

# KUTATÁSTERVEZÉS

► **Kis János**

► **SZIE** ► **ÁOTK**  
► **Biológiai Intézet**

**2010-11-16**

<ftp://www.univet.hu/biostat/>



# TEMATIKA

## I. A kutatómunka

1. Alapkutatás – alkalmazott kutatás – alkalmazás
2. A kutatómunka természete: jelenség, modell

## II. Tervezés

1. Kérdésfeltevés
2. Miért érdekes?
3. Hipotézisek, jóslások ( $\rightarrow$  változók)
4. Mérések: becslés, mérési hibák, ismételhetőség
5. Mintavétel
6. Megközelítési módok: megfigyelés, kísérlet, filogenetikai összehasonlítások, formális modellek, szimuláció
7. Engedélyek: állatvédelem, természetvédelem
8. Adatok kezelése

# I. Mi a kutatómunka?

*“Now my own suspicion is that the Universe is not only queerer than we suppose, but queerer than we **can** suppose.” Haldane 1927*

*[Alice to the Red Queen] “ ‘There’s not use trying,’ she said: ‘one can’t believe impossible things.’*

*‘I daresay you haven’t had much practice,’ said the Queen. ‘When I was your age, I always did it for half-an-hour a day. Why, sometimes I’ve believed as many as six impossible things before breakfast.’” Carroll 1871*

*“Researchers who find their task boring and uninspiring are unlikely to be very creative. A sense of wonder is important.” Maindonald 2000*

# I. Mi a kutatómunka?

## 1. Alap kutatás – alkalmazott kutatás – alkalmazás

|                 | <b>Alap</b>                           | <b>Alkalmazott</b>                                    | <b>Alkalmazás</b>      |
|-----------------|---------------------------------------|-------------------------------------------------------|------------------------|
| <b>Cél</b>      | A világ megismerése                   | Ismeretek alkalmazhatóságának kitalálása, bizonyítása | Ismeretek alkalmazása  |
| <b>Kérdés</b>   | Hogyan működik?<br>Miért úgy működik? | Hogyan hasznosítható?                                 | Hogyan kell használni? |
| <b>Eredmény</b> | Új ismeret, elmélet                   | Új módszer, alkalmazás                                | Új termék              |

# Alap kutatás – alkalmazott kutatás – alkalmazás

## Alap

## Alkalmazott

## Alkalmazás

|                 |                                            |                                                                                    |                                              |
|-----------------|--------------------------------------------|------------------------------------------------------------------------------------|----------------------------------------------|
| <b>Kérdés</b>   | Milyen szerkezetű az anyag?                | Hogyan lehet az $E_r$ -t hasznosítani?                                             | Hogyan kell egy erőművet felépíteni?         |
| <b>Eredmény</b> | Radioaktivitás felfedezése                 | $E_r \rightarrow E_e$ , biztonságos erőmű                                          | $E_e$ -ellátás atomerőművel                  |
| <b>Kérdés</b>   | Hogyan képesek vírusok emlősökben működni? | Hogyan lehet oltóanyagot készíteni?                                                | Hogyan lehet egy konkrét járványt elkerülni? |
| <b>Eredmény</b> | Immunrendszer megértése                    | Vakcinák                                                                           | Állatállomány oltása                         |
| <b>Kérdés</b>   | Hogyan szaporodnak a hangyaboglárkalepkék? | Miért pusztultak ki Angliából?                                                     | Mit kell tenni a visszatelepítésükért?       |
| <b>Eredmény</b> | Lepke-tápnövény-hangya kapcsolat           | Legeltetés felhagyása $\rightarrow$ magas növényzet $\rightarrow$ hangyák eltűnnek | Legeltetni kell – nem őrködni                |

**E területek között a határok sokszor nem különíthetők el élesen!**

# Egy kis kitérő:

## definíciók – verbális modellek – fogalmi gondolkodás

- A definíció nem cél, hanem eszköz: ha nagyon mereven használjuk, a megismerés helyett a definíciókon rágódunk
- A definíció is egy modell: addig kell használni, amíg operatív
- Szükséges: ismereteinket a nyelv útján csereberéljük, ha egy szón mindenki mást ért, mindenki mondja a magáét, nincs gondolatcsere  
→ akkor érdemes szembemenni a szakmai konszenzussal, ha új definíciónk lényegesen hozzájárul egy adott probléma megoldásához
- “**Ösztön**” ~ “öröklött viselkedésforma”: '70-es évekig még használták a viselkedésbiológiában, máig használják pszichológusok  
– nincs jó definíció: nem eléggé használható
- **DE** az **időt** sokkal előbb tudták mérni, mint definiálni!

# MI A KUTATÓMUNKA?

## 2. A kutatómunka természete: modellkészítés

► Gotelli & Ellison 2004. pp. 79-106.

### induktív módszer: egyetlen hipotézis

indukció: konkrét → általános

**jelenség** → ötletek, intuitív modellek

→ kérdés → **1 hipotézis** (modell) → jóslások (predikciók)

→ vizsgálat → értékelés: **egyeznek-e** az eredmények az elvárattal?

– nem → új modell → új kérdés

– igen → aktuális “igazság”

→ új kérdés: mit nem tudunk még a jelenségről?

*... és addig kutattak, míg meg nem haltak –  $\infty$  ciklus*

# INDUKTÍV MÓDSZER

- Egyetlen kiindulási hipotézis
- Cél a hipotézis megerősítése (verifikálása)
- Elsősorban egyetlen hipotézis több adatsoron történő ellenőrzése alkalmas

## Előnyei

- hangsúlyozza az adatok és az elmélet szoros kapcsolatát
- meglévő ismeretekre alapozza a hipotézisek készítését és változtatását

## Hátrányai

- egyetlen kiindulási hipotézist állít, másik csak ennek tarthatatlanságával jelenik meg
  - könnyen félrevisz, kedvenc hipotézis
- elméletet kizárólag tapasztalatra alapozza
  - ↔ sokszor új elméletek vezetnek új tapasztaláshoz

Maynard Smith (1982, lásd még 1989) *kő–papír–olló játék*: nincs ESS

Első empirikus eredmény: kaliforniai foltosoldalú gyíkoknál *Uta stansburiana* ♂ polimorfizmus

Sinervo & Lively 1996

### ESS: evolúciósan stabil stratégia

Olyan stratégia, amelyet ha egy populáció egyedeinek túlnyomó többsége követ, rovására semmilyen más ritka mutáns stratégia nem képes elterjedni

# A KUTATÓMUNKA TERMÉSZETE: MODELLKÉSZÍTÉS

## hipotetikus-deduktív módszer: alternatív hipotézisek

dedukció: általános → konkrét

**jelenség** → ötletek, intuitív modellek

→ kérdés → **alternatív hipotézisek** (modellek)

→ jóslások (predikciók): melyik modell egyezik leginkább a  
“valósággal”? → vizsgálat

→ értékelés: jóslások ellenőrzése, **hipotézisek cáfolata**  
(*falszifikálása*) és/vagy **”legjobb hipotézis”** megtartása

– nem létezik → új hipotézisek → új kérdés

– létezik → aktuális “igazság”

→ új kérdés: mit nem tudunk még a jelenségről?

# HIPOTETIKUS-DEDUKTÍV MÓDSZER

- Több alternatív kiindulási hipotézis
- Cél a hipotézisek megcáfolása (*falszifikálása*): elfogadjuk, ha nem tudjuk cáfolni
- Elsősorban több hipotézis egyetlen adatsoron történő ellenőrzésére alkalmas

## Előnyei

- eleve elvárja több hipotézis megfogalmazását
- kiemeli az alternatív hipotézisekből eredő eltérő elvárásokat a vizsgált rendszer működésével szemben
- tapasztalati ismeretek nélkül is használható

## Hátrányai

- nem mindig lehet megfogalmazni alternatív hipotéziseket
- akkor működik jól, ha az “igaz” hipotézis is a cáfolandók között szerepel

*“Les théories passent, la grenouille reste” - “Az elméletek jönnek és mennek – a béka marad” Jean Rostand 1959. Carnets d'un biologiste*

## **INDUKTÍV & HIPOTETIKUS-DEDUKTÍV MÓDSZEREK**

- **A kutatómunka során gyakran észrevétlenül használjuk mindkettőt együtt**
- **A jelenség vizsgálata hosszú távon nem megúszható**
- **Az elméleti alapok kidolgozása hosszú távon nem megúszható**  
**⇒ a kettőt együtt kell használni a kutatómunka során:**  
**érdemes csoportban dolgozni**
- **A modellezés közelítő eljárás**

**SOHA nem derül ki, mi az “IGAZSÁG” – különböző minőségű modellek versenyeznek egymással**

*“Imagine that everything is completely known and that science has nothing more to discover: what a nightmare!” Dobzhansky 1973*

## II. TERVEZÉS

### 1. Kérdésfeltevés

- **Miért kell kérdést feltenni?**

- kérdés nélkül nincs válasz: a módszerünk az, hogy a vizsgált rendszernek kérdéseket teszünk fel, és megpróbáljuk válaszra bírni

- **Honnan lesz kérdésünk?**

- saját megfigyelések, feltáró elemzések, gondolatok
- “*brainstorming*” – csevegés kollégákkal, konferenciák
- ► irodalom
- témavezető

- **Milyen legyen a kérdés?**

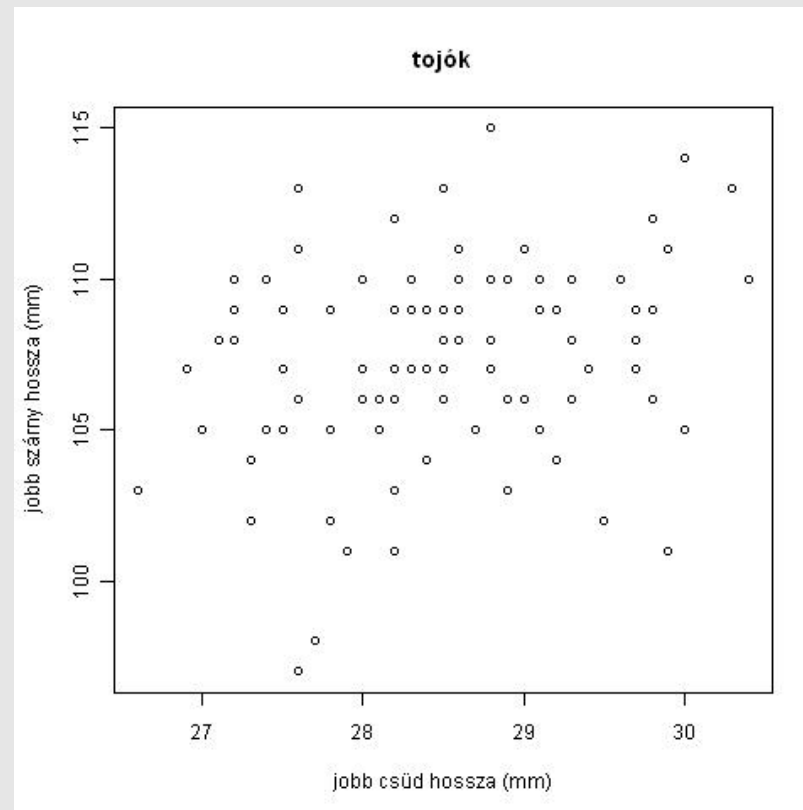
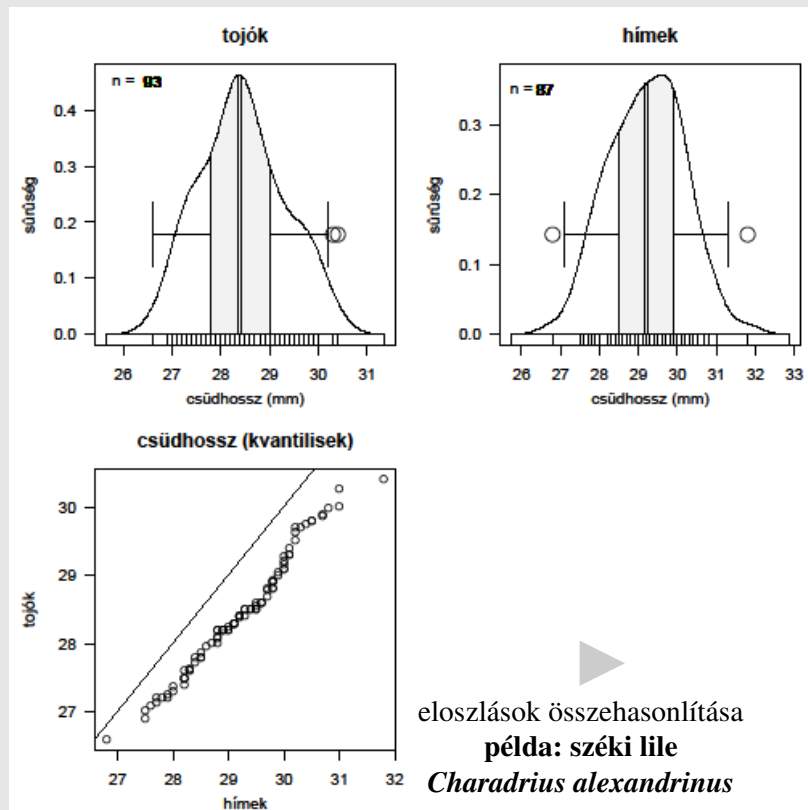
- a vizsgált rendszer szempontjából adekvát
- a lehető legpontosabb
- érdekes

# KÉRDÉSFELTEVÉS

## Feltáró (exploratív) elemzések pl. elővizsgálatokból *pilot study*

### Mire figyelünk?

- eloszlások: 1–több populáció, vagy csak nem rétegeztünk pl. ivar szerint?
- kiugró értékek: hiba az adatbevitelben, vagy extrém mérték?
- középértékek, szóródás: mennyire tér el csoportok között?
- változók közötti kapcsolat – szórásdiagrammok: milyen? allometria?



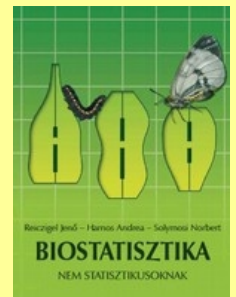
▶ R-letöltések

▶ RKWard  
(Linux)

▶ Tinn-R  
(MS-Windows)

▶ R/Tinn-R segéd

▶ R-gyakorlat  
▶ tankönyv



## **2. MIÉRT ÉRDEKES A KÉRDÉSÜNK?**

- **Ennek megválaszolása elengedhetetlen, ha**
  - **új kutatásba kezdünk**
  - **meg szeretnénk győzni másokat, hogy vegyenek részt a munkában**
  - **közölni szeretnénk eredményeinket**
  - **pénzt akarunk szerezni kutatásunkra**
  - **indokolni szeretnénk, hogy mire költöttük pl. az adófizetők pénzét**
- **El kell dönteni, alap- vagy alkalmazott kutatást végzünk-e, és annak szempontjai szerint kezdjük az indoklást – ha a másik szempont is érdekes, fontos azt is megemlíteni**
- **Mielőtt belevágunk egy új kutatómunkába, nézzük meg a szakirodalomban, hogy kérdésünk mennyire új, hol tartanak mások**
  - **Attól, hogy valami teljesen új kérdést vizsgálunk, még nem biztos, hogy érdekes, vagy tudományos szempontból jelentős ( ► Ig Nobel-díjat viszont még nyerhetünk)**
- **Mi az a tudományos kontextus, amibe munkánk illeszkedik?**

### 3. A TUDOMÁNYOS HIPOTÉZIS

- ellenőrizhető: elvileg cáfolható vagy bizonyítható
- **oksági** viszonyt ír le

⇒ tudományos hipotézis  $\neq$  statisztikai hipotézis

Miért gondozzák a ♂ ► széki lilék egyedül a fiókáikat gyakrabban / ♀♀?

► Székely *et al.* 1999  
Székely & Cuthill 1999

H<sub>1</sub>: Széki lile ♂♂ költsége a család elhagyásából > ♀♀, **MERT** több felnőtt ♂ van a populációban / ♀, így a ♂♂-nek tovább tart új párt találniuk / ♀♀-nek

h<sub>0</sub>: A széki liléknél az újrapárosodási idő átlaga a ♂♂-nél = a ♀♀-nál

$$\mu_{t(\text{ÚP-}\sigma)} = \mu_{t(\text{ÚP-}\varphi)}$$

h<sub>1</sub>: A széki liléknél az újrapárosodási idő átlaga a ♂♂-nél > a ♀♀-nál

$$\mu_{t(\text{ÚP-}\sigma)} > \mu_{t(\text{ÚP-}\varphi)}$$

H<sub>2</sub>: Széki lile ♂♂ nyeresége a fiókgondozásból > ♀♀, **MERT** a ♂♂ hatékonyabb szülők, azaz több fiókát képesek egyedül felnevelni / ♀♀

Student-féle *t*-próbára  
 $\mu$ : becsült populációátlag

#### Jelölések

(nem konvenció!)

**H**: tudományos hipotézis

**h**: statisztikai hipotézis

/: valamihez képest

# A TUDOMÁNYOS HIPOTÉZIS ÉS A JÓSLÁSOK

A jóslás egy elvárás: mi várható a vizsgált rendszertől, ha a hipotézis igaz  
→ a jóslások megfogalmazásával válik ellenőrizhetővé egy hipotézis

Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?

$H_1$ : Széki lile ♂♂ költsége a család elhagyásából > ♀♀, **MERT** több felnőtt ♂ van a populációban / ♀, így a ♂♂-nek tovább tart új párt találniuk / ♀♀-nek

$P_1$ : A felnőtt ivararány a ♂♂ felé eltol

$h_0$ : ivararány = 0.5

$$p_{\text{ivararány}} = p_0 = 0.5$$

$h_1$ : ivararány > 0.5

$$p_{\text{ivararány}} > p_0, p_0 = 0.5$$

Binomiális próbára  
 $p$ : populációban mért arány

$P_2$ : A ♂♂-nek tovább tart új párt találni / a ♀♀-nek

$h_0$ : az újrapárosodási idő átlaga

♂♂-nél és ♀♀-nél egyenlő

$$\mu_{t(\text{ÚP-}\sigma)} = \mu_{t(\text{ÚP-}\varphi)}$$

$h_1$ : az újrapárosodási idő átlaga

♂♂-nél > ♀♀-nél

$$\mu_{t(\text{ÚP-}\sigma)} > \mu_{t(\text{ÚP-}\varphi)}$$

$$\text{ivararány} = \frac{n_{\sigma}}{n_{\sigma} + n_{\varphi}}$$

## Jelölések

(nem konvenció!)

**H**: tudományos hipotézis

**P**: tudományos hipotézishez tartozó predikció

**h**: statisztikai hipotézis

/: valamihez képest

# A TUDOMÁNYOS HIPOTÉZIS ÉS A JÓSLÁSOK

Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?

