Kettőnél több csoport vizsgálata

Makara B. Gábor

Három gyógytápszer elemzéséből az alábbi energia tartalom adatok származtak (kilokalória/adag egységben)

Három gyógytápszer elemzésébő		
A	B	C
230	225	200
210	205	208
240	245	203
250	235	190

Kérdések:

1. Van-e különbség A, B, és C kalóriatartalma között?
2. Milyen sorrend állítható fel a gyógytápszerek között?

Több csoport statisztikai elemzése: pajzsmirigy gyulladás - szérum hormon értékek

Csoport	fT4	fT3	fT4/fT3
Kontroll (20)	14 ± 1.0	4.2 ± 0.5	3.3 ± 0.5
Beteg (18)	14 ± 2.0	4.4 ± 0.5	3.2 ± 0.4
Beteg – kezelt (35)	16 ± 2.0	4.0 ± 0.5	4.0 ± 0.6

Több csoport statisztikai elemzése: pajzsmirigy gyulladás - szérum hormon értékek

Csoport	fT4	fT3	fT4/fT3
Kontroll (20)	14 ± 1.0	4.2 ± 0.5	3.3 ± 0.5
Beteg (18)	14 ± 2.0	4.4 ± 0.5	3.2 ± 0.4
Beteg – kezelt (35)	**16 ± 2.0****	**4.0 ± 0.5***	**4.0 ± 0.6****

* $p<0.05$, ** $p<0.01$ vagy akár ***0.001

Miért nem hasonlítunk össze minden csoportot párosával?

- Mert rossz hatásfokú
- Mert torzíthatja döntéseinket
- Minden páros összehasonlításnál 1:20 arányban (vagy a szignifikancia szinttől függő arányban) van esélyünk hibás döntést hozni.
  - Az egy vizsgálatban végzett több páros összehasonlítás hibái halmozódnak
  - A vizsgálat összes , valós első fajú hibája (α) magasabb, mint 0,05

- A tudományos irodalom reprodukálhatósági krízisének egyik fontos oka a rejtett vagy tudatlan hibahalmozás a szignifikancia utáni hajszában

A több csoport elemzése két lépésből áll

1. Van-e különbség
   a csoportosított eredmények halmazában
2. <u>Ha van különbség</u>,
   akkor megvizsgáljuk a csoportok közötti
   lehetséges, vagy tervezett vizsgálatokban
   a páros vagy többes eltérések valószínűségét

A variancia elemzés alapgondolata: az összes mintában a varianciát két független módszerrel becsüljük

- Az ANOVA alkotója **R.A. Fisher** (biológus és statisztikus, egy angliai mezőgazdasági kísérleti állomáson, 1918-25 között).
- Zseniális felismerése:
  Több csoporton együtt végzett kísérletben a null hipotézis, H_0 úgy is vizsgálható, hogy
  a populáció varianciát **becsüljük két független módszerrel**.
- Az egyik módszer: a mintákon/csoportokon belüli (belső) szóródásból következtet a varianciára
- A másik módszer ugyanazon minták átlagainak szóródásából következtet ugyanarra a varianciára.
- A varianciák hányadosának várható értéke 1, eloszlása az F-eloszlás (Fisher-Snedecor eloszlás)
- A t próba ekvivalens a két csoporton végzett varianciaanalízissel

Az ANOVA és a t próba kapcsolata

- A t próba képletében a nevezőben a 2 csoport együttes belső szórása van
- A számlálóban is szórásnak megfelelő érték, a 2 minta átlagának különbsége van
- Ez nem más, mint a két szám eltérése az átlaguktól, osztva n-1 -el, ami n=2 esetben nem más, mint n-1=1
- A számlálóban és a nevezőben ugyanazon értékre 2 független becslés szerepel, melynek négyzeteinek hányadosa F eloszlású

