikus szabály:

- Ha  $p < 0.05$ , akkor elvetjük a nullhipotézist  $\rightarrow$  van kapcsolat.
- Itt pedig  $p = 0.000000\dots$  (88 nulla)... 174  $\rightarrow$  jóval kisebb.

Ezért:

Nagyon erős bizonyíték van arra, hogy a két változó között valódi kapcsolat van.  
Nem pusztán a véletlen műve, hanem tényleg összefüggenek.

### 5) 🎯 Mi a helyzet a bizonytalansággal? (Konfidencia-intervallum)

Bár az  $r = -0.97$  nagyon erős érték, mégis csak egy minta alapján becsültük ( $n = 150$ ).

A valóságos korrelációt  $\rho$ -val jelöljük:

$$\rho \quad (\text{populációs korreláció})$$

A konfidencia-intervallum megmondja:

Milyen tartományba esik  $\rho$  nagy valószínűséggel?

A mintád alapján:

$$95\% \text{ CI} \approx [-0.978 ; -0.959]$$

Ez azt jelenti:

- a valódi kapcsolat szinte biztosan nagyon erős negatív,
- még ha újra is vennénk a mintát, az eredmény végig ebben a tartományban lenne,
- az  $r = -0.97$  jó és stabil becslés.