ANOVA
Nocions bàsiques

Exemple

Es jutja amb més benevolència a una persona segons com somriu?

136 persones, dividides a l'atzar en 4 grups de 34.

A les persones de cada grup se'ls passà un dossier on s'acusava a un home d'una falta greu i sel's demanà 5 preguntes sobre la culpabilitat de l'interfecte i el càstig.

A partir de les respostes de cada subjecte, es calculà un índex de benevolència de com havia jutjat l'acusat.

Els dossiers eren idèntics, excepte la foto de l'acusat: mateix home, però diferent tipus de somriure.

Exemple

Sincer

Fals

Compungit

Neutre

Exemple

Sincer

Fals

Compungit

Neutre

Dades a https://raw.githubusercontent.com/AprendeR-UIB/MatesIIAD/master/dades/smiles.txt

M. LaFrance, M. Hecht, *Personality and Social Psychology Bulletin* 21 (1995) 207–214

Exemple

Resultats:

Tipus de somriure			
Sincer	Fals	Compungit	Neutre
6.5	4.5	5.0	4.0
4.5	3.5	6.5	2.5
4.0	7.5	6.0	2.5
7.0	6.5	5.0	2.5
5.5	3.0	5.5	2.0
5.0	3.0	3.5	5.0
4.5	5.5	4.0	2.5
6.0	6.0	5.5	6.0
⋮	⋮	⋮	⋮

Exemple

Resultats:

Tipus de somriure			
Sincer	Fals	Compungit	Neutre
6.5	4.5	5.0	4.0
4.5	3.5	6.5	2.5
4.0	7.5	6.0	2.5
7.0	6.5	5.0	2.5
5.5	3.0	5.5	2.0
5.0	3.0	3.5	5.0
4.5	5.5	4.0	2.5
6.0	6.0	5.5	6.0
⋮	⋮	⋮	⋮

L'índex mitjà de benevolència, depèn del tipus de somriure de la foto?

Exemple

Variables d'interès

- X_s: Índex de benevolència d'una persona a qui s'ha mostrat una foto amb un somriure sincer de l'acusat
- X_f, X_c, X_n: Ídem, amb somriure fals, compungit i neutre

Siguin $\mu_s, \mu_f, \mu_c, \mu_n$ les seves mitjanes

Contrast

$$\begin{cases} H_0 : \mu_s = \mu_f = \mu_c = \mu_n \\ H_1 : \text{Hi ha somriures } i, j \text{ tals que } \mu_i \neq \mu_j \end{cases}$$

Problema general

- Tenim una variable quantitativa X mesurada sobre $k > 2$ poblacions, usualment subpoblacions d'una població definides per nivells (tractaments) de factors.

- Volem decidir si el valor mitjà de X és el mateix a totes aquestes poblacions o no

- Siguin $\mu_1, \ldots, \mu_k$ les mitjanes d'aquesta variable en aquestes poblacions. Volem fer el contrast:

$$\left\{\begin{array}{l} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 : \text{Hi ha } i, j \text{ tals que } \mu_i \neq \mu_j \end{array}\right.$$

- Prenem una mostra aleatòria de cada població, i a partir d'aquestes mostres ho decidim

Problema general

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 : \text{Hi ha } i, j \text{ tals que } \mu_i \neq \mu_j \end{cases}$$

H_1 no és

totes les mitjanes són diferents

sinó

hi ha almenys dues mitjanes que són diferents

Problema general

En capítols anteriors: Per comparar les mitjanes de dues vv.aa., calculàvem les mitjanes de dues mostres i les comparàvem

Per comparar les mitjanes de $k \geqslant 3$ vv.aa., podríem fer-ho per parelles, però hauríem de fer $\binom{k}{2}$ contrastos i augmenta la probabilitat d'error

I les hem de comparar totes amb totes, perquè pot passar que

- No poguem rebutjar que $\mu_1 = \mu_2$
- No poguem rebutjar que $\mu_2 = \mu_3$
- *Sí* que poguem rebutjar que $\mu_1 = \mu_3$

Volem un test que ens digui en un pas si totes són iguals, o si n'hi ha de diferents (en aquest darrer cas, ja cercarem les diferents si volem)

ANOVA

La tècnica més usual és l'Anàlisi de la Variància (ANOVA, de l'anglès *ANalysis Of VAriance*)

Aquesta tècnica es pot aplicar sota diferents dissenys experimentals:

- Segons quants factors empram per separar la població en subpoblacions
- Segons com triam els nivells dels factors
- Segons com prenem les mostres

Veurem els dissenys més bàsics

ANO. . . VA?