A t próba képlete, és annak átalakítása

$$t = \frac{m_1 - m_2}{\dfrac{s_{1,2}}{\sqrt{n_1 + n_2 - 2}}}$$

Ha a képlet mindkét oldalát négyzetre emeljük:

$$t^2_{n_1+n_2-2} = \frac{\dfrac{(m_1 - m_2)^2}{1^2}}{\dfrac{s^2_{1,2}}{n_1 + n_2 - 2}}$$

akkor a jobb oldalon
két variancia hányadosát kapjuk, azaz

$$t^2_{n_1+n_2-2} = F_{1,n_1+n_2-2}$$

A négyzetes összeg additív elemekre bontható

- A minta elemeinek távolságát a teljes minta „nagy átlagától” becsüljük a négyzetes összeggel
- $\Sigma\,(x_{nagy\text{-}átlag} - x_i)^2$
- A négyzetes összeg particionálható (additív módon részekre bontható)
- Az egyes részeket úgy bontjuk részekre, hogy azok a szórás meghatározott értelmezésű részeinek feleljenek meg,
- Összes szórás = A csoportok (belső) szórása + a csoportátlagok szórása
- Illusztráció:
  $\Sigma\,(x_{nagy\text{-}átlag} - x_i)^2 =$
  $= \Sigma\,(\Sigma[x_{nagy\text{-}átlag} - csoportátlag_j\,]^2 + \Sigma[csoportátlag_j - x_i]^2\,)$

A variancia elemzés gondolatmenete

(angol rövidítés - analysis of variance - ANOVA)

- A minták normális eloszlásból származnak (n darab)
- Független minták
- Véletlen minták (randomizálás)
- H_0: a null hipotézis: a minták közös sokaságból/populációból származnak ($v_1=v_2=v_3=\ldots=v_n$)
- H_1: legalább egy minta egy másik sokaságból származik
- A null hipotézis következménye: $s_1^2=s_2^2=s_3^2=\ldots=s_n^2$)
azaz a varianciák homogenitási feltétele (homoszcedaszcitás)
- A mintákból két független becslést készítünk
a populáció varianciájára (σ^2)
- A két variancia becslés hányadosa az $F_{1,2}$ eloszlást követi
($F_{1,2} = s_1^2/s_2^2$)

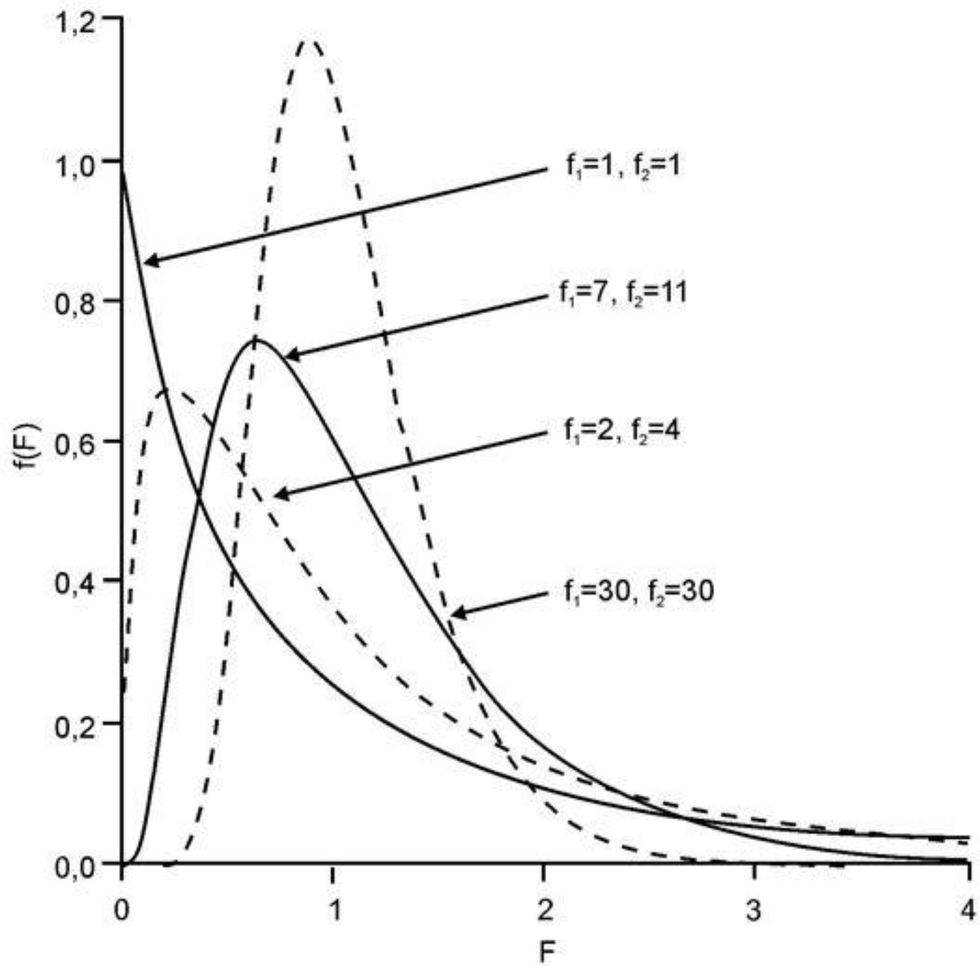
A variancia elemzés gondolatmenete (folytatás)