$H_2$ : Széki lile ♂♂ nyeresége a fiókagondozásból  $>$  ♀♀, **MERT** a ♂♂ hatékonyabb szülők, azaz több fiókát képesek egyedül felnevelni / ♀♀

Ugyanehhez a predikcióhoz még több változót rendelhetünk, pl. *fiókatúlélés [nap]*, *fiókák növekedése (csüd hossz [mm], tömeg [g])*

$P_1$ : Az egyedül a ♂-gondozta széki lile családoknál nagyobb a szaporodási siker / ♀ gondozta családoknál

$h_0$ : A kirepült fiókák száma a ♂-gondozott családokban átlagosan megegyezik a ♀-gondozottakéval

$$\mu_{n(\text{fióka}/\text{♂-gondozott})} = \mu_{n(\text{fióka}/\text{♀-gondozott})}$$

$h_1$ : A kirepült fiókák száma a ♂-gondozott családokban átlagosan  $>$  ♀-gondozottakénál

$$\mu_{n(\text{fióka}/\text{♂-gondozott})} > \mu_{n(\text{fióka}/\text{♀-gondozott})}$$

$\Rightarrow$  a jóslások megfogalmazásával kitalálható, milyen változók mérésére van szükség

# KÉRDÉS - HIPOTÉZIS - JÓSLÁS

- A pontosan megfogalmazott kérdésből könnyen állíthatunk hipotéziseket, ezekből következnek jóslásaink, a jóslások pedig utalnak arra, milyen változókat kell mérnünk hipotéziseink teszteléséhez – végső soron kérdésünk megválaszolásához

**kérdés → hipotézisek → jóslások → változók**

- Gyakran a jóslások megfogalmazásának nehézségéből derül ki, hogy pontatlanul fogalmaztuk meg hipotéziseinket, és/vagy a kérdésünk nem világos
- Számos esetben végzünk olyan tudományos munkát, ami nem hipotézis tesztelés, pl. kutatómunka kezdetén egy ismeretlen rendszer leírásakor – elmélyülve a munkában azonban hamar eljuthatunk a hipotézisteszteszteléshez

## 4. VÁLTOZÓK

*„Szép Helénáért ezer hajó szállt vízre (másodikos anyag, John M. Prausnitz). Legyen akkor 1 millihelén azon nő szépségének foka, akiért egy darab hajót bocsátanánk vízre. Ez természetét tekintve vektor. A negatív millihelén akkor úgy értelmezendő, hogy egy hajót el kell süllyeszteni, olyan szép.” ► Esterházy 2000*

# VÁLTOZÓK TÍPUSAI

- Először a **válasz változóinkat** kell pontosan meghatároznunk, amelyek leírják az általunk vizsgált rendszert, aminek a változatosságát magyarázni szeretnénk. Statisztikai modelljeinkben ezek lesznek a **függő változók**
- Meg kell határoznunk a lehetséges **magyarázó változókat** is, ezeknek a válasz változókra gyakorolt hatásaira vagyunk kíváncsiak:
  - van-e kapcsolat a válasz és a magyarázó változók között?
  - milyen mértékben magyarázzák a magyarázó változók a válasz változók varianciáját?
  - Statisztikai modelljeinkben a magyarázó változók a **független változók**
- A **külső változók** kapcsolatban állnak a válasz változóinkkal de a erre a hatásukra nem vagyunk kíváncsiak.
  - ha tudjuk hatásukat kontrollálni, **kontroll változók**
  - ha nem tudjuk hatásukat kontrollálni, **zavaró változók**

# VÁLTOZÓK

## Mérési skálák

- **Nominális skála:** megnevezés, csoportba sorolás pl. ivar, alak, területegység
- **Ordinális skála:** csoportba sorol, de számít a sorrend pl. madarak zsírindexe:  
protokoll alapján becsüljük 0–5 (nincs zsír – nagyon sok van) között,  $5 > 0$ , de 5 & 0 között a távolság nem értelmezhető
  - átlagszámítás ezért ÉRTELMETLEN, gyakoriságokat számolhatunk
- **Intervallumskála:** nincs abszolút 0-pont, összeadás, kivonás értelmes, szorzás, osztás, arányítás NEM. pl. Celsius-skála, fészkelési időszak: nincs értelme annak a kijelentésnek, hogy X madár kétszer később kezdett el tojást rakni Y-nál, a 0-pont meghatározása önkényes
- **Arányskála:** van abszolút 0-pont, szorzás és osztás értelmes. pl. Kelvin-skála, hosszúság: van értelme annak a kijelentésnek, hogy X madár begytolla kétszer hosszabb Y-énál, a 0-pont jól definiált

# VÁLTOZÓK FELÍRÁSA

- Statisztikai modell megadása:

\* **interakció:** magyarázók kapcsolt hatása válaszra

**Y:** Függő változó ~ **X:** Független változó-1 + Fnv-2 +...Fnv-1 \* Fnv-2 +...

tengelyeken ábrázolandó

Befolyásolja-e a ♂ bóbicek *Vanellus vanellus* fészekaljvédelmét a költési időszak?

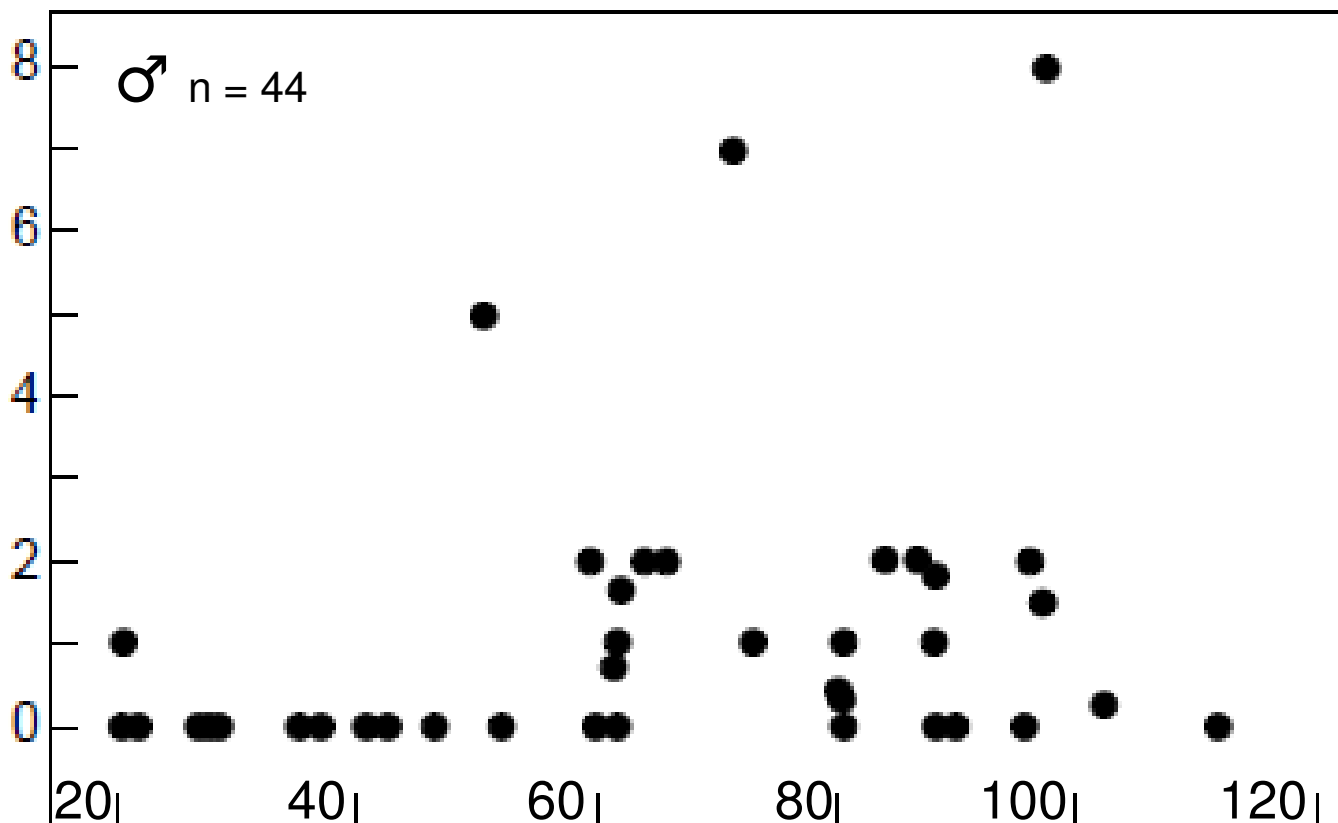
► Kis et al. 2000

♂ Támadások gyakorisága (TGY) ~ Fészkelési időszak (FID)

+ Fészeksűrűség (FS): zavaró változó

→ statisztikai kontroll, ha lehetséges

Támadások gyakorisága (db támadás / óra)



Fészkelési időszak: Március 1-től eltelt idő (nap)

**Spearman rang-korrelációk**

**TGY ~ FID**

$r_s = 0.401$ ,  $n = 44$ ,  $P < 0.01$

**DE: FS ~ FID**

$r_s = -0.654$ ,  $n = 34$ ,  $P < 0.001$

**TGY ~ FS**

$r_s = -0.683$ ,  $n = 32$ ,  $P < 0.001$  !

?

**KONTROLL:**  
parciális korreláció

**TGY ~ FID | FS**

$r_s = -0.058$ ,  $n = 29$ , n.s.

**TGY ~ FS | FID**

$r_s = -0.597$ ,  $n = 29$ ,  $P < 0.001$

⇒ **TGY ~ FS !**

# AZ ELTEL IDŐ, MINT ZAVARÓ VÁLTOZÓ

- Az eltelt idő nagyon sok esetben fellép zavaró változóként
  - két tetszőleges időben trendszerűen változó, növekedő vagy csökkenő mennyiség között látszólagos korreláció jön létre, ha az eltelt időre nem kontrollálunk
  - a lineáris regresszió nem alkalmas eszköz ennek kontrollálására
    - **idősorelemzés** *time series analysis*
      - nagy tapasztalatot igényel: forduljunk statisztikushoz

pl. magyar államadósság és India lakosságának növekedése  $r \sim 0.9$

# VÁLTOZÓK MÉRÉSE

*“Szíves érdeklődésére nb. tudomására hozom, hogy háromheti működés után szaldómon két lelőtt oroszlán fungál, az alábbi részletezés szerint:*

|                                |                   |
|--------------------------------|-------------------|
| <i>1 db nőstény</i>            | <i>2 méter</i>    |
| <i>1 db hím (vagy udvarló)</i> | <i>2,5 méter</i>  |
| <i>Összesen:</i>               | <i>4,5 méter”</i> |

► *Rejtő 19??*

# VÁLTOZÓK MÉRÉSE

- **Minden mérés becslés**
  - Minden becslésnek van hibája, ami 2 komponensre bontható
    - **pontosság:** mekkora a *becslésből származó értékek véletlen szóródása* a becslés várható értéke (a képzeletben nagyon sokszor, független mérésekre megismételt becslések átlaga) körül
      - a pontosság növelhető a mintaelemszám növelésével
    - **torzítás:** *mennyivel tér el szisztematikusan, azaz NEM a véletlen következtében*, a becsült középérték a statisztikai populáció valódi középértékétől
      - a torzítás speciális statisztikai módszerekkel csökkenthető, pl. *jackknife*-fal
- ⇒ jó a becslés, ha nagy a pontossága & kicsi a torzítása
- hogy mi számít nagynak vagy kicsinek, az a problémánktól függ

A becslés

a-c: rossz *not accurate*

d: jó *accurate*

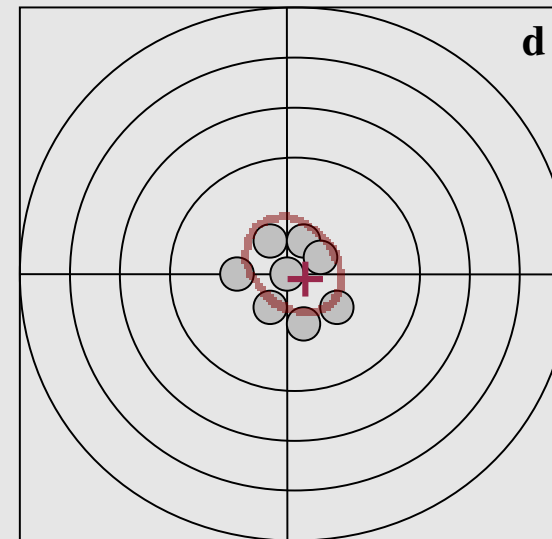
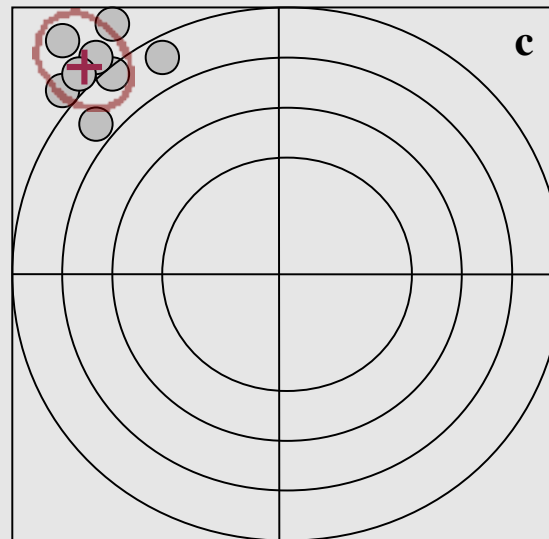
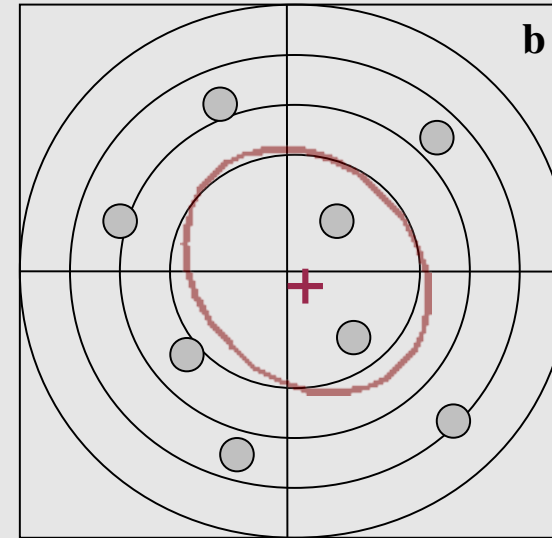
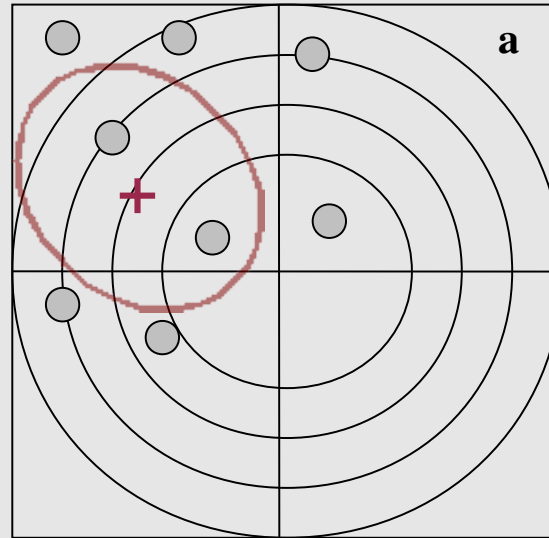
+ becsült átlag

○ szórás

TORZÍTÁS *bias* ( $\leftrightarrow$  “MEGBÍZHATÓSÁG”)

nagy kicsi

kicsi nagy



nagy

MINTA SZÓRÁSA  
véletlen hibából *random error*

kicsi

PONTOSSÁG / PRECIZITÁS  
*precision*

kicsi

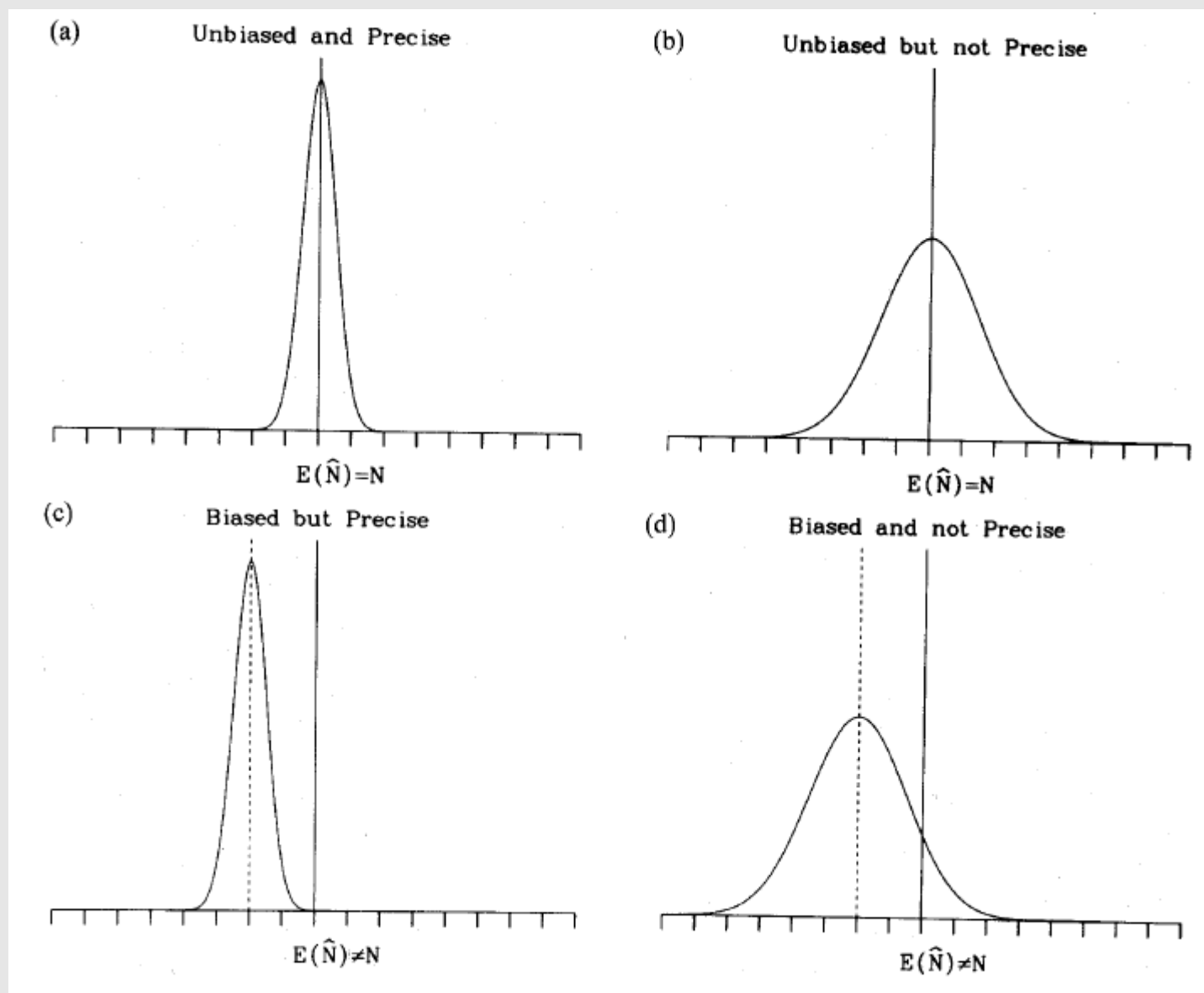
nagy

A **b** ábrán látható becslés  
jósága növelhető a  
mintaelemszám növelésével,  
a szórás csökkentésén  
keresztül