Per comparar les mitjanes de 3 o més poblacions, ens fixam en les fonts de variabilitat de les dades:

- Variabilitat total de les dades (respecte de la mitjana global)
- Variabilitat dins cada mostra (respecte de la mitjana mostral corresponent)
- Variabilitat de les mitjanes mostrals (respecte de la mitjana global)

Si les mitjanes mostrals tenen molta variabilitat, ho prendrem com a senyal que les mitjanes poblacionals no poden ser totes iguals

ANO. . . VA?

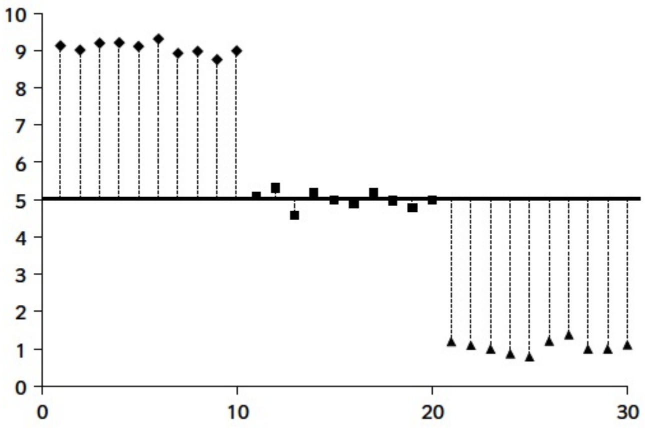
Amb definicions adients de les “variabilitats”:

$$\begin{aligned}&\text{Variabilitat total}\\&\qquad = \text{Suma de les variabilitats dins cada mostra}\\&\qquad\qquad + \text{Variabilitat de les mitjanes mostrals}\end{aligned}$$

Idea bàsica:

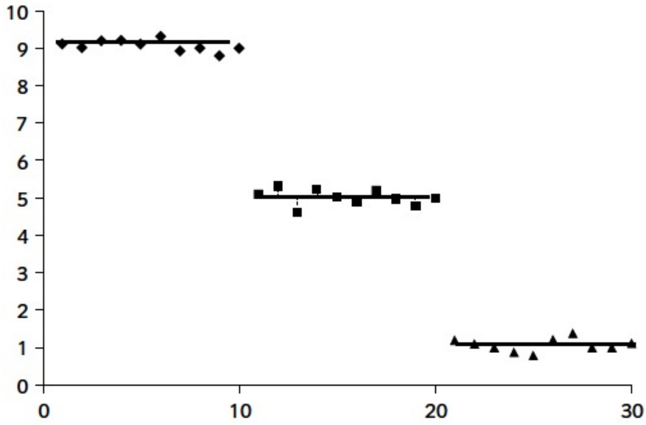
Si la variabilitat total de les dades és molt més gran que la suma de les variabilitats dins cada mostra, hi haurà una gran variabilitat de les mitjanes mostrals, i ho prendrem com a senyal que les mitjanes poblacionals no són iguals

ANO. . . VA?