- Ha a minták egy sokaságból valók (a null hipotézis igaz), akkor $F_{1,2}$ eloszlásának várható értéke $v(F_{1,2}) = 1$
- Ha $p<0,05$ arra, hogy $F_{1,2} = 1$,
  akkor elvetjük a nullhipotézist
- Ha elvetettük a null hipotézist, akkor megkeressük, mely csoportokra mondhatjuk ki, hogy
  nem egy eloszlásból származnak?
- Előre tervezett (a priori), vagy utólagos (a posteriori) összehasonlításokat végzünk

Két variancia hányadosának eloszlása a Fisher–Snedecor (F) eloszlás

Normális eloszlású mintákból képzett
varianciák hányadosa

$F_{(m,n)}=s_{1(m)}{}^2/s_{2(n)}{}^2$



A több csoport elemzésének második lépése

- Van-e különböző a csoportok eredményeinek halmazában
- Ha van, akkor megkeressünk a csoportok közötti valószínű eltérés(eke)t
  - Az eltérés nem csak párok közötti különbség formájában lehet, hanem többes kontrasztokban is
  - Kontrasztra példa:
    két csoport együttesen eltér másik két csoporttól, párosával nincs szignifikáns különbség

Kontraszt – páros és többes összehasonlítás

- Ha 3 csoport átlagait m_1, m_2 és m_3 jelöli, akkor a lehetséges összehasonlítások, a null hipotézis érvényessége mellett a következők:
- $0= c_1*m_1+c_2*m_2+c_3*m_3$, ahol a $c_1+c_2+c_3=0$ teljesül
- m1 és m2 különbsége: m_1-m_2, ami kontrasztként felírva: $1*m_1-1*m_2+0*m_3$, ahol $c_1=1$, $c_2=-1$ és $c_3=0$
- (m_1 és m_2) együttes eltérése (m_3)-tól: $c_1=1$, $c_2=1$, $c_3=-2$, azaz $m_1+m_2-2*m_3$
- Ha m_1 egy fociklubhoz, m_2 másik fociklubhoz és m_3 egy kézilabdaklubhoz tartozó mérések átlagai, akkor az $m_1+m_2-2*m_3$ kontraszt adja a foci és a kézilabda összehasonlítását ($c_1=1$, $c_2=1$, $c_3=-2$, összegük nulla).

A több csoport elemzésének veszélyei

- A feltételek nem teljesülnek
- Szignifikáns ANOVA, nem találunk szignifikáns páros különbséget
- Többszörös összehasonlításra módszerválasztás tervezett vagy utólagos elemzés

Több független statisztikai döntés egy vizsgálatban?

- Mi történik az első fajú hibával, ha két teljesen független vizsgálatot végzünk?
- Ha két egymástól független hipotézisvizsgálatot végzünk, és két szignifikancia vizsgálatot, mindegyiket az **α=0,05** szinten. Miután két független vizsgálatról van szó, ezért a két szignifikancia vizsgálat is függetlennek tekinthető.
- Ha mind a két null hipotézis igaz, akkor annak valószínűsége, hogy legalább egyik nullhipotézist (hibásan) elvetjük:
    - $P(s_1)=0,05$ az első teszt esetében a fenti valószínűséget,
    - $P(s_2)=0,05$ a második teszt fenti valószínűségét.
    - $P(s_1)*P(s_2)$ a két esemény együttes előfordulásának valószínűsége, ami 0,05*0,05=**0,0025**
- A három lehetséges esemény: s_1 önmagában, s_2 önmagában, s_1 és s_2 együtt fordul elő
- A két független kísérlet esetében annak valószínűsége, hogy legalább az egyikben hibásan elvetjük a null hipotézist:
- p= 0,05+0,05-0,0025= **0,0975**, ami **lényegesen magasabb**, mint az egy szignifikancia teszt esetében elfogadott 0,05.
- ... és, ha a kísérletek és az összehasonlítások nem függetlenek?