nagy

kicsi

MINTA- & POPULÁCIÓÁTLAG ELTÉRÉSE szisztematikus hibából *systematic error*



# MÉRÉSEK / BECSLÉSEK HIBÁI

► Arnqvist & Mårtensson 1998

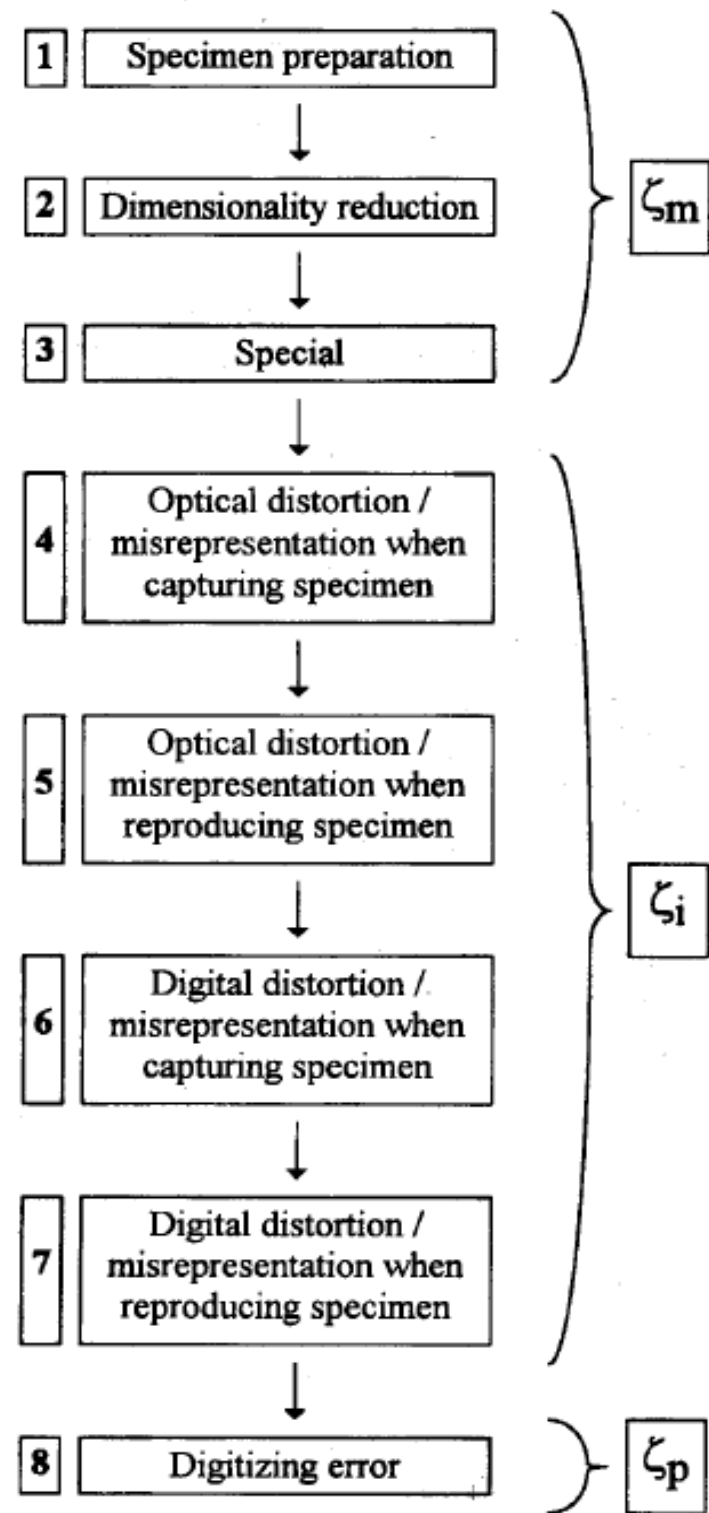
$\zeta$  := zéta

$$\zeta = \zeta_m + \zeta_i + \zeta_p$$

$\zeta_m$  := módszer hibája (method)

$\zeta_i$  := eszköz hibája (instrument)

$\zeta_p$  := mérő személy hibája (person)



# MÉRÉSEK ISMÉTELHETŐSÉGE (“megbízhatósága”) $r$

repeatability, reliability Fleiss *et al.* 2003, ► Lessells & Boag 1987

Osztályon belüli (intraklassz) korrelációs együttható intraclass correlation coefficient

Egyutas ANOVA varianciakomponenseiből adódik:

$$r = \frac{S_A^2}{S^2 + S_A^2}$$

$S_A^2 :=$  csoportok közötti

variancia komponens

$S^2 :=$  csoporton belüli

$$S^2 = MS_W$$

$$S_A^2 = \frac{MS_A - MS_W}{n_0}$$

$MS_A :=$  csoportok közötti

négyzetátlag

$MS_W :=$  csoporton belüli

$n_0 :=$  csoportonkénti mintaelemszámtól ( $n_i$ ) függő  
együttható

ha  $n_i = n_j$ ,  $n_0 = n_{\text{átlag}}$

ha  $n_i \neq n_j$ ,  $n_0 < n_{\text{átlag}}$

## a mérés ismételhetősége

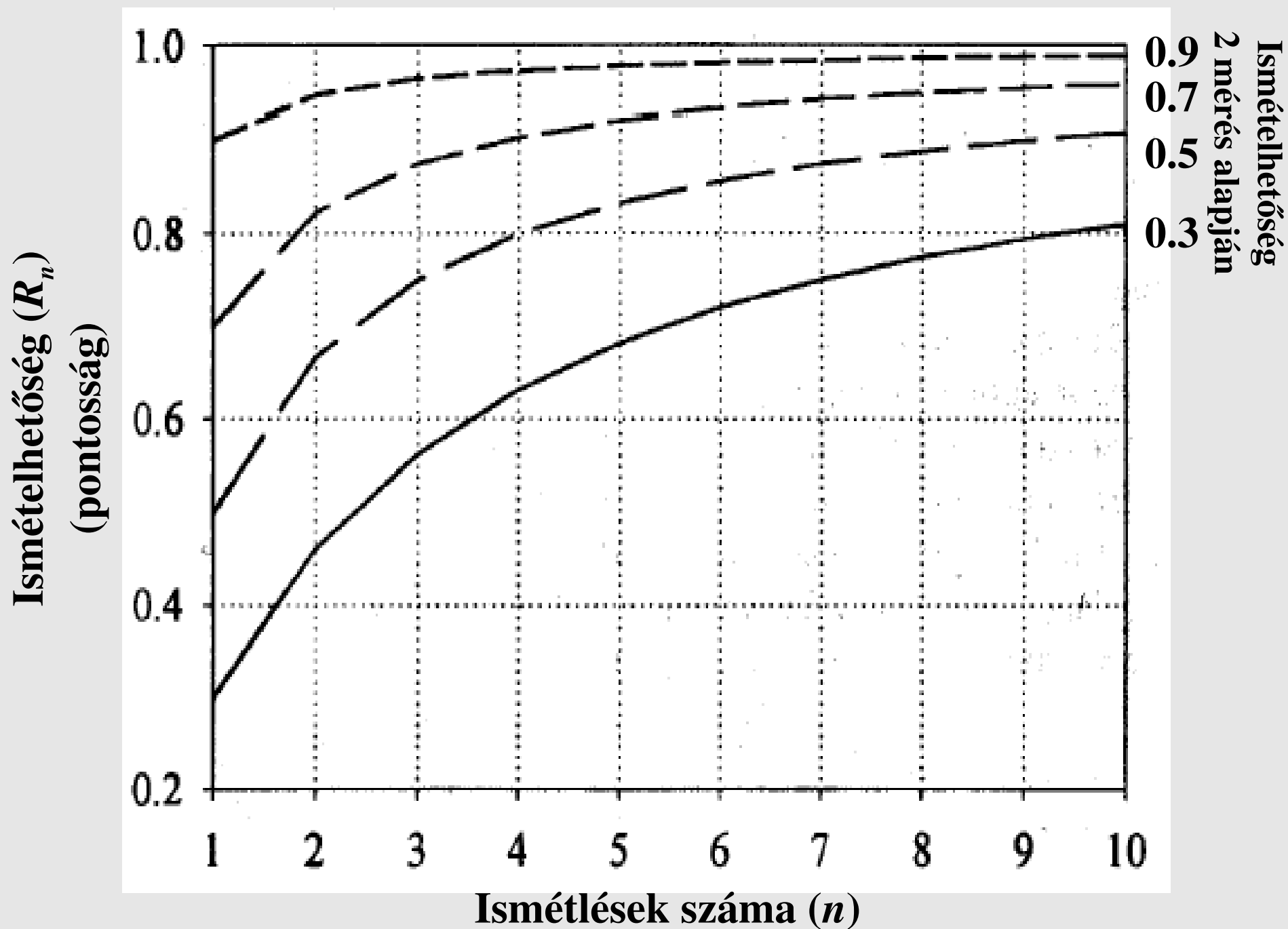
kiváló:  $r_{CIL} > 0.75$

elfogadható:  $0.4 \leq r_{CIL} \leq 0.75$

gyenge:  $r_{CIL} < 0.4$

Ahol  $r_{CIL}$  az ismételhetőségre  
adott egyoldali **alsó** 95 %-os  
konfidencia-határ

**A példa folytonos mérési skálán mért változókra vonatkozik!**



# ISMÉTELHETŐSÉG MÉRÉSE R-REL

► [irr package & irr.pdf](#) ► [R-letöltések](#)



icc (intraclass correlation coefficient) paranccsal

## Lehetséges beállítások

```
icc(ratings, model = c("oneway", "twoway"),  
type = c("consistency", "agreement"),  
unit = c("single", "average"), r0 = 0, conf.level = 0.95)
```

**Adatok:** kis Apolló-lepke  
*Parnassius mnemosyne*  
erényöv hossza (mm)  
2 személy mérése tolómérővel

|   | 1. mérő | 2. mérő |
|---|---------|---------|
| 1 | 9.70    | 9.75    |
| 2 | 6.30    | 6.40    |
| 3 | 10.15   | 10.25   |
| 4 | 10.10   | 9.90    |
| 5 | 8.80    | 8.80    |

## R-parancs:

```
icc(sphragis.repeatability, conf.level = 0.9)
```

## R-output:

Single Score Intraclass Correlation

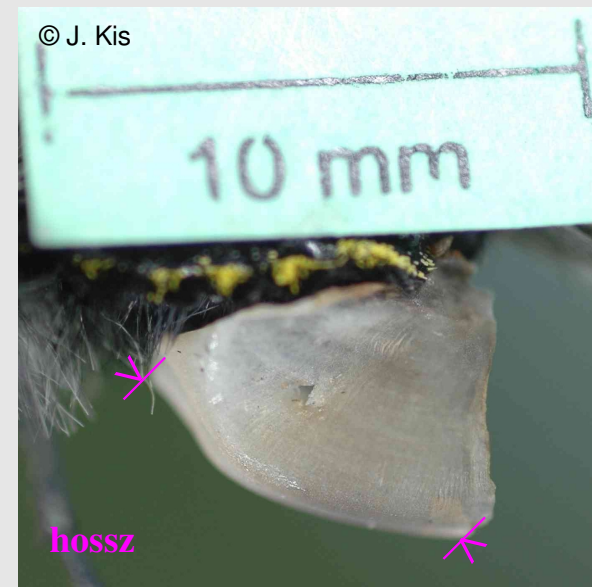
Model: oneway  
Type : consistency

Subjects = 5  
Raters = 2  
ICC(1) = 0.998

F-Test,  $H_0: r_0 = 0$   
 $F(4,5) = 802$ ,  $p = 6.68e-07$

!!! 90%-Confidence Interval for ICC Population Values:  
 $0.987 < ICC < 1$

**Nem folytonos skálán mért változókra más módszerek alkalmazandóak – lásd irr.pdf!**



► gyakorlat

# ISMÉTELHETŐSÉG MÉRÉSE R-REL

## Mi fontos az R-outputban?

- 95 %-os alsó konfidencia-határ
- elfogadható ismételhetőség:  
az alsó konfidencia-határ  $> 0.4$

## Mi irreleváns az R-outputban?

- felső konfidencia-határ
- $p$ -érték

## Mire jó az ICC?

- 2 v. több mérő becslésének az összehasonlítására

## Mire NEM jó az ICC?

- olyan csoportok összehasonlítására, amelyek varianciája egymástól eltér
- pl. műszerek megbízhatóságának tesztelésére

## R-parancs:

`icc(sphragis.repeatability, conf.level = 0.9)`

## R-output:

Single Score Intraclass Correlation

Model: oneway

Type : consistency

Subjects = 5

Raters = 2

ICC(1) = 0.998

F-Test, H0:  $r_0 = 0$

$F(4,5) = 802$  ,  $p = 6.68e-07$

!!! 90%-Confidence Interval for ICC Population Values:  
 $0.987 < ICC < 1$

### Miért 90 %-os KI-mal számolunk?

Az ICC() kétoldali szimmetrikus KI-at számol, minket viszont csak az alsó határ érdekel, mivel arra vagyunk kíváncsiak, *nagyobb-e  $r$  egy kritikus értéknél*, pl. 0.4-nél vagy 0.75-nél. Ezért az alsó határt szeretnénk 5 %-os hibával ismerni, a felsőre nincs szükségünk, egyoldali KI-ot számolnánk. Ezt az ICC-nek úgy tudjuk megadni, hogy 90 %-os határokat számolunk, aminek az alsó határa megfelel a 95 %-os alsó határnak egyoldali KI esetén, a felső határa 1. Így végül **a 95 %-os alsó határt kapjuk meg, noha 90 %-ot kértünk és az output is annyit ír.**

# STATISZTIKAI HIPOTÉZISTESZTELÉSBŐL FAKADÓ HIBÁK

- További hibák származnak a statisztikai elemzésekből: amikor statisztikai hipotézisek elfogadásáról döntünk, valamekkora valószínűséggel hozhatunk hibás döntést.
- A  $p$ -érték annak a valószínűsége, hogy a véletlen  $h_0$ -nak legalább annyira ellentmondó mintát eredményez, mint amit megfigyeltünk, noha  $h_0$  igaz: azt mutatja, hogy a megfigyelt minta mennyire erős bizonyíték  $h_1$  mellett,  $h_0$  ellenében  
 $\Rightarrow$  ha a  $p$ -érték az elfogadási tartományba esik, elvetjük  $h_0$ -t, azonban ha ezen kívül esik, ez nem elég erős bizonyíték  $h_0$  mellett!
- **1-fajta** (1-fajú) **hiba** type-1 error: elutasítjuk  $h_0$ -t, pedig igaz
- **2-fajta** (2-fajú) **hiba** type-2 error: elfogadjuk  $h_0$ -t, pedig hamis
- Az 1-fajta hiba minimalizálásával a 2-fajta hiba növekedni fog: célunk egy optimális szint, amelyen a 2 hiba együttese minimális
- A **teszt ereje** power annak a valószínűsége, hogy  $h_0$ -t elvetjük, ha  $h_1$  igaz, azaz jó döntést hozunk

## 5. MINTAVÉTEL

- A legtöbb esetben vizsgálataink során nem egy teljes **statisztikai populáción** végezzük méréseinket, hanem annak egy részét mintázzuk
- A statisztikai populáció  $\neq$  genetikai vagy ökológiai populációval, noha esetenként egybe eshet valamelyikkel
  - ebben az anyagban populáción a statisztikai populációt értem
- A mintavétel során arra törekszünk, hogy **mintánk**, a **vizsgált populáció** azon része, amelyen a méréseket elvégezzük,
  - minél *pontosabban* és minél *kisebb torzítással* írja le a vizsgált populációt,
  - legyen **reprezentatív** a mért változókra nézve: e változók eloszlása a mintában legyen azonos a statisztikai populációban észlelttel
    - gyakorlati közelítés: minden egyed azonos valószínűséggel kerüljön be a mintába
- A vizsgált populáció kiválasztásánál arra törekszünk, hogy jól jellemezze a **célpopulációt**, amire a vizsgálat irányul

# MINTAVÉTEL

**Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?**

- Erre a kérdésre keresve a választ célpopulációunkat a világon élő széki lilék képezik.  
Általános választ szeretnénk kapni kérdésünkre a széki lile fajra vonatkozóan.
- Vizsgált populációnk a törökországi Tuzla-tó becsléseink szerint 1000 költő párból álló széki lile állománya.
- Mintánk az a néhány száz széki lile, amelyről adataink vannak.