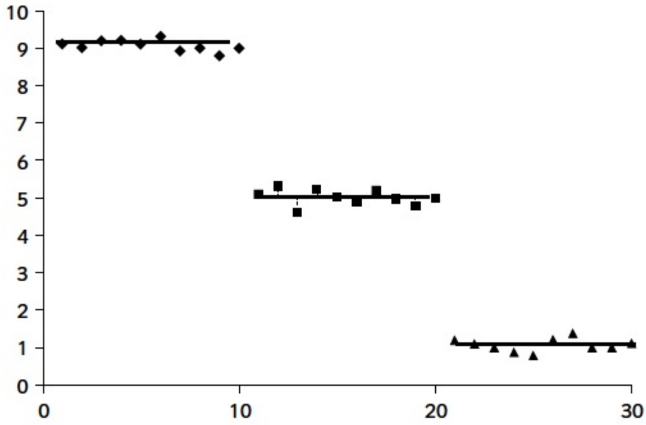
Molta variabilitat global. . .

ANO. . . VA?



. . . i molt poca dins cada nivell

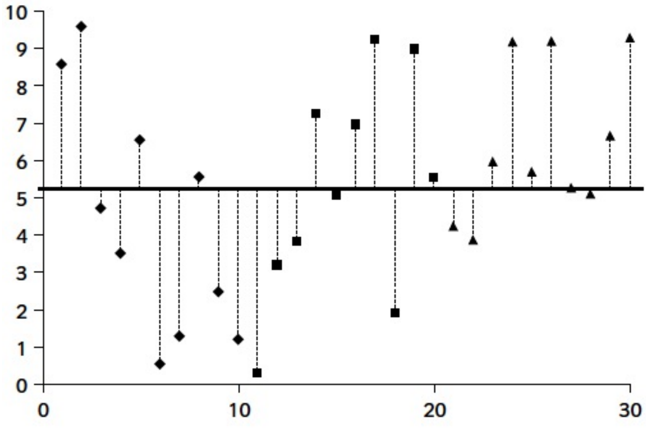
ANO. . . VA?



. . . i molt poca dins cada nivell

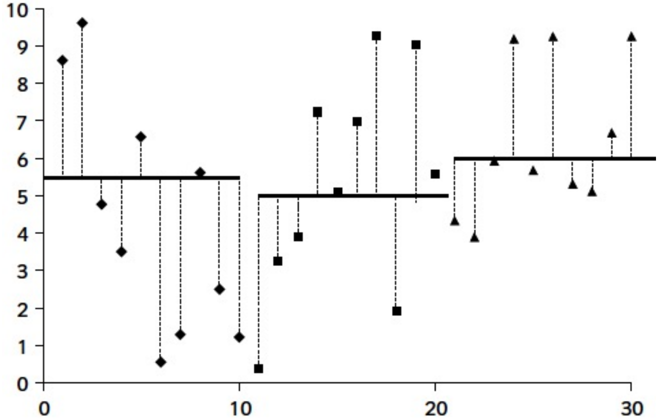
Evidència que les mitjanes són diferents

ANO...VA?



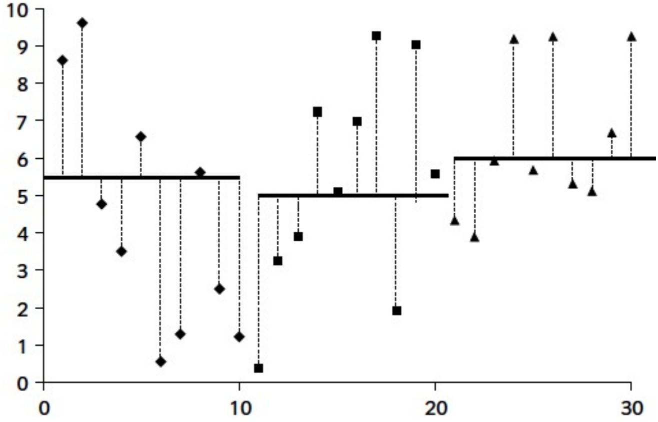
Molta variabilitat global...

ANO. . . VA?



. . . però també molta dins cada nivell

ANO. . . VA?



. . . però també molta dins cada nivell

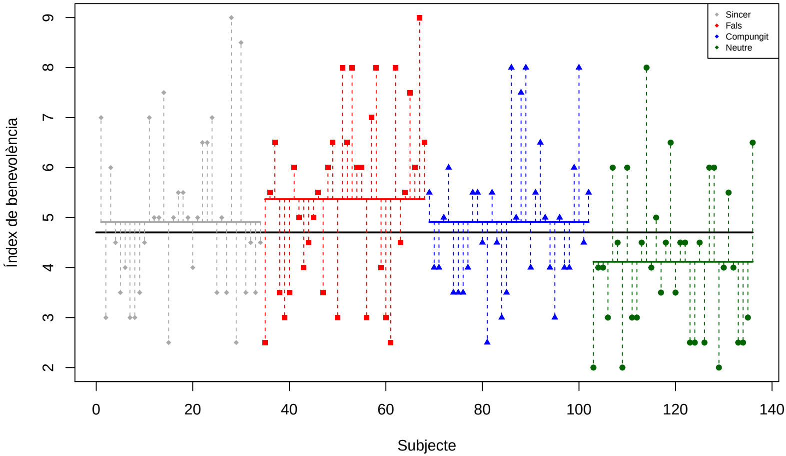
No hi ha evidència que les mitjanes siguin diferents

ANOVA

En resum: si la suma de les variabilitats dins els nivells és "suficientment més petita" que la variabilitat global, serà evidència que las mitjanes són diferents

ANOVA

En resum: si la suma de les variabilitats dins els nivells és "suficientment més petita" que la variabilitat global, serà evidència que las mitjanes són diferents



ANOVA d'1 via

Exemple

L'índex mitjà de benevolència depèn del tipus de somriure de la foto?

Variables d'interès:

- X_s: Prenc una persona a qui s'ha mostrat una foto amb un somriure sincer de l'acusat i mesur el seu índex de benevolència
- X_f, X_c, X_n: Ídem, amb somriure fals, compungit i neutre

Siguin $\mu_s, \mu_f, \mu_c, \mu_n$ les seves mitjanes

Contrast

$$\left\{ \begin{array}{l} H_0 : \mu_s = \mu_f = \mu_c = \mu_n \\ H_1 : \text{Hi ha somriures } i, j \text{ tals que } \mu_i \neq \mu_j \end{array} \right.$$

Exemple

Dades:

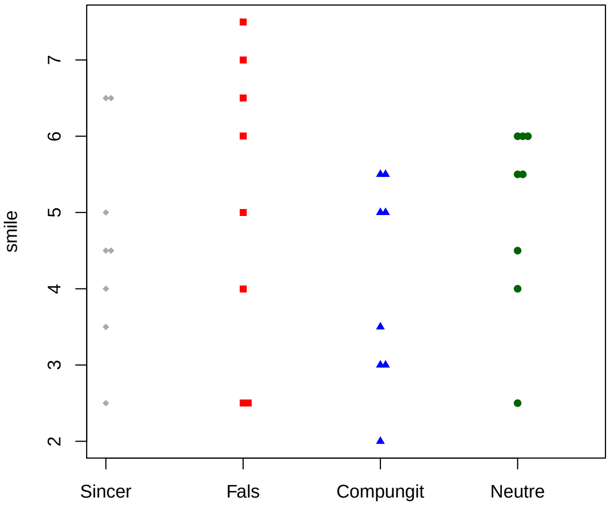
Tipus de somriure			
Sincer	Fals	Compungit	Neutre
6.5	4.5	5.0	4.0
4.5	3.5	6.5	2.5
4.0	7.5	6.0	2.5
7.0	6.5	5.0	2.5
5.5	3.0	5.5	2.0
5.0	3.0	3.5	5.0
4.5	5.5	4.0	2.5
6.0	6.0	5.5	6.0
⋮	⋮	⋮	⋮

Exemple

Una mostra de 8 resultats de cada, per fer a mà:

Tipus de somriure			
Sincer	Fals	Compungit	Neutre
6.5	4.5	5.0	4.0
4.5	3.5	6.5	2.5
4.0	7.5	6.0	2.5
7.0	6.5	5.0	2.5
5.5	3.0	5.5	2.0
5.0	3.0	3.5	5.0
4.5	5.5	4.0	2.5
6.0	6.0	5.5	6.0