Ha sok a csoport?

- A fenti gondolatmenet k=10 független teszt elvégzése esetén $p=1-(1-0,05)^{10}=0,4$
- A független vizsgálatok számának növelésével jelentősen növeljük annak valószínűségét, hogy olyan hatások létezését mondjuk ki, amelyek a valóságban nem léteznek (fals pozitív döntés(ek)))
- Minden lehetséges szignifikancia tesztet tekintve a tesztek nem függetlenek, noha a minták azok voltak. Példa a közös kontroll…

Ismételt páros összehasonlítások, együttes valószínűségek, az adott vizsgálati adatok tükrében

Független döntések száma	*Névleges szignifikanciaszint*	*Tényleges H0 igaz valószínűsége*	*Tényleges szignifikancia szint*
1	**0,05**	**0,950**	**0,050**
2	**0,05**	**0,903**	**0,098**
3	**0,05**	**0,857**	**0,143**
4	**0,05**	**0,815**	**0,185**
5	**0,05**	**0,774**	**0,226**
6	**0,05**	**0,735**	**0,265**
7	**0,05**	**0,698**	**0,302**
8	**0,05**	**0,663**	**0,337**
9	**0,05**	**0,630**	**0,370**
10	**0,05**	**0,599**	**0,401**
20	**0,05**	**0,358**	**0,642**
40	**0,05**	**0,129**	**0,871**

Lehetséges megoldások a többes összehasonlítások feladatára

1. Az egyedi összehasonlításokban az egyes döntésekre vonatkozó *küszöbértékeket módosítjuk* úgy, hogy
   a teljes eljárásban (egy vizsgálatban)
   az összes összehasonlításra
   együttesen érvényes küszöbérték(ek)et alkalmazunk
2. Olyan *specifikus* eljárásokat készítünk és alkalmazunk, amelyek ismert közös valószínűséggel dolgoznak

Általános eljárások többes összehasonlításokra

Ezek az eljárások függetlenek a p értéket szolgáltató eljárásoktól

- Bonferroni eljárás
  - A részdöntésenkénti szint alacsonyabb, mint a kisérletenkénti szint
  - A részdöntésekben egy közös szintet alkalmazunk mindenütt
  - $\alpha^*=1-(1-\alpha)^{1/k}$ másképen az (1-α) k-adik gyökét kell vonnunk, és az kivonni 1-ből
  - Ha a függetlenség is bizonytalan, akkor $\alpha^*=\alpha/k$ (k=5 esetben 0,01)
- Holm eljárása
  - Részdöntésenként változó szinteket alkalmazunk, rendezzük a p értékeket
  - k összehasonlítás esetén α/k, α/(k-1), α/(k-2), α/(k-3), …. , α/2, α ha k=5, akkor α/5, α/4, α/3, α/2, α elsőfajú hibát alkalmazunk

Hátrányok - előnyök,
általános vagy specifikus eljárás?

- Előny: kontrolláljuk az elsőfokú hibát az egész vizsgálatra vonatkozóan
- Hátrány: konzervatívak vagyunk a másodfokú hiba tekintetében
- Holm eljárása ugyanolyan *biztonságos* az elsőfajú hiba tekintetében, és jobb hatásfokú, mint a konzervatívabb Bonferroni eljárás
- Általános eljárás, nem használja ki az ANOVA ismert tulajdonságait
- Vannak specifikus eljárások az ANOVA-ra optimalizálva
  - Newman-Keuls, Tukey, Scheffé, Dunnett, kicsit eltérő alkalmazásokhoz optimalizálva

Az ANOVA eljárásához készített többes összehasonlító eljárások

- A különböző eljárások eltérő elvi megoldásokat alkalmaznak
- Az adott helyzetre és feladatra optimális eljárást érdemes ismerni és alkalmazni
- Példa: egy kontroll csoporthoz több csoportot hasonlítunk: optimális a Dunnett próba, legjobb, ha a kontroll csoportba nagyobb elemszámot teszünk, mint a többibe.
- A Scheffé próba a legáltalánosabb többes összehasonlító eljárás, eredménye ekvivalens az ANOVA döntéseivel, azaz szignifikáns ANOVA esetén biztos lesz szignifikáns páros vagy többes kontraszt