# MINTAVÉTEL

## vizsgálat érvényessége

- Azt, hogy vizsgálatunk mennyire alkalmas az általunk levont következtetések alátámasztására, a vizsgálat **érvényessége (validitása)** jellemzi.
- A vizsgálat **belső érvényessége** mutatja, valóban az a hatás okozza-e válaszváltozónk változatosságát, amelyről ezt feltételezzük. Ez gyakran csak kísérletes körülmények között biztosítható megnyugtatóan kontrollcsoport, alapszint baseline használatával, de ezt szolgálja az alternatív hipotézisek versenyeztetése is.
- A **konstrukció érvényesség** mutatja, helyes-e az eredményekből levont következtetésünk.
  - tágabb értelemben: helyes-e a teljes vizsgálat koncepciójának érvelése
- A **statisztikai érvényesség** statisztikai módszereink alkalmazhatóságát mutatja.
- A vizsgálat **külső érvényessége** azt jelenti, hogy mennyire általánosíthatók következtetéseink a nem vizsgált populációkra

# BELSŐ ÉRVÉNYESSÉG

**Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?**

- Két alternatív hipotézis vizsgálata két kísérletben (**K**)
  - **H1**: A ♂♂ többen vannak, ezért több idő új párt találniuk /♀♀
  - **H2**: A ♂♂ jobb szülők, több fiókájuk él túl, ha egyedül gondoznak /♀♀
  - **K(H1)**: pár-& fészekalj-eltávolított ♂♂ & ♀♀, újrapárosodási idő mérése; ivarok egymás kontrolljai → ♂ *újrapárosodási idő* > ♀,  $n♂ > n♀$
  - **K(H2)**: vagy ♂ vagy ♀ szülő eltávolítása családtól, fiókák kirepülése, túlélése, növekedése; különböző ivarú szülők vezette családok egymás kontrolljai → ♂ & ♀ *gondozta családok nem különböznek*
- ⇒ **H1** igaz, **H2** nem
- belső érvényesség rendben (meggyőző kísérleti elrendezés, alternatív hipotézisek; az eredmények sokkal meggyőzőbbek, mint a bóbicek példában, amelyben csak statisztikailag tudtunk egy zavaróváltozóra kontrollálni, mire kiderült, hogy valószínűbb a hatása az eredetileg számba vett magyarázóváltozónál, nem kizárható más, nem mért változók lényeges hatása sem, amelyek kapcsolatban állnak a ható változónak tekintettel)

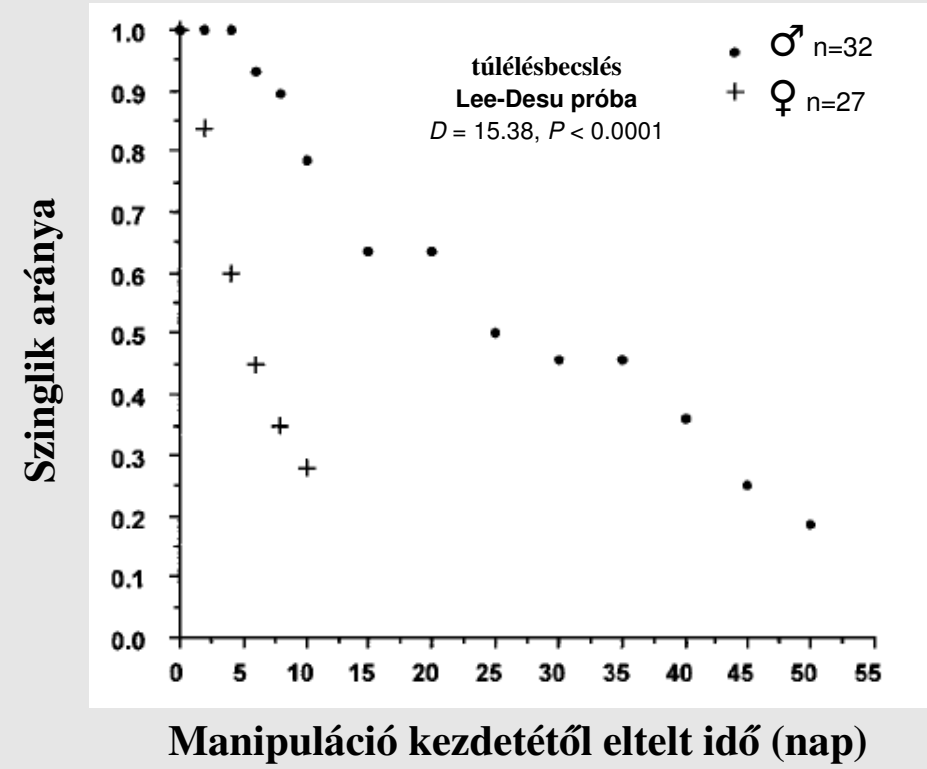
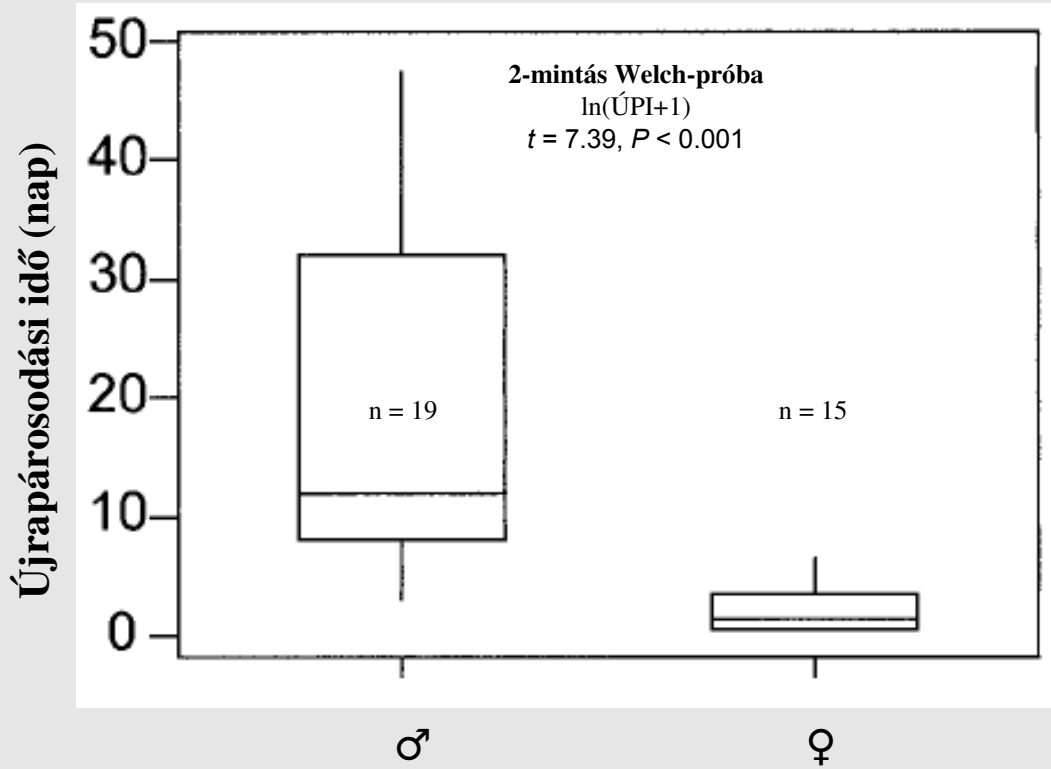
# KONSTRUKCIÓ-ÉRVÉNYESSÉG

Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?

- Két kísérlet (K) eredményei
  - $K(H_1)$ : ♂ *újrappárosodási idő* > ♀,  $n\sigma > n\phi$
  - $K(H_2)$ : ♂ & ♀ *gondozta családok nem különböznek*
  - $\Rightarrow H_1$  igaz,  $H_2$  hamis
- Következik-e az eredményekből, hogy
  - igaz  $H_1$ : A ♂♂ többen vannak, ezért több idő új párt találniuk /♀♀
  - hamis  $H_2$ : A ♂♂ jobb szülők, több fiókájuk él túl, ha egyedül gondoznak /♀♀ ?

# STATISZTIKAI ÉRVÉNYESSÉG

Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀? ► Székely et al. 1999



- A Welch-próba a  $t$ -próba általános formája, ami nem feltételezi a szórások homogenitását
- A Welch próba csak azokat a mintaelemeket képes kezelni, amelyeknél az újrapárosodási idő ismert. Ez torzítja becslésünket, mert a minta csak a sikeresebb madarakat tartalmazza.
- A túlélésbecslés **cenzorált adatként** kezeli azon mintaelemeket, amelyekről annyit tudunk, mi az a minimális idő, ameddig biztosan nem kaptak új párt. Ezek figyelembevételével becsli a medián újrapárosodási időt a 2 ivarra, így a becslés torzítása << Welch-próbáé

|                | ♂    | ♀   |   | ♂  | ♀  |
|----------------|------|-----|---|----|----|
| medián ÚPI     | 12.0 | 1.5 | n | 19 | 15 |
| TBK medián ÚPI | 25.4 | 5.3 | n | 32 | 27 |

TBK: túlélésbecsléssel korrigált

# KÜLSŐ ÉRVÉNYESSÉG

**Miért gondozzák a ♂ széki lilék egyedül a fiókáikat gyakrabban / ♀♀?** ► Kis 2003

- Vizsgált populációnk a törökországi Tuzla-tó becsléseink szerint 1000 költő párból álló széki lile állománya.
- Célpopulációkat a világon élő széki lilék képezik.

⇒ Mennyire általánosíthatóak a tuzla-tavi állományon végzett vizsgálatok eredményei a széki lile fajra?

- Több felnőtt ♂ / ♀ Tuzlán és Kaliforniában
- Tuzlán kétszülős és egyszülős ♂-gondozás fiókás családoknál, egyszülős ♀-gondozás ritka vagy nem fordul elő  
↔ a Kiskunságban, Camargue-on és Kaliforniában nem ritka az egyszülős ♀-gondozás sem

⇒ Ahhoz, hogy a széki lilékre mint fajra általánosítsuk következtetéseinket, más populációkat is vizsgálni kell!

# MINTAVÉTEL

## Mire figyelünk?

- A mintavétel tervezése során célunk, hogy
  - a vizsgált populáció jól reprezentálja a célpopulációt kérdésünk szempontjából
    - irodalom, kollégák, elővizsgálatok
  - mintánk legyen jól definiált és legyen reprezentatív a vizsgált populációra
    - szisztematikus hibák elkerülése: **minden egyed ugyanakkora valószínűséggel kerüljön be a mintánkba:**
- **véletlen mintavétel (randomizáció):** véletlen számok segítségével döntjük el, mely egyedek kerüljenek a mintába, melyek ne – pl. “szép” egyedek nem tudatos beválogatása elkerülhető: pl. múzeumi gyűjteményekben a lepkék szárnya többnyire ép, míg a szabadban sok töredezett szárnyút fogunk
- **rétegzett mintavétel:** mintánkat egy faktor, pl. ivar, fajta, területi egység szerint részekre bontjuk, és rétegenként választjuk be az egyedeket a mintánkba pl. véletlen választással, hogy rétegenként egyenlő arányban legyenek egyedeink – ha elővizsgálatok során többszűs eloszlásokat kapunk egy változóra, érdemes gondolkodnunk azon, vajon milyen faktor szerint bonthatnánk azt egyszűsűakra, mi szerint rétegezhethetnénk a mintavételt

# MINTAVÉTEL

## Mire figyelünk?

- **Szisztematikus hibák elkerülése:** minden egyed ugyanakkora valószínűséggel kerüljön be a mintánkba:
- **blokk-elrendezés:** a mintavételt blokkokba rendezzük, pl. hogy mindegyik kezelés ugyanabban az időszakban kezdődjön a blokkban

Miért gondolják a ♂ széki lilék egyedül a fiókákat gyakrabban / ♀♀?

| blokk | terület | dátum | fészekID | eltávolított szülő ivara |
|-------|---------|-------|----------|--------------------------|
| 1     | A       | 04-12 | 24       | ♀                        |
| 1     | A       | 04-13 | 31       | ♂                        |
| 1     | B       | 04-12 | 49       | ♂                        |
| 1     | B       | 04-12 | 33       | ♀                        |
| 2     | A       | 04-13 | 27       | ♂                        |
| 2     | A       | 04-14 | 36       | ♀                        |
| 2     | B       | 04-14 | 40       | ♂                        |
| 2     | B       | 04-15 | 46       | ♀                        |
| 3     | A       | 04-15 | 38       | ♀...                     |

A blokkokba területenként és időszakonként véletlen szerint válogatjuk a fészkeket, és hogy melyik fészeknél távolítjuk el pl. a ♂-et. Mivel az A24-esnél a sorsolás alapján a ♀-t távolítottuk el, az A-ban a következő fészeknél a ♂-et fogjuk, de a B-ben újra sorsolunk.

# MINTAVÉTEL

## Mire figyelünk?

- **Szisztematikus hibák elkerülése:** minden egyed ugyanakkora valószínűséggel kerüljön be a mintánkba:
  - **vak minta:** a megfigyelő elvárásai a megfigyelt rendszer működésével kapcsolatban torzíthatják a mintavételt. Ilyen esetben olyan megfigyelőt kell választani, aki nem ismeri a megfigyelés célját, csak alkalmazza a megfigyelési protokollt, amit a kutatás tervezése során készítünk

# MINTAVÉTEL

## Egy gyakori hiba

- A mintánk nem jól definiált: következtetéseink érvénytelenek

Egy erdőben elhelyezünk 100-100 kisemlőscsapdát a tölgyesbe és a bükkösbe, és a fogások alapján kijelentjük, hogy az egyik pocokfaj a tölgyesekre, a másik a bükkösökre jellemző.

Mekkora a mintaelemszám?  **$n = 1$**

Miért? Egyetlen erdő egyetlen tölgyesét és bükkösét mintáztuk, amelyek különböző részei nem függetlenek egymástól. Vizsgálatunk külső érvényessége csak a mintázott erdőrészekre vonatkozik.

Hasonló ehhez, amikor WISTAR patkányokon végzett viselkedésvizsgálataink eredményeit érvényesnek tartjuk a patkányokra általában. A WISTAR egy klón, amit éppen azért hoztak létre, hogy az egyedek közötti varianciát minimalizálják. Ez jól megfelelhet gyógyszerkutatási céloknak, és nem várható, hogy általában a patkányok viselkedéséről megtudunk belőle valamit.

# 6. MEGKÖZELÍTÉSI MÓDOK

- **Megfigyelés:** a jelenségből indulunk ki, nem avatkozunk a rendszer működésébe
  - a rendszer természetes működését észleljük
  - sok zavaróváltozó hatásáról még csak tudomást sem szerzünk
- **Kísérlet:** a jelenségből indulunk ki, próbálunk minden külső változóra kontrollálni, és egy, vagy kevés magyarázó változót manipulálni, amelyek hatását a válaszváltozókra vizsgálni szeretnénk
  - zavaróváltozók nem befolyásolják eredményeinket
  - a kísérleti körülmények gyakran befolyásolják a rendszer működését, így műterméket kaphatunk
- **Metaelemzés:** publikált adatokból dolgozunk: számos megfigyelés és kísérlet eredménye értékelhető egyszerre
  - sok populációval dolgozhatunk: a külső érvényesség nagy
    - vagy skálát váltunk: faj helyett fajcsoport vizsgálata
  - a különböző forrásból származó adatok gyűjtéséből adódó módszertani különbségek torzíthatják következtetéseinket: belső érvényesség kicsi

# MEGKÖZELÍTÉSI MÓDOK

- **Evolúciós összehasonlító** (komparatív filogenetikai) **módszerek**
  - legtöbb esetben metaelemzés, azaz mások adataival dolgozunk
  - megfigyelés
  - törzsfamodellel dolgozunk általában a ma élő fajok tulajdonságainak felhasználásával
  - jellegek közötti összefüggések elemzése
  - jellegek evolúciós rekonstrukciója
- **Modellezés:** matematikai nyelven megfogalmazott, **formális modellek** készítése – sokkal pontosabb verbális vagy grafikus modelleknél
  - elméleti konstrukcióból indulunk ki
  - legyen matematikailag koherens
  - analitikus megoldása van (levezetés) vagy szimulációs modell
  - problémák végiggondolására alkalmas

# MEGKÖZELÍTÉSI MÓDOK

## Megfigyelés vagy kísérlet?

- **Megfigyelés:** változók közötti kapcsolat keresése; nincs kezelés
- **Kísérlet:** ok-okozati összefüggés keresése; kezelés  $\rightarrow$  független változó

# MEGFIGYELÉS

- **Változók közötti kapcsolat keresése**
- Nincs kezelés
- **Előnyei**
  - „természetes” rendszerek vizsgálata
  - természetes variabilitás a vizsgált változókbán
  - természetes környezet
  - egyszerű módszerek és eszközök
  - nagy mintaszám
  - etikai problémák csökkenthetők
- **Hátrányai**
  - korreláló változók
  - zavaró változók
    - standardizált mintavétel
    - megfigyelési blokkok
    - reziduumok elemzése: kontroll zavaróváltozókra
  - ritka események vizsgálata
  - megfigyelő hatása

# MEGFIGYELÉS

- **Milyen kérdéseket vizsgálhatunk megfigyelésekkel?**
  - elővizsgálatok
  - tudományos hipotézisek tesztelése: sok predikció egyidejű ellenőrzése
  - kísérletes vizsgálatok érvényességének igazolása: *Vajon a természetes állapot megfigyelése alátámasztja kísérletből származó következtetéseinket, vagy kísérletes eredményeink műtermékek?*
  - nincs lehetőség kísérletezésre

# MEGFIGYELÉS

Befolyásolja-e a ♂ bíbicek *Vanellus vanellus* fészekaljvédelmét a költési időszak?

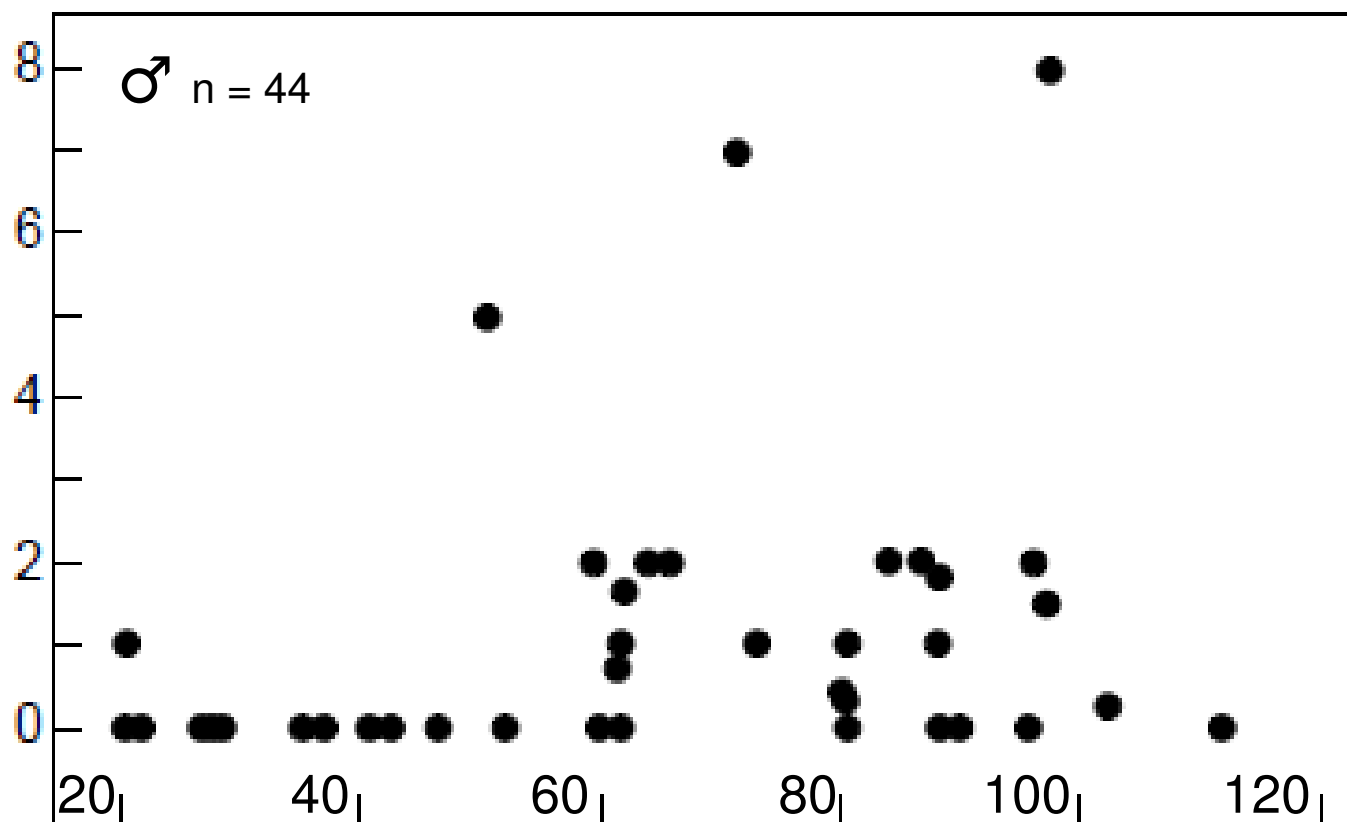
► Kis et al. 2000

♂ Támadások gyakorisága (TGY) ~ Fészkelési időszak (FID)

+ Fészeksűrűség (FS): zavaró változó

→ statisztikai kontroll, ha lehetséges

Támadások gyakorisága (db támadás / óra)



Fészkelési időszak: Március 1-től eltelt idő (nap)

**Spearman rang-korrelációk**

**TGY ~ FID**

$r_s = 0.401, n = 44, P < 0.01$

**DE: FS ~ FID**

$r_s = -0.654, n = 34, P < 0.001$

**TGY ~ FS**

$r_s = -0.683, n = 32, P < 0.001 !$

?

**KONTROLL:**

**parciális korreláció**

**TGY ~ FID | FS**

$r_s = -0.058, n = 29, \text{n.s.}$

**TGY ~ FS | FID**

$r_s = -0.597, n = 29, P < 0.001$

**⇒ TGY ~ FS !**

- A megfigyelések természetes körülmények között zajlottak, a viselkedés természetes
- Lehet olyan figyelmen kívül hagyott zavaróváltozó, amely kapcsolatban áll FS-gel, és az hat TGY-re

# KÍSÉRLET

- **Ok-okozati összefüggés keresése**
- A **válaszváltozót** a kísérlet során a kezeléssel **befolyásoljuk**
- A független változó a **kezelés**
  - a statisztikai modellekben a kezelés egy **faktor**
  - **szintjei:** kezelési csoportok száma
  - a kezelés hatását **kontrollcsoport**hoz vagy ismert **alapszinthez** baseline viszonyítjuk
- A kérdés: a kezelés hatására kapott eredmény mekkora valószínűséggel tulajdonítható a kezelésnek vagy a véletlennek
- **A kísérletek típusai:**
  - **egyszerű:** (megfigyelés) – kezelés – megfigyelés: viszonyítás alapszinthez
  - **összehasonlító:** viszonyítás kontroll csoport

# KÍSÉRLET

- **A kísérletezés alapelemei:**
  - standardizálás (kontrollálás): kezelésen kívül a csoportok lehetőleg mindenben azonosak
  - előre eltervezett kezelési csoportok (blokkok)
  - zavaró változók statisztikai kontrollja
  - randomizálás
    - alanyok csoportba sorolásához
    - kezelések sorrendjének eldöntéséhez
  - ismétlés

# KÍSÉRLET

- **Gyakori kísérleti elrendezések**

- Cél: a kísérleti alanyok kezelési csoportok közötti változatosságának csökkentése

- tökéletesen random elrendezés
- blokk elrendezés
- latin négyzet
- faktoriális elrendezés

- független csoportok vagy ismételt mérések?

- A kísérletezés korlátai

- megfigyelő elfogultsága: „vakon” végzett kísérletek
- érvényessége korlátozott: többféle kísérlet ugyanarra a kérdésre
- a probléma nem vizsgálható kísérlettel

# KÍSÉRLET

**Miért nagy a különbség a ♂ széki lilék pártalálási sikerében?** ► Lendvai *et al.* 2004

**H<sub>1</sub>:** A nagyobb begycsíkú hímek jobb pasik, ezek hamarabb találnak párt a kisebb begycsíkúaknál

**H<sub>2</sub>:** A hosszabb begytollú hímek jobb apák, ezek hamarabb találnak párt a rövidebb begytollúaknál

- mindkét változónak nagy a természetes varianciája

## Terepi kísérlet

**2 × 2 faktoriális  
elrendezés**

**Begytoll  
vágott kontroll**

|                 |          |           |           |
|-----------------|----------|-----------|-----------|
| <b>Begycsík</b> | Növelt   | <b>Nv</b> | <b>Nk</b> |
|                 | Kontroll | <b>Kv</b> | <b>Kk</b> |

### Blokkosítás

időben (szezón)  
területenként

### Randomizálás

melyik fészket (egyedet) használjuk fel  
melyik ♂ melyik kezelést kapja

## Terepmunka beosztása

**Latin négyzet elrendezés**

|                |           |                |          |          |
|----------------|-----------|----------------|----------|----------|
|                |           | <b>Terület</b> |          |          |
|                |           | <b>A</b>       | <b>B</b> | <b>C</b> |
| <b>Napszak</b> | <b>DE</b> | K              | M        | K        |
|                | <b>DU</b> | M              | K        | M        |

K: kísérlet

M: megfigyelés

# KÍSÉRLET

**Miért nagy a különbség a ♂ széki lilék pártalálási sikerében?** ► Lendvai *et al.* 2004

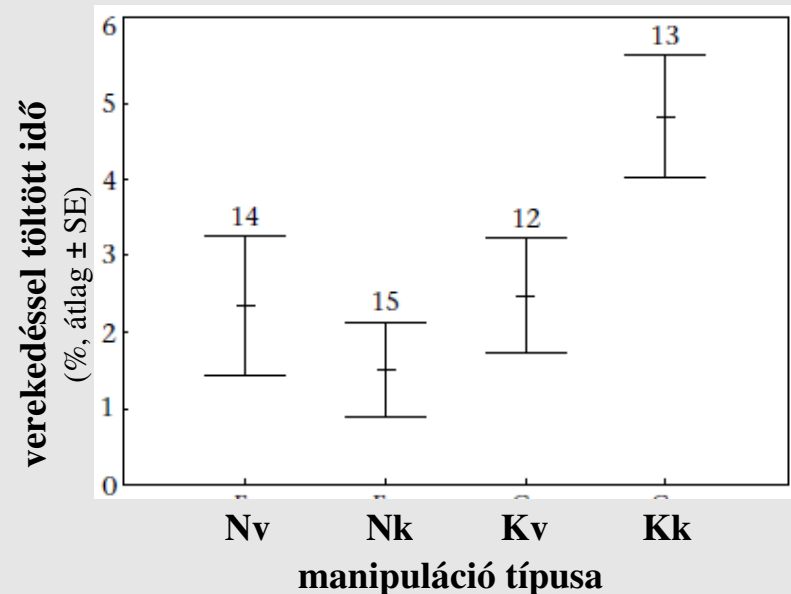
**H<sub>1</sub>:** A nagyobb begyecsíkú hímek jobb pasik, ezek hamarabb találnak párt a kisebb begyecsíkúaknál

**H<sub>2</sub>:** A hosszabb begytollú hímek jobb apák, ezek hamarabb találnak párt a rövidebb begytollúaknál

- mindkét változónak nagy a természetes varianciája

Nincs kezelések között statisztikailag szignifikáns különbség az újrapárosodási időben

⇒ sem **H<sub>1</sub>**-t, sem **H<sub>2</sub>**-t nem támasztják alá e vizsgálatok



Begyecsík-növeltek kevesebb időt töltöttek verekedéssel / begyecsík-kontroll

$$F_{1,48} = 7.418, P = 0.009 \text{ ANOVA}$$

⇒ **H:** begyecsík ♂ státuszát jelzi

## Terepi kísérlet

**2 × 2 faktoriális  
elrendezés**

**Begytoll**  
vágott      kontroll

**Begyecsík**

|          |    |    |
|----------|----|----|
| Növelt   | Nv | Nk |
| Kontroll | Kv | Kk |

# METAELEMZÉS

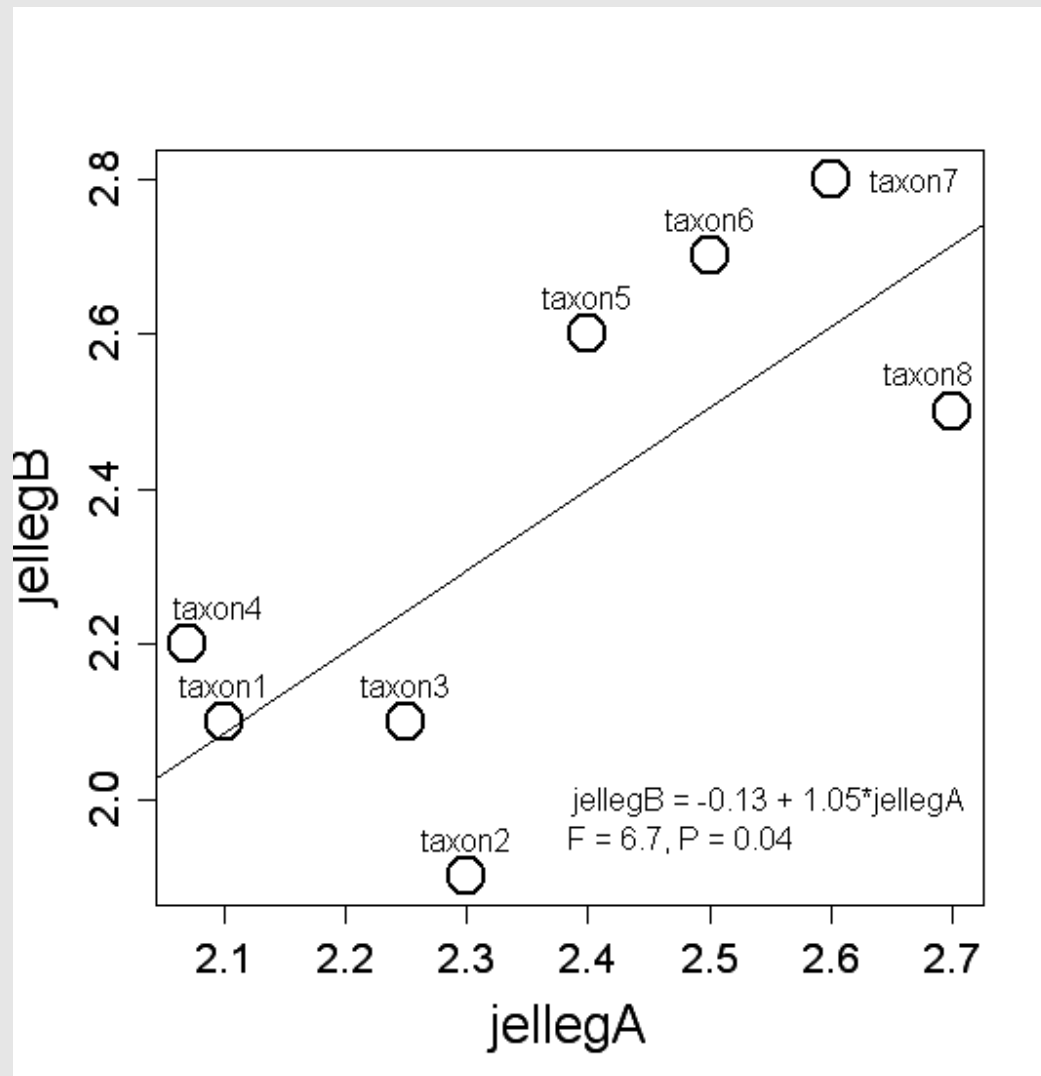
- **Megfigyelés**
- publikált adatokból dolgozunk: számos megfigyelés és kísérlet eredménye értékelhető egyszerre
- sok populációval dolgozhatunk: a külső érvényesség nagy
  - gyakran skálát váltunk: faj helyett fajcsoport vizsgálata
  - a különböző forrásból származó adatok gyűjtéséből adódó módszertani különbségek torzíthatják következtetéseinket: belső érvényesség kicsi
  - az átlagok csak homogén esetben reprezentálják az egyesített populációt  
→ homogenitásteszt; homogén: *fix hatásos metaelemzés*
  - szabályosan heterogén populációk: *random hatásos metaelemzés*  
→ gondosan kell kiválasztani a megfigyelési egységet  
→ *sok tényezőre kell kontrollálni:*
    - kutatási befektetés pl.
      - földrajzi egységenként
      - taxononként
      - jelenségenként
- egy gyakori alosete: filogenetikai vizsgálatok

DerSimonian-Laird 1986

változhat

# Evolúciós összehasonlító (komparatív) vizsgálatok

fajok összehasonlítása hagyományos módszerekkel (pl. lineáris regresszió)



Virtuállatok *Simulalia*

- hullámtermészetűek
- korpuszkuálisak

Feltevésünk szerint a nagyobb tömegű  
(*jellegA*) fajok nagyobb amplitúdóval  
(*jellegB*) hullámszanak

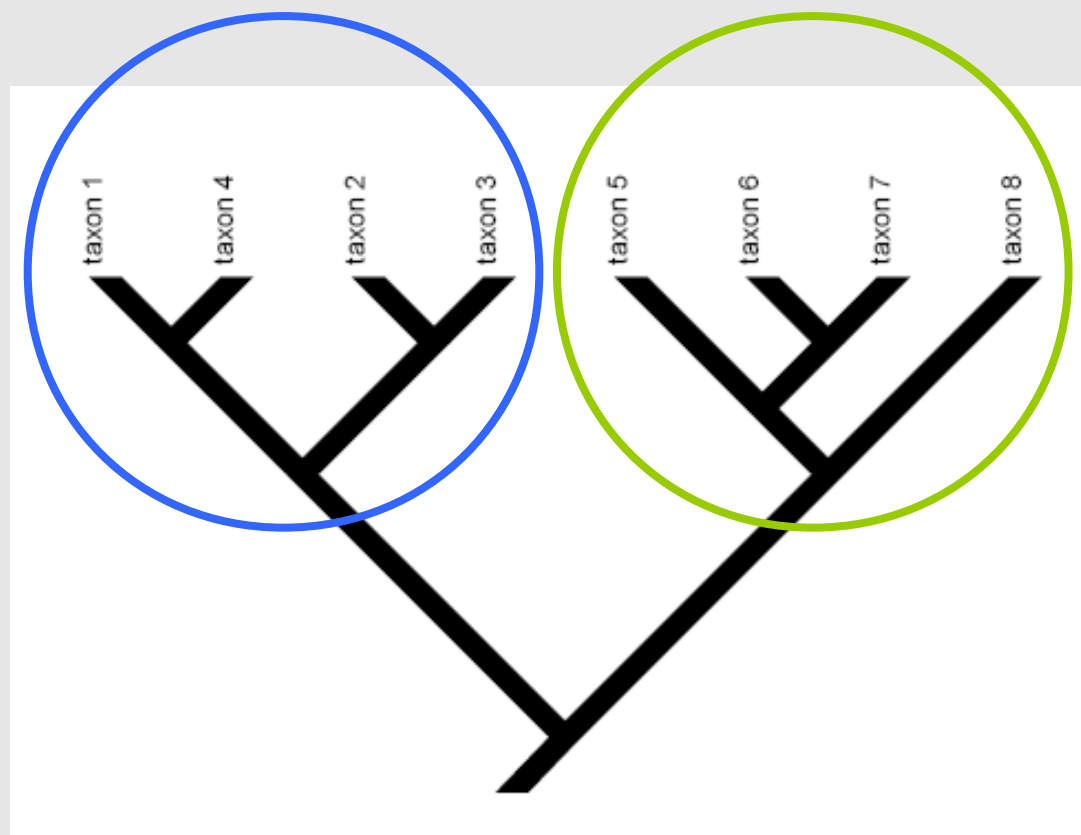
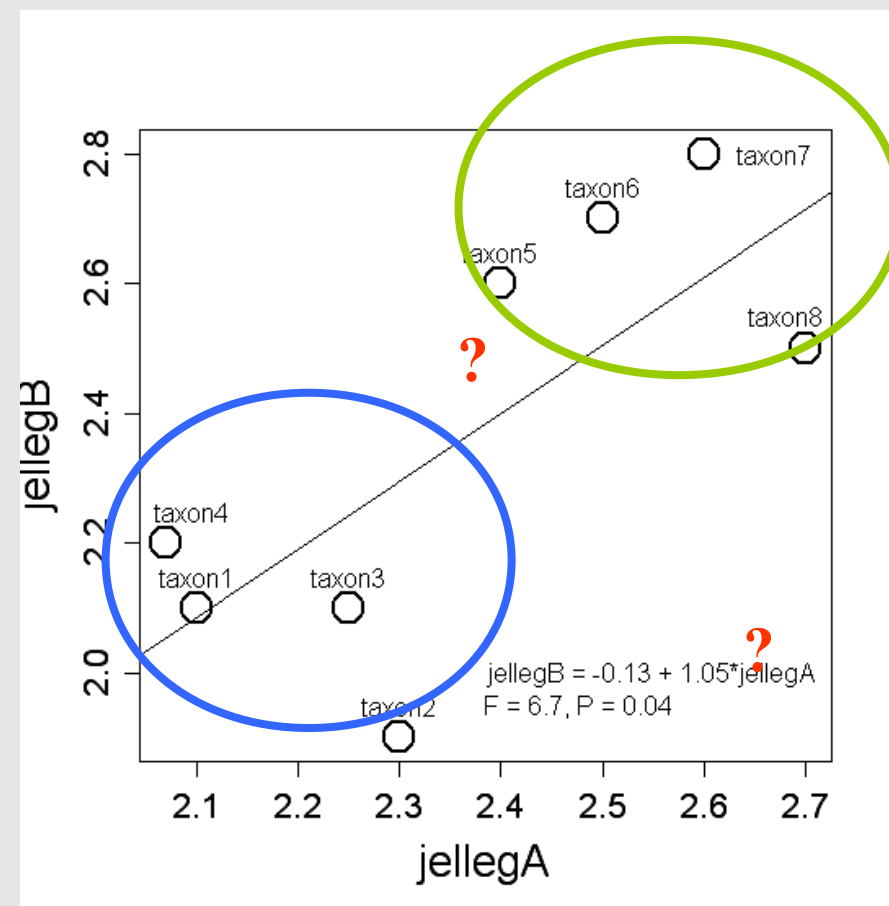
Eredményünk is ezt igazolja

# Evolúciós összehasonlító vizsgálatok

**Probléma: fajok adatai közvetlenül nem összehasonlíthatóak**

- nem teljesül a számos statisztikai próba általános feltétele, a mintaelemek egymástól való függetlensége az evolúciós rokonság miatt