Exemple



Exemple

En aquest exemple:

- Hem emprat un únic factor per classificar la població en 4 subpoblacions (el tipus de somriure)
- Hem pres una m.a.s. cada subpoblació, independents les unes de les altres

Exemple

En aquest exemple:

- Hem emprat un únic factor per classificar la població en 4 subpoblacions (el tipus de somriure)
- Hem pres una m.a.s. cada subpoblació, independents les unes de les altres

En un experiment amb disseny d'ANOVA de 1 via (*1 way*):

- Empram un únic factor per classificar la població en k subpoblacions
- Prenem una m.a.s. de cada subpoblació, independents les unes de les altres

Situació general

- μ: mitjana poblacional del conjunt de la població (ignorant els nivells)
- μ_i: mitjana poblacional dins el nivell i-èsim, $i = 1, \ldots, k$

Contrast:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 : \text{Hi ha } i, j \text{ tals que } \mu_i \neq \mu_j \end{cases}$$

Situació general

- μ: mitjana poblacional del conjunt de la població (ignorant els nivells)
- μ_i: mitjana poblacional dins el nivell i-èsim, $i = 1, \ldots, k$

Contrast:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 : \text{Hi ha } i, j \text{ tals que } \mu_i \neq \mu_j \end{cases}$$

Si $H_0 : \mu_1 = \cdots = \mu_k$ és certa, aleshores

$$\mu_1 = \cdots = \mu_k = \mu$$

Situació general

Tenim les dades en una taula:

Tractaments			
1	2	$\ldots$	k
X_{11}	X_{21}	$\cdots$	X_{k1}
X_{12}	X_{22}	$\cdots$	X_{k2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	X_{kn_k}
X_{1n_1}	$\vdots$	$\vdots$	
	X_{2n_2}		

on

- n_i: mida de la mostra del nivell i (no tenen per què ser iguals, però millor si ho són: la potència depèn del mínim)
- X_{ij}: valor de la característica sota estudi al subjecte j del nivell i

Situació general

- $N = n_1 + \cdots + n_k$: la mida total de la mostra
- $\overline{X}_i$: Mitjana mostral del nivell i-èsim, estima μ_i

$$\overline{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$$

- $\overline{X}$: Mitjana mostral de totes les dades, estima μ

$$\overline{X} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij}}{N}$$

$$\begin{array}{cccc}
\multicolumn{4}{c}{\text{Nivell del factor}} \\
\hline
1 & 2 & \cdots & k \\
\hline
X_{11} & X_{21} & \cdots & X_{k1} \\
X_{12} & X_{22} & \cdots & X_{k2} \\
\cdots & \cdots & \cdots & \cdots \\
X_{1n_1} & X_{2n_2} & \cdots & X_{kn_k} \\
\hline
\multicolumn{4}{c}{\underbrace{\overline{X}_1 \qquad \overline{X}_2 \qquad \cdots \qquad \overline{X}_k}_{\overline{X}}}
\end{array}$$

Exemple

Emmagatzemam les dades de la taula curta en un *dataframe* amb dues variables:

- `IB`: l'índex de benevolència
- `TS`: el tipus de somriure (S,F,C,N)

Tipus de somriure			
Sincer	Fals	Compungit	Neutre
6.5	4.5	5.0	4.0
4.5	3.5	6.5	2.5
4.0	7.5	6.0	2.5
7.0	6.5	5.0	2.5
5.5	3.0	5.5	2.0
5.0	3.0	3.5	5.0
4.5	5.5	4.0	2.5
6.0	6.0	5.5	6.0