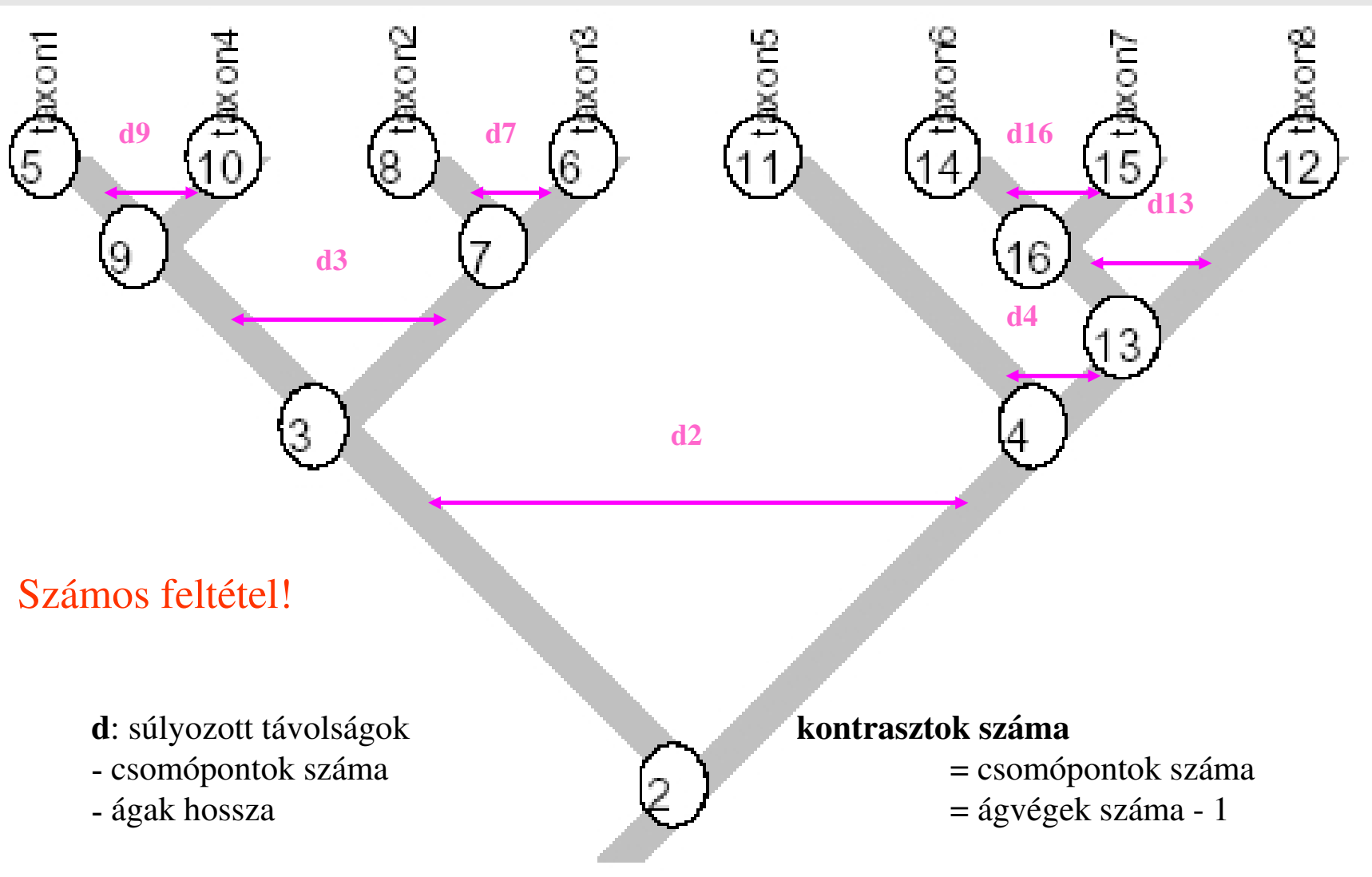
Valaki azonban megszekvenálta a virtuállatokat és elkészítette a ma ismert legjobb *Simulalia* törzsfát



# Evolúciós összehasonlító vizsgálatok

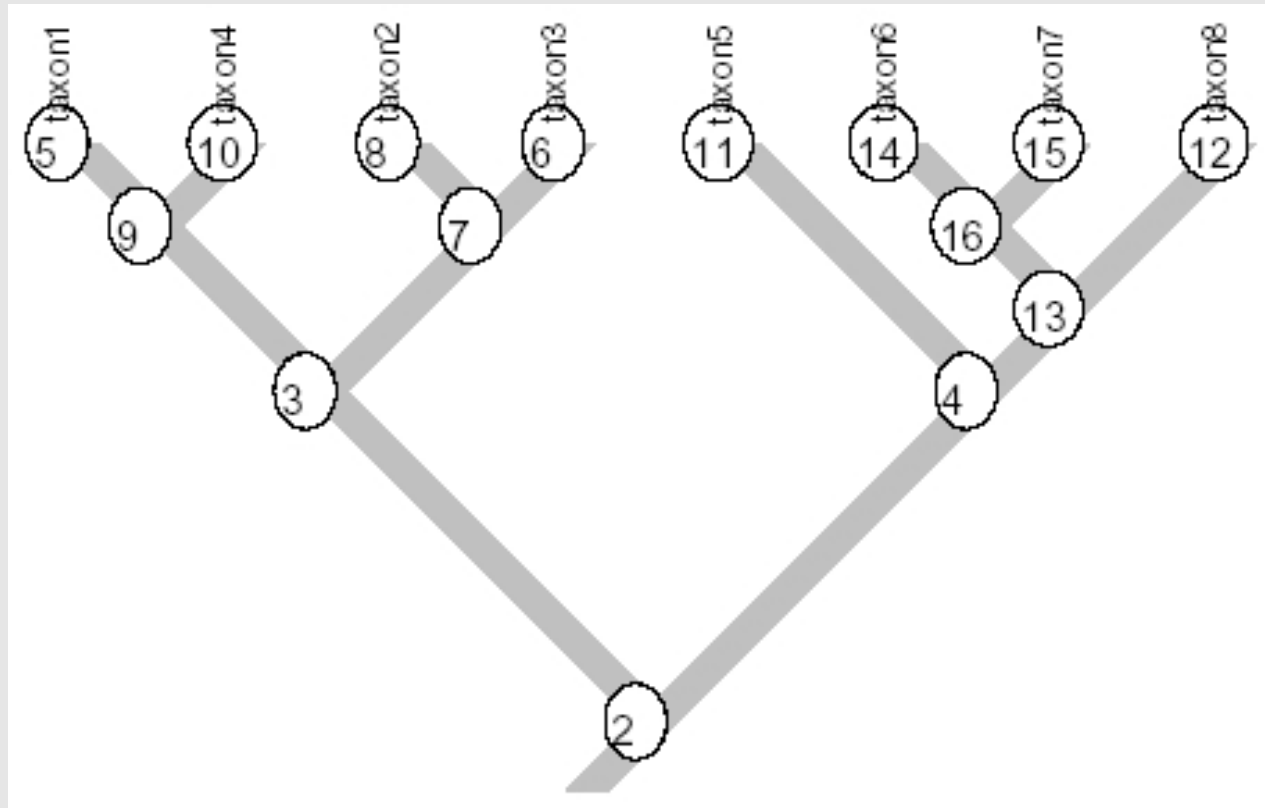
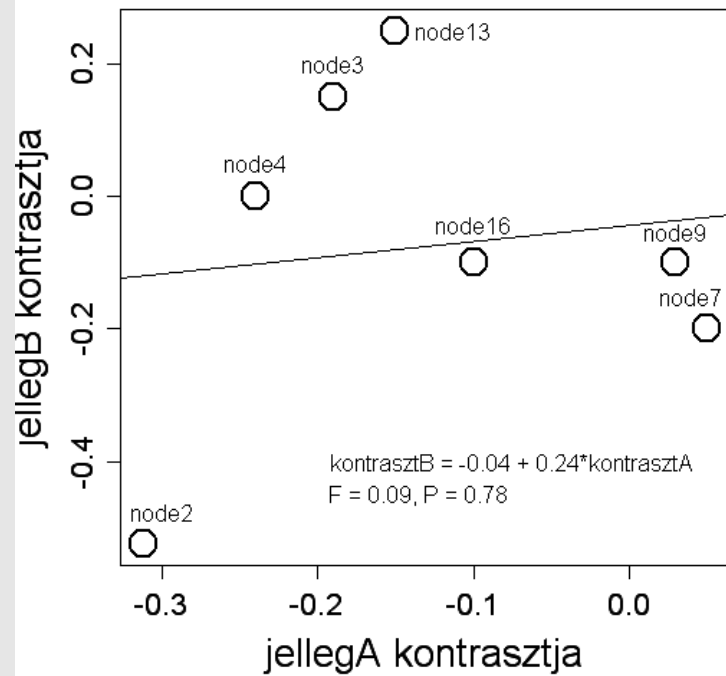
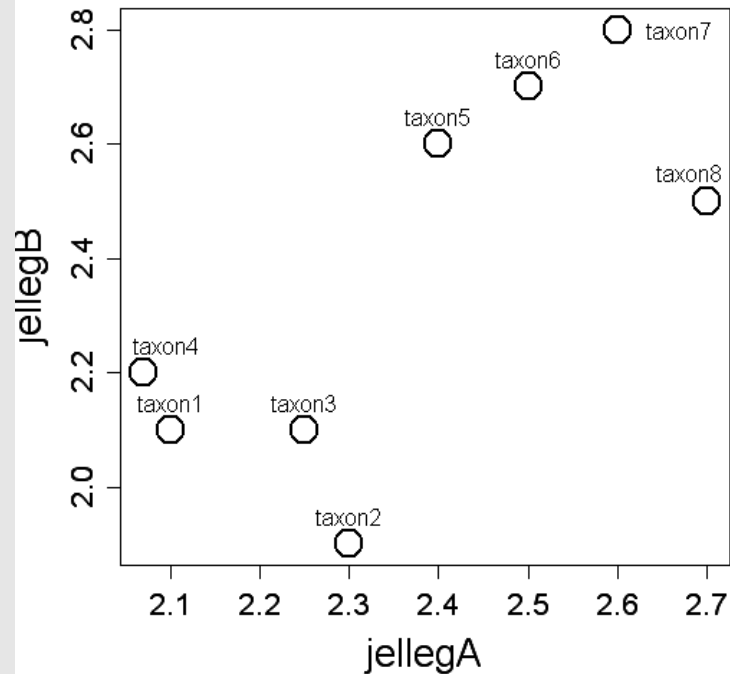
Egy megoldás: független kontrasztok számítása az egyes jellegekre

Felsenstein 1985



# Evolúciós összehasonlító vizsgálatok

## Eredeti és kontraszt változók függésének összehasonlítása



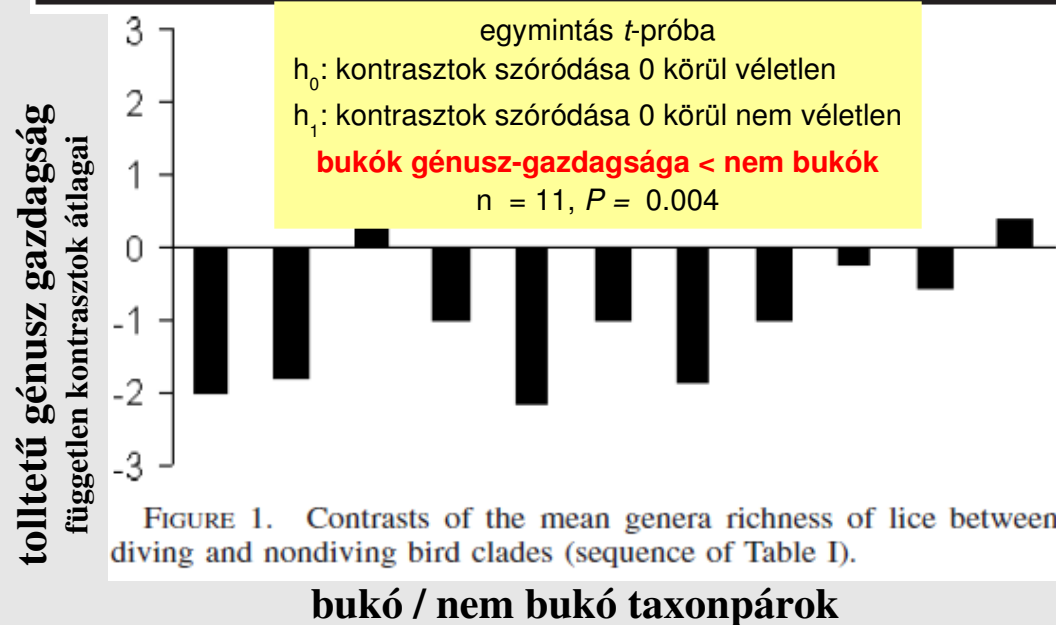
Még szerencse, hogy időben megszereztük az új törzsfát, mielőtt leközlött volna fals eredményeinket. Sokkal többet ugyan nem tudunk a virtuállatokról, de mord szerkesztőkkel sem kell interferálnunk.

# METAELEMZÉS

Csökken-e a víz alá bukó táplálkozással a tolltetvek fajgazdagsága a nem bukó, vízből táplálkozó madarakhoz képest? ► Felső & Rózsa 2006

TABLE I. Host sister clades included in the analyses. See the text for definitions and sources.

| Diver clade                         | Mean genera richness of lice | Mean species richness of host genera | Mean body mass (g) | Nondiver clade                | Mean genera richness of lice | Mean species richness of host genera | Mean body mass (g) |
|-------------------------------------|------------------------------|--------------------------------------|--------------------|-------------------------------|------------------------------|--------------------------------------|--------------------|
| <i>Alcedo, Megaceryle</i>           | 1                            | 10.50                                | 94.86              | <i>Merops</i>                 | 3                            | 22                                   | 32.83              |
| <i>Aptenodytes</i> and allies       | 1.20                         | 2.40                                 | 7,159.27           | <i>Fregata</i>                | 3                            | 5                                    | 1,238.75           |
| <i>Cinclus</i>                      | 3                            | 3                                    | 60.25              | <i>Turdus, Mimus, Sturnus</i> | 2.72                         | 14.55                                | 67.68              |
| <i>Fulica</i>                       | 5                            | 11                                   | 677.60             | <i>Gallinula</i>              | 6                            | 9                                    | 333.79             |
| <i>Gavia</i>                        | 1                            | 5                                    | 3,239.80           | <i>Hydrobatidae</i>           | 3.14                         | 3                                    | 40.21              |
| <i>Netta, Aythya</i>                | 5                            | 7.50                                 | 899.63             | <i>Anser, Branta, Cygnus</i>  | 6                            | 7                                    | 3,754.37           |
| <i>Pandion</i>                      | 3                            | 1                                    | 1,485.50           | <i>Accipitridae</i>           | 4.88                         | 7.94                                 | 1,498.47           |
| <i>Pelecanus occidentalis</i>       | 2                            | 1                                    | 3438               | Other <i>Pelecanus</i> spp.   | 3                            | 6                                    | 6,550              |
| <i>Sterna</i> and allies            | 3.75                         | 10.50                                | 158.49             | <i>Larus</i> and allies       | 4                            | 15.33                                | 643.61             |
| <i>Sula, Anhinga, Phalacrocorax</i> | 2                            | 12.25                                | 1,674.03           | <i>Ardeidae</i>               | 2.56                         | 5.55                                 | 764.26             |
| <i>Uria</i> and allies              | 2.40                         | 2.30                                 | 420.82             | <i>Glareola</i>               | 2                            | 7                                    | 73.60              |



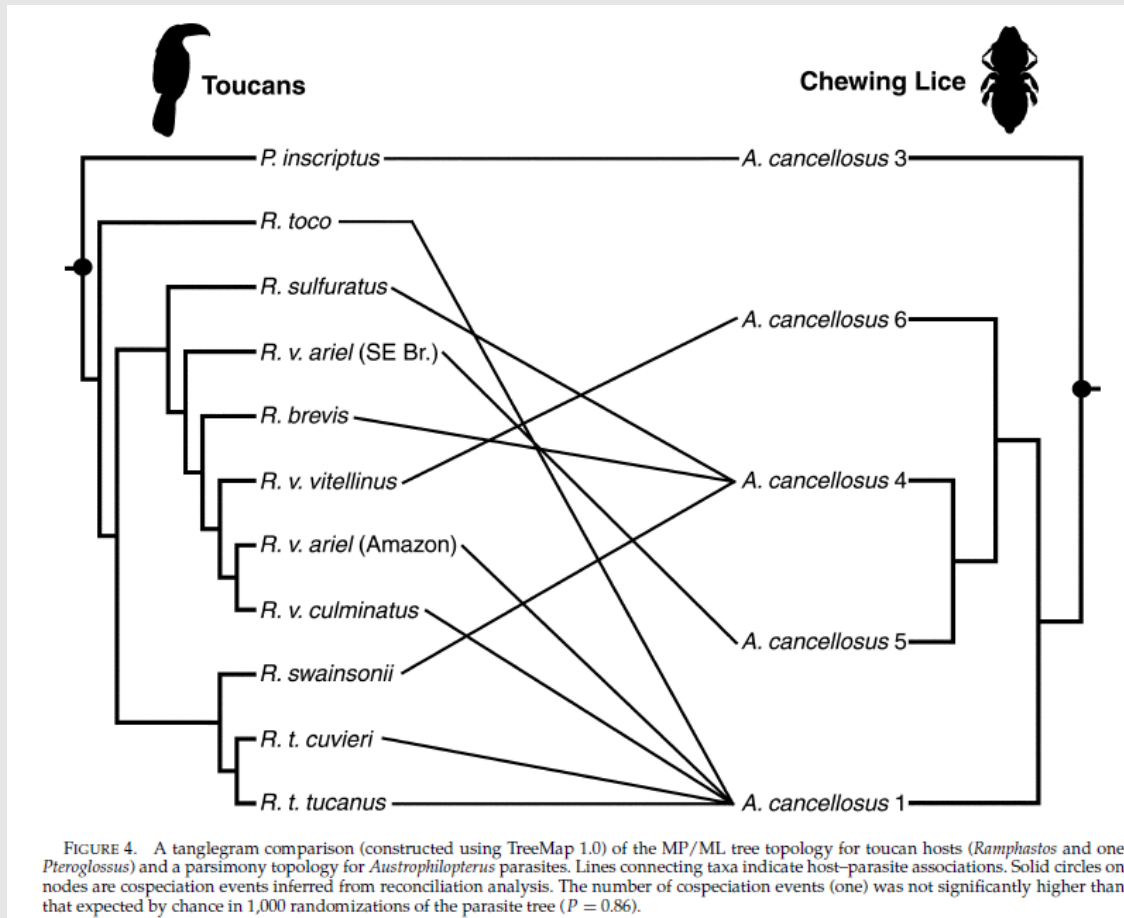
**H:** a víz alá bukás miatt a tollazat mikroklímája nedves, ezért kevesebb tetűfaj képes benne élni a nem bukó, ezért szárazabb tollazatú madarakhoz képest

- törzsfa: Sibley & Alquist (1990) DNS-hibridizáció
- kevésbé kutatott földrészek kizárva
- faj helyett génusz: torzítás csökkentése
  - nagyobb elterjedésű madarakon régióként más-más tetűfajok
  - egyes tetűfajok taxonómiai helyzete tisztázatlan

# METAELEMZÉS

## Koevolválódtak-e a *Ramphastos* tukánok és *Austrophilopterus* tolltetvek?

Weckstein 2005



**H<sub>1</sub>: kospeciáció:** a tolltetvek evolúciója követi a tukánokét, ezért a gazdák és parazitáik törzsfáinak topológiája hasonló

**H<sub>2</sub>: gazdaváltás:** a tolltetvek többször gazdát váltottak, ezért a tukánok és parazitáik törzsfáinak topológiája különböző

- törzsfa:
  - **tukánok:** 2493 bp mtDNS  
MP & ML alapján
  - **tolltetvek:** 379 bp mtDNS & 347 bp fehérjét kódoló nukleáris DNS
- 1 koevolúciós esemény: nem különbözik a véletlen változás esélyétől ( $P \geq 0.89$ )
- több gazdaváltás

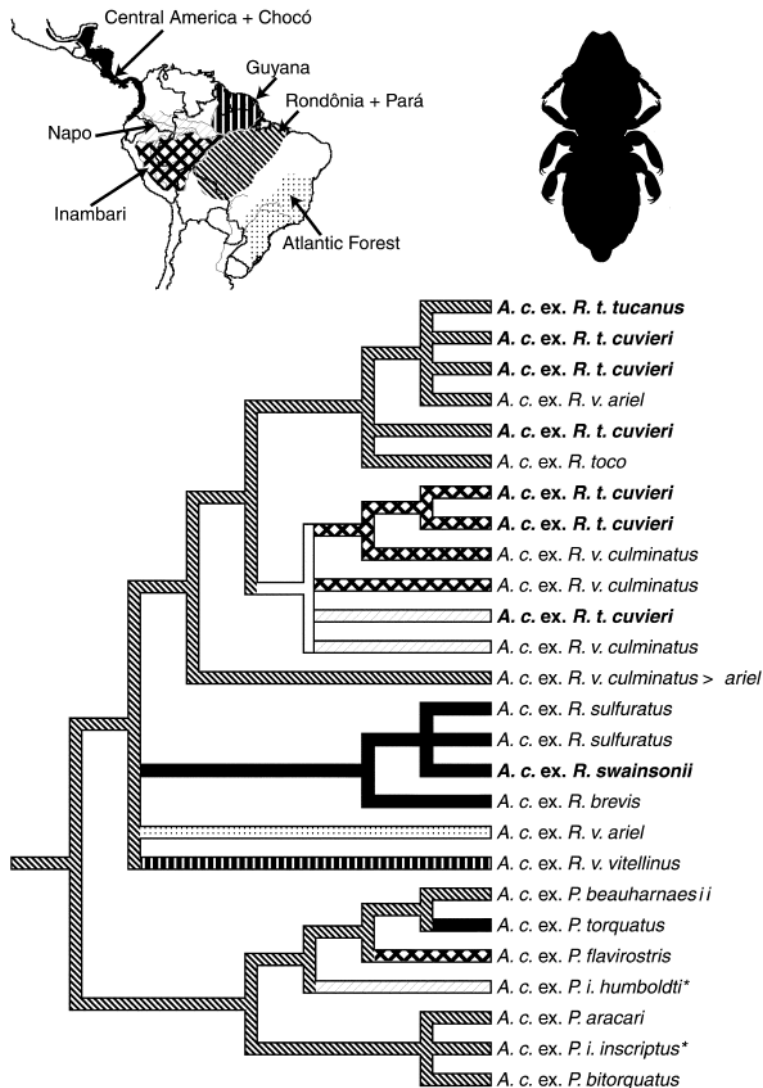
|        |                     |
|--------|---------------------|
| MP:    | maximum parszimónia |
| ML:    | maximum likelihood  |
| bp:    | bázispár            |
| mtDNS: | mitokondriális DNS  |

# METAELEMZÉS

## Koevolválódtak-e a *Ramphastos* tukánok és *Austrophilopterus* tolltetveik?

Weckstein 2005

Hogyan változott az *Austrophilopterus* tolltetvek elterjedése evolúciójuk során Közép- és Dél-Amerikában?



- Közeli rokon tolltetűfajok nem ugyanarról a tukánfajról, viszont azonos régióból származnak

⇒ ezek alapján a gazdaváltás hipotézis a valószínűbb, nincs koevolúció

FIGURE 5. Biogeographic regions mapped onto the MP topology. Patterns on branches match approximated biogeographic regions as indicated on the map. These biogeographic regions are based on Cracraft's (1985) areas of endemism, which are combined in some cases. Names in bold identify taxa collected from the smooth-billed yelping *Ramphastos* clade. Asterisks identify *A. cancellus* from two closely related *Pteroglossus* aracarís.

# MODELLEK

- A **verbális modell** egy konkrét jelenség/rendszer leírása szavakkal, pl. egy hipotézis vagy elmélet megfogalmazása
  - sokkal kevésbé pontos a formális modelleknél
- A **grafikus modell**
  - szemléletes, formális modell hiányában segíthet a verbális modelleknél pontosabban megfogalmazni, így ellenőrizni is hipotéziseket
  - formális modellek szemléltetésére is használható: segíti a megértést azok számára, akik kevésbé járatosak a matematikában
- A **formális modell** egy konkrét jelenség/rendszer matematikai leírása – lehet olyan rendszer is, ami formálisan jön létre, amihez nem ismerünk a természetben jelenséget

# MODELLEK

## Mikor érdemes egy madárszülőnek fészekalját védenie ragadozókkal szemben?

► Montgomerie & Weatherhead 1988

### Verbális

A madarak fészekaljvédő viselkedésének intenzitása függ attól, mennyivel növeli a védelem a jelenlegi fészekaljból származó fitnessznyereséget, azaz a fiókák túlélését és várható szaporodását, és hogyan befolyásolja a szülők várható szaporodását a jövőben. A védelem költségei pl. sérülés, ami jelentősen csökkentheti mind a jelenlegi, mind a jövőben várható szaporodási sikert.

### Grafikus

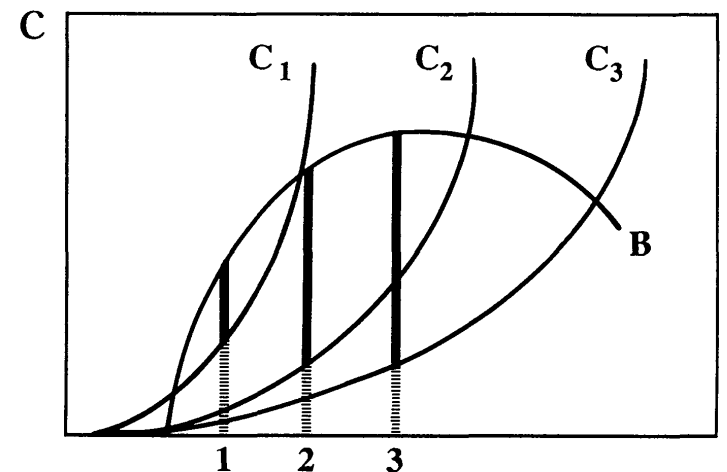
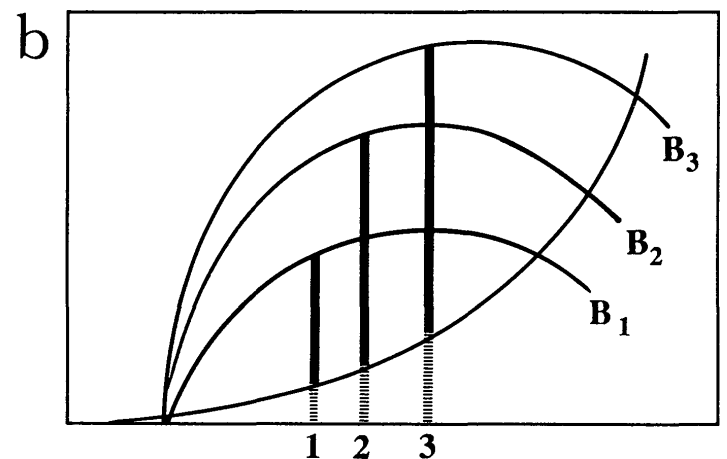
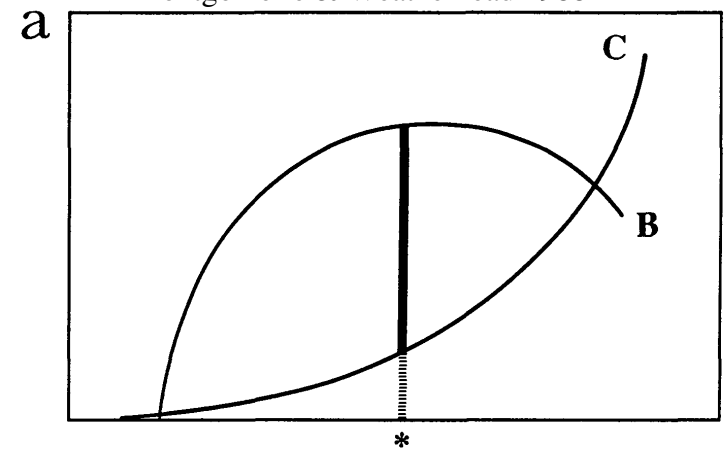
Eltérő költségek / nyereségek esetén más lesz a viselkedés optimuma is \* : optimum

### Formális

$$V_x = b_x + \sum \frac{l_t}{l_x} b_t$$

Fitnessz: reprodukív érték (V) = jelenlegi (b) + reziduum (várható, Σ...)  
x: jelen, t: jövő, b: szaporodás, l: túlélés

Montgomerie & Weatherhead 1988



Fitnessz költség (C) és nyereség (B)

Fészekaljvédő viselkedés intenzitása

# MODELLEZÉS

- A **formális modell** egy konkrét jelenség/rendszer matematikai leírása – lehet olyan rendszer is, ami formálisan jön létre, amihez nem ismerünk a természetben jelenséget
  - nem matematikailag is jól interpretált
  - a rendszer egyes tulajdonságait változókkal írjuk le, és keressük e változók közötti összefüggéseket
  - a rendszer más tulajdonságait elhanyagoljuk: szükségszerű; a feltevések pontos definiálása alapvető
    - ⇒ *a modell minden esetben egyszerűsítés*
    - a modell működtetésekor, következtetések levonásánál a feltevéseket figyelembe kell venni
- Biológiai modell: paraméterek, összefüggések biológiai jelentése értelmezhető, van szemléletes jelentése.

# MODELLEZÉS

- Mikor készítsünk formális modellt?
  - rendszerek általános tulajdonságaira vagyunk kíváncsiak: minél kevesebb specifikus paraméterre van szükségünk
  - kérdésünk pontos megfogalmazása modell nélkül lehetetlen
- vannak modellezők, akik szerint az általános tudományos kérdések mind ilyenek: attól lesz valami tudomány, hogy formális modelleket használ

# MODELLEZÉS

A modell megadása gyakran egy képlet:  $E = mc^2$

- A modell megoldása

- számítás

- képlet esetén ritkán, egyszerű esetekben *analitikus*: a képlet megoldása *szimbolikus számolás*, pl. egyensúlyi pont kifejezése képletből
    - ha túl bonyolult a modell, *numerikus*: számolás számítógéppel ~szimuláció!
    - *szimulációs modell*: ha nincs analitikus megoldás, vagy nagy összetett rendszert vizsgálunk, vagy szeretnénk megtudni, stabil-e egy stratégia 10000 generáción keresztül
    - egyszerű, általánosan programozható, kész szoftverek
    - a szimulációs modell és a numerikus megoldás hátránya, hogy végig kell pásztázni a paraméter-teret, vagyis sok paraméter-kombinációból és kezdeti feltételből kell kiindulni és kiszámoltatni a megoldás egyensúlyi helyzetét.
    - az egyensúly stabilitásának megállapításához gyakran nem szükséges analitikusan megoldani a modellt, konkrét változók értékeire nézve

# JÁTÉKELMÉLETI MODELL

Mikor éri meg egy szülőnek gondozni utódait? Maynard Smith 1977

## Szülői befektetés ESS modellje

### *Paraméterek és feltételezések*

#### tojások túlélése

$P_0$ : nincs gondozás,  $P_1$ : egy szülő gondoz,  $P_2$ : kétszülős gondozás

$$P_0 < P_1 < P_2$$

#### tojások száma

$w_d$ : dezertáló ♀,  $w_g$ : gondozó ♀

$$w_d > w_g$$

#### ♂ újrapárosodási esélye

$p$

**Valuta:** fitnessz (utódszám)

### **Egyszerűsítés**

csak a ♂ párosodhat újra, a ♀ nem

#### **ESS: evolúciósan stabil stratégia**

Olyan stratégia, amelyet ha egy populáció egyedeinek többsége követ, rovására egy újonnan megjelenő, eltérő, ritka stratégia nem tud elterjedni.

Feltétele:  $\text{freq}(\text{ESS}) \gg \text{freq}(\text{mutáns}) \Rightarrow$  a rezidens határozza meg a rezidens és a mutáns környezetét

# JÁTÉKELMÉLETI MODELL

Mikor éri meg egy szülőnek gondozni utódait? Maynard Smith 1977

## Szülői befektetés ESS modellje

### Kifizetési mátrix *payoff*

|   |          | ♀        |           |
|---|----------|----------|-----------|
|   |          | nyereség |           |
| ♂ | gondoz   | ♂        | $w_g P_2$ |
|   |          | ♀        | $w_g P_2$ |
|   | dezertál | ♂        | $w_d P_1$ |
|   |          | ♀        | $w_d P_1$ |

### ESS lehet

#### szülő viselkedése

#### feltétel

#### alternatíva

1. ♂ d, ♀ d

$P_0(1+p) > P_1$

vagy a ♂ g

2. ♂ g, ♀ d

$w_d P_0 > w_g P_1$

vagy a ♀ g

3. ♂ d, ♀ g

$P_1 > P_0(1+p)$

vagy a ♂ d

4. ♂ g, ♀ g

$w_d P_1 > w_g P_2$

vagy a ♀ g

$P_1(1+p) > P_2$

vagy a ♂ g

$w_g P_1 > w_d P_0$

vagy a ♀ d

$P_2 > P_1(1+p)$

vagy a ♂ d

$w_g P_2 > w_d P_1$

vagy a ♀ d

### alternatív stratégiák:

♂ d, ♀ d – ♂ g, ♀ g &  
♂ g, ♀ d – ♂ d, ♀ g

- ♂ g, ♀ d pl., ha a ♀ sok tojást rak és nem képes gondozni:

$$w_g \ll w_d, P_1 \gg P_0, P_1 \sim P_2$$

halak

- ♂ d, ♀ g pl., ha a ♂♂ újrapárosodási esélye nagy  
madarak, emlősök

- ♂ g, ♀ g, ha  $P_2 \gg P_1$ ,  
énekesmadarak

# SZIMULÁCIÓS MODELL

Hogyan építik a *Polistes* papírdarazsak fészkeiket? ► Karsai & Pénzes 1998

## Különböző *Polistes* fajok fészkeinek alakja

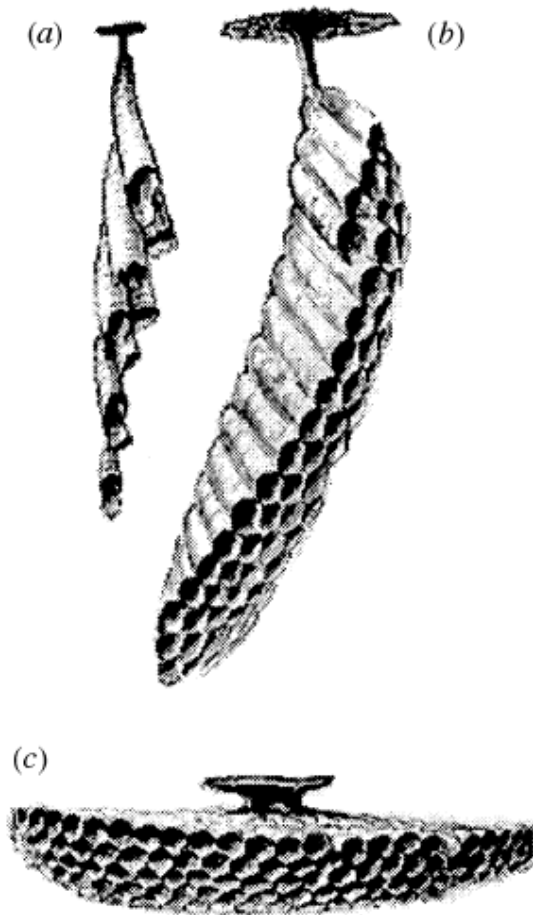


Figure 1. Side view of *Polistes* nests. (a) *P. (A.) goeldii*: eccentric, slender, inclined design; (b) *P. (A.) infuscatus*: eccentric, oval, inclined design; (c) *P. (F.) flavus*: centred, round, horizontal design. (Redrawn from Evans & West-Eberhard 1970).

## Fészkek alakok a modellekből

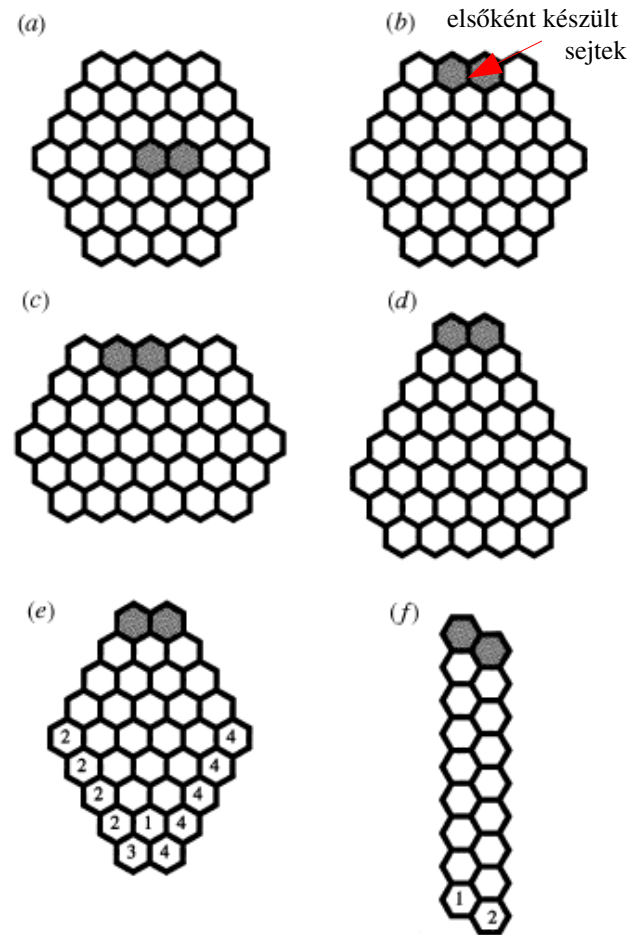


Figure 4. Nests are generated by different models. An example of species whose nests resemble a given structure is given in parentheses (see also table 2). The first two cells are shaded. Numbers denote the building blocks and the initiation sequence (see model section). (a) Model 1: *P. (P.) dominulus*. (b) Model 2: *P. (A.) exclamans*. (c) Model 3: *P. (F.) fuscatus*. (d-f) Model 4: *P. (A.) annularis*, *P. (A.) canadensis*, and *P. (A.) goeldii*, respectively).

- 1000 Monte Carlo szimuláció
- 4 fészkepítési modell 500 sejtes állapotig
  - a–c: 1–3. modell
  - d–f: 4. modell
  - mind a 6 alak valamelyik faj fészkére jellemző

# SZIMULÁCIÓS MODELL

Hogyan építik a *Polystes* papírdarazsak fészkeiket? ► Karsai & Péntes 1998

Fészekátmérők arányainak változása  
a sejtszám növekedésével

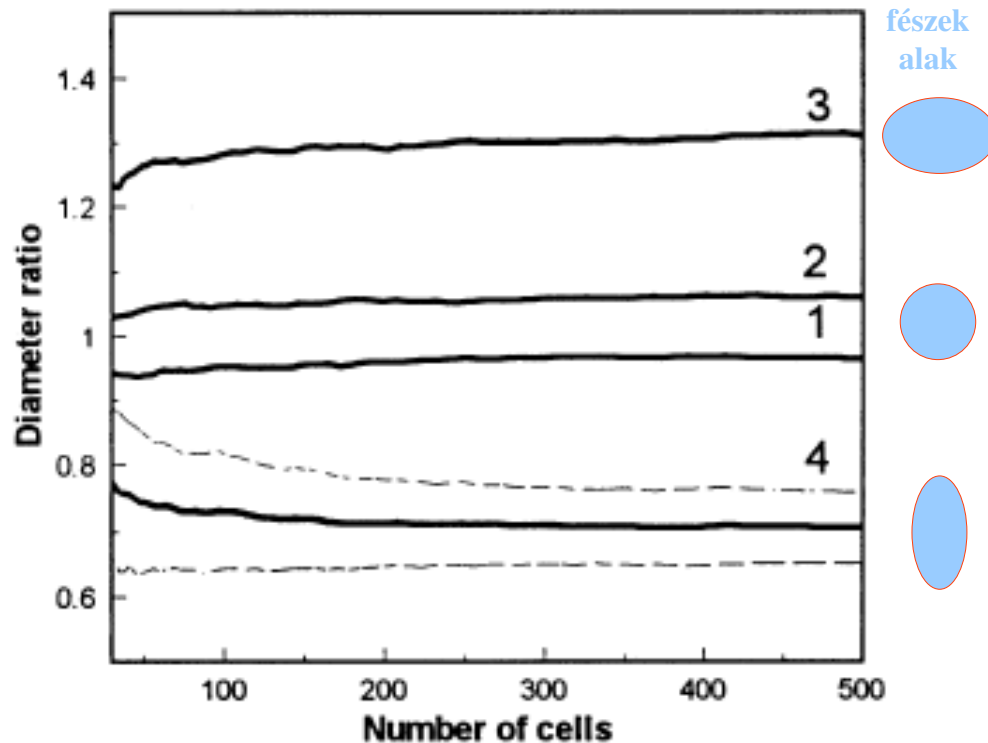
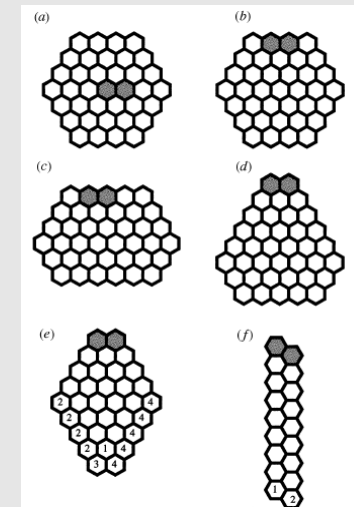


Figure 5. Average diameter ratio (thick line) and its standard deviation (broken line for model 4—the standard deviations for other models had the same magnitude, but they omitted to simplify the figure) of structures generated by 1000 Monte-Carlo simulations according to the models (numbers) (see also figure 4 and the text).

## Modellek közti különbség:

- Mely falak mellé lehet újabb falat építeni, melyek mellé nem
  - **1. modell** (a): minden oldalra egyforma valószínűséggel épül új  
 $w_1 = w_2 = \dots = w_6$
  - **2. modell** (b): az egyik oldal tiltott, egy másik 2× gyakoribb / többi nem tiltott  
 $w_1 = 0, w_2 = w_3 = \dots = w_5 > 0, w_6 = 2w_2$
  - **3. modell** (c): az egyik oldal tiltott  
 $w_1 = 0, w_2 = w_3 = \dots = w_6 > 0$
  - **4. modell** (d–f): 3 szomszédos oldal tiltott  
 $w_1 = w_2 = w_3 = 0, w_4 = w_5 = w_6 > 0$



# SZIMULÁCIÓ

*sensu* Maynard Smith (1974)

- A **szimuláció**  $\neq$  szimulációs modell; egy konkrét rendszer ismeretében alkalmas a rendszer működésének előrejelzésére (és nem célja új elmélet kidolgozása)
- minél több részletet adunk meg (minél több paramétert ismerünk, amik jól írják le rendszerünk működését), annál pontosabb előrejelzést kapunk  
 $\Rightarrow$  ***pontos ismeretekre van szükség a hatékony jósláshoz***

**Mennyi időbe telik egy terület veszettség-mentesítése, ha orális vakcinát használunk?**

**Adjuk meg**

- a terület nagyságát
- a veszettséget terjesztő fajokat és állományaik méretét
- e fajok egyedeinek földrajzi terjedését gátló tényezőket (pl. magashegység, folyó)
- az oltóanyag terjesztésének módját
- stb...

## 7. ENGEDÉLYEK

- Kísérletekhez állatvédelmi, etikai engedély kell gerincesekre (*Vertebrata*)  
Munkahelyi Állatkísérleti Bizottságon keresztül  
ehhez dokumentumok: <http://www.univet.hu/intra>  
kari megbízott (2008) Dr. Vajdovich Péter (Belgyógyászati Tanszék és Klinika)
- Védett fajokkal végzett munkához
- Védett területen történő munkához
  - ▶ természetvédelmi hatósági engedély szükséges
- Védett területeken az utak használatát gyakran az erdészetek felügyelik, gépjárművel történő behajtáshoz engedély kell
- Magánterületek gazdáitól kérjünk szóban engedélyt a belépéshez

# 8. ADATOK

- Minél előbb vigyük adatainkat számítógépre (pl. *Calc*, *Excel*)
  - minél egyszerűbb adatstruktúra: minden változó egy oszlop, változók kategóriás felosztása csoportokra külön csoportosító változóval történjen
  - alapadat file-ban ne legyen semmi más (pl. összesítés átlagokról, ábra)
  - minden változást az alapadat file-ba kell vezetni
- metaadatok: ki (elérhetőséggel), mikor, hol gyűjtötte az adatokat, milyen módszerrel, eredeti adattároló (pl. terepfüzet) lelőhelyével
- txt-formátum általánosan használható, operációs rendszerek, szoftverek közt jól közlekedik
- **Egyetlen példány NEM ELÉG!!!**
  - leég az intézmény, ellopják a gépet, vagy tönkremegy az adattároló egységünk, oda a munkánk: földrajzilag távol eső helyeken tartsunk másolatokat terepi és labor jegyzőkönyvekből és elektronikusan tárolt adatainkból is
- Elektronikus adathordozók nagyon gyorsan változnak, ehhez folyamatosan alkalmazkodni kell – ma már külön kell a számítógéphez floppy drive-ot rendelni

[Jó összefoglaló a ► *digiKam Handbook* Kulzer et al. 2010 *Using digiKam > Digital Asset Management (DAM) with digiKam > Protect your images* fejezetében]

# Digitális adatvesztés okai

- adathordozó öregedése
- átírási hibák másolásakor
- hosszú távú digitális file-formátumok hiánya
- ősi hardver (floppy)

| adatvesztés milyen hibából fakad | érzés (%) | tapasztalat (%) |
|----------------------------------|-----------|-----------------|
| hardver-, rendszer               | 78        | 56              |
| felhasználó                      | 11        | 26              |
| szoftver                         | 7         | 9               |
| vírus                            | 2         | 4               |
| katasztrófa                      | 1-2       | 1-2             |

a Kroll Ontrack Inc. adatai; *In: Kulzer et al. 2010*

# Digitális archíválás

## CD/DVD/BlueRay

- megbízhatóság néhány év
- megbízhatóság archíváláshoz készült minőségénél ~ 100 év
- kémiai öregedés, UV, por, karcolás
- jó minőségű író/olvasó és adathordozó számít
- adatjavítás: ► *IsoBuster* (*MS, Linux – Wine* alatt)

Leglassabb írás jó minőségű íróval, archíváláshoz való lemezre szabad file-formátumban

- visszaolvasás
- címkézés (tartalom leírás, dátum, szerző)
- tárolás tiszta, száraz, sötét helyen elzárva
- átírás új médiumfajtára, mielőtt e régit olvasót kidobnánk

# Digitális archíválás

## Merevlemez (HDD)

- 35 – 45 °C optimális üzemi hőmérséklet;
- VESZÉLY: rázkódás, ütődés, hideg, sok ki- & bekapcsolás
- megbízhatóság 4 - 5 év
- ha a HDD furcsa hangot ad: mielőbbi adatmentés ► *dd utility* szoftverrel

## Solid state drive (SSD; USB-kulcsok, memóriakártyák, ...)

- nincs mozgóalkatrész: kevesebb lehetőség meghibásodásra / HDD
- főleg hordozható adathordozókban: elejtés, lopás
- gyakori adatvesztés: nem megfelelő leválasztás PC-ről
- nem írható végtelenszer újra még elvben sem

# Digitális archíválás

## Web tárhelyek (HDD)

- nagy cégeknél biztonságos: adatok több szerveren, rendszeres back-up, átírási hibákat javítja a TCP-protokoll
- viszonylag drága
- adatcsere lassú

## Átírási hibák

szoftver  $\Rightarrow$  OS  $\Rightarrow$  I/O kontroller  $\Rightarrow$  HDD

- számos helyen történhet hiba: pl. képfájl-oknál átlag 3000-ből 1
- TCP-protokollhoz hasonló javítórendszerek fejlesztése folyamatban

# Digitális archíválás

## Túlfeszültség

- adatvesztések 1 %-a
- egy számítógép napi  $2 - 3 \times$  kap túlfeszültséget az elektromos hálózathoz
- villám

⇒ villám elleni védelemmel ellátott túlfeszültségvédő *surge protector*

- *MINDEN perifériára (nyomtató, modem, stb.)*

## Felhasználó hibái

- lopás & baleset az adatvesztés fő okai: notebook 86 %, desktop 46 %,
  - csak lopás notebooknál 50 %
- vírusok, stb.: kevésbé jelentős; főleg *MS*, ritka *MAC*, *GNU/Linux*
- *pánik* gyakori probléma: ha baj van, jobb megállni egy pillanatra

# Digitális archíválás

## Back-up!!!

- **rendszer adatmentés szükséges egymástól távol eső helyeken!**

# TERVEZÉS

## Mire figyelünk?

- A kutatómunka tervezése során célunk, hogy minél szélesebb körben általánosíthatóak legyenek majd eredményeinkből levonható következtetések. Ehhez:
  - mintavételünk legyen reprezentatív, határozzuk meg, mi a mintánk, hogy a megfelelő mintaelemszámokat tervezhessük
  - megfigyeléseinkhez, ha mód van rá, tervezzünk kísérletet
  - ne feledkezzünk meg a kontrollról / alapszint meghatározásáról
  - statisztikai ismereteink alapján tervezzük meg előre az elemzéseket, így számításba vehetjük elvárásainkat a változóinkkal szemben
  - a statisztikai elemzést MINDIG feltáró elemzéssel kezdjük, ÁBRÁKKAL
  - a statisztikai hipotézisvizsgálatot MINDIG előzze meg az alkalmazni kívánt modellek feltételeinek ellenőrzése – a  $p$ -értéket csak akkor nézzük meg, amikor már megbizonyosodtunk a feltételek teljesüléséről. Ha a feltételek sérülnek, a  $p$ -értékre kapott becslés torzít, abból következtetéseket levonni CSALÁS!

# KÖSZÖNETNYÍLVÁNÍTÁS

## Szakmai lektorok:

- Lang Zsolt, SZIE ÁOTK ► Biomatematikai és Számítástechnikai Tanszék
- Szabó Péter, SZIE ÁOTK ► Biológiai Intézet

## Felhasznált óravázlatok:

- Liker András 2000. Kutatástervezés. SZIE ÁOTK Zoológiai Intézet, Ökológiai Tanszék.
- Kövér Szilvia 2007. Modellek az ökológiában és az evolúciógenetikában SZIE ÁOTK ► Biológiai Intézet, *bemutató*.
- Zimmer Márta 2007. A kísérletezés módszertana. BME GTK ► Kognitív Tudományi Tanszék *bemutató*.

## Felhasznált szabad szoftverek:

- ImageJ
- IrfanView
- KCharSelect
- Mesquite
- Mozilla Firefox / FireFTP
- Okular
- Open Office / Sun Presentation Minimizer
- R / irr
- RKWard
- Tinn-R

## Szabad operációs rendszerek:

- Gnome
- GNU
- KDE
- Kubuntu
- Linux
- Ubuntu

## Szabad betűtípusok:

LINUX LIBERTINE O C  
DejaVu Sans  
DejaVu Serif  
FreeMono  
FreeSans  
FreeSerif

# HIVATKOZÁSOK

- Arnqvits G & Mårtensson T 1998. Measurement error in geometric morphometrics: empirical strategies to assess and reduce its impact on measures of shape. *Acta Zool. Acad. Sci. H.* 44, 73-96.
- Carroll L 1871. *Through the Looking Glass and What Alice Found There.* Macmillan
- DerSimonian R & Laird N 1986. Meta-analysis in clinical trials. *Controlled Clin. Trials* 7, 177-88.
- Dobzhansky TD 1973. Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teacher* 35, 125-129.
- Felsenstein J 1985. Phylogenies and the comparative method. *Am. Nat.* 125, 1-15.
- Esterházy P 2000. *Harmonia Cælestis.* Magvető Könyvkiadó, Budapest, pp. 64-65.
- Felső B & Rózsa L 2006. Reduced taxonomic richness of lice (Insecta: Phthiraptera) in diving birds. *J. Parasitol.* 92, 867-869.
- Fleiss JL, Levin B & Paik MC 2003. *Statistical Methods for Rates and Proportions.* 3<sup>rd</sup> ed. John Wiley & Sons, Inc. Hoboken, New Jersey
- Gotelli NJ & Ellison AM 2004. *A Primer of Ecological Statistics.* Sinauer, Sunderland, USA.
- Haldane JBS (1927) 2002. *Possible Worlds.* Transaction Publishers New Brunswick, New Jersey, p. 286.
- Karsai I & Péntes Zs 1998. Nest shapes in paper wasps: can the variability of forms be deduced from the same construction algorithm? *P. Roy. Soc. Lond. B* 265, 1261-1268.
- Kis J 2003. Parental behaviour in Kentish plover and northern lapwing. PhD-thesis, Eötvös University, Budapest
- Kis J, Liker A & Székely T 2000. Nest defence by lapwings: observations on natural behaviour and an experiment. *Ardea* 88, 155-163.
- Kulzer G, Caulier G, Stone E, Doerr O, Hoelzer R & Ahrens J 2010. *The digiKam Handbook.*
- Lessells CM & Boag PT 1987. Unrepeatable repeatabilities: a common mistake. *Auk* 104, 116-121.
- Lendvai ÁZ, Kis J, Székely T & Cuthill IC 2004. An investigation of mate choice based on manipulation of multiple ornaments in the Kentish plover. *Anim. Behav.* 67, 703-709.
- Maynard Smith J 1974. *Models in Ecology.* Cambridge University Press, Cambridge.
- Maindonald J 2000. *The Design of Research Studies – A Statistical Perspective.* Part I. 2<sup>nd</sup> ed., unpublished manuscript
- Maynard Smith J 1977. Parental investment: A prospective analysis. *Anim. Behav.* 25, 1-9.
- Maynard Smith J 1982. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge.
- Maynard Smith J 1989. *Evolutionary Genetics.* Oxford University Press, Oxford pp. 129-130.
- Montgomerie RD & Weatherhead PJ 1988. Risks and rewards of nest defence by parent birds. *Q. Rev. Biol.* 63, 167-187.
- Reiczigel J, Harnos A & Solymosi N 2007. *Biostatistika – nem biostatistikusoknak.* Pars Kft, Nagykovácsi.
- Rejtő J 19???. Vanek úr Párizsban. Alexandra, Pécs (2006) pp. 30-31.
- Sinervo B & Lively CM 1996. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380, 240-243.
- Székely T, Cuthill IC & Kis J 1999. Brood desertion in Kentish plover: sex differences in remating opportunities. *Behav. Ecol.* 10, 185-190.
- Székely, T & IC Cuthill 1999. Brood desertion in Kentish Plover: the value of parental care. *Behav. Ecol.* 10, 190-197.
- Weckstein JD 2004. Biogeography explains cophylogenetic patterns in toucan chewing lice. *Syst. Biol.* 53, 154 - 164.
- White GC, Anderson DR & Burnham KP 1982. *Capture – recapture and removal methods for sampling closed populations.* Los Alamos National Laboratory, New Mexico.

# Copyright

Ez a Mű a Creative Commons Nevezd meg! - Ne add el! - Így add tovább! 2.5 Magyarország Licenc feltételeinek megfelelően szabadon felhasználható. A licenc szövegének megtekintéséhez látogasd meg az alábbi oldalt,

<http://creativecommons.org/licenses/by-nc-sa/2.5/hu/>

vagy küldj egy levelet a következő címre: Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

## ***Így idézd:***

Kis J 2010. Kutatástervezés – előadásvázlatok. SZIE ÁOTK Biológiai Intézet, Budapest

URL: [ftp://www.univet.hu/biostat/kutternv/Kutatastervezes\\_AOTK-PhD.pdf](ftp://www.univet.hu/biostat/kutternv/Kutatastervezes_AOTK-PhD.pdf)