Exemple

- IB: l'índex de benevolència
- TS: el tipus de somriure (S,F,C,N) com a factor

```
> IB=c(6.5,4.5,5.0,4.0,4.5,3.5,6.5,2.5,4.0,7.5,
  6.0,2.5,7.0,6.5,5.0,2.5,5.5,3.0,5.5,2.0,5.0,
  3.0,3.5,5.0,4.5,5.5,4.0,2.5,6.0,6.0,5.5,6.0)
> TS=factor(rep(c("S","F","C","N"),times=8),
          levels=c("S","F","C","N"))
> Benev=data.frame(IB,TS)
> str(Benev)
'data.frame': 32 obs. of  2 variables:
 $ IB: num  6.5 4.5 5 4 4.5 3.5 6.5 2.5 4 7.5
   ...
 $ TS: Factor w/ 4 levels "S","F","C","N": 1 2 3
    4 1 2 3 4 1 2 ...
> N=dim(Benev)[1]; k=length(levels(Benev$TS))
> c(N,k)
[1] 32 4
```

Exemple

- Les $\overline{X}_i$:

```
> Xb.i=aggregate(IB~TS,data=Benev,mean)
> Xb.i
  TS      IB
1  S  5.3750
2  F  4.9375
3  C  5.1250
4  N  3.3750
```

- La $\overline{X}$:

```
> Xb=mean(Benev$IB)
> Xb
[1] 4.703125
```

$\overline{X}_s$	$\overline{X}_f$	$\overline{X}_c$	$\overline{X}_n$	$\overline{X}$
5.375	4.9375	5.125	3.375	4.703

Condicions necessàries

Per poder fer una ANOVA d'1 via, cal que:

- Les k mostres siguin m.a.s. independents
- $N \geqslant k+1$
- Cadascuna de les k poblacions segueix una llei normal
- Homogeneïtat de les variàncies, o homocedasticitat: Totes aquestes poblacions tenen la mateixa variància σ^2

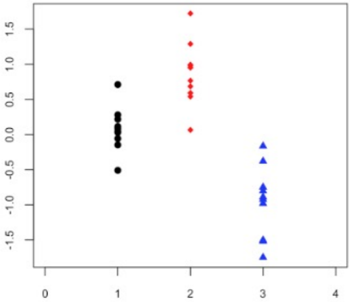
No basten mostres grans: Aquí no juga cap paper el TCL

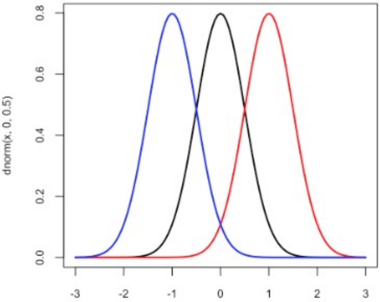
- En general, l'ANOVA és bastant robust contra falta de normalitat
- però sensible a l'heterocedasticitat

Homocedasticitat

Homocedasticitat: Totes les subpoblacions tenen la mateixa variància (**NO**: totes les mostres tenen la mateixa variància)

Homocedasticitat

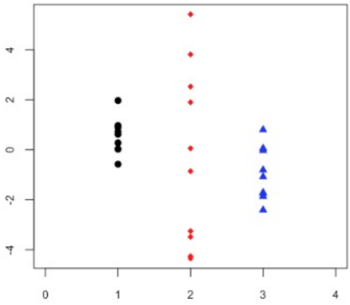


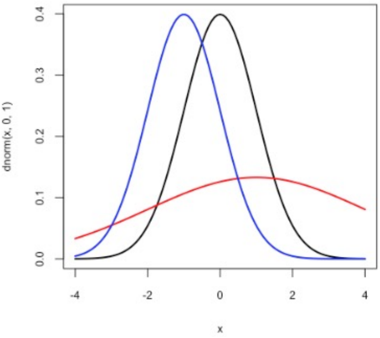


Homocedasticitat

Homocedasticitat: Totes les subpoblacions tenen la mateixa variància (**NO**: totes les mostres tenen la mateixa variància)

Heterocedasticitat





Model

$$X_i - \mu = (X_i - \mu_i) + (\mu_i - \mu), \;\; i = 1, \ldots, k,$$

on

- X_i: la variable corresponent al nivell i-èssim
- $X_i - \mu$: la desviació del valor de X_i en un subjecte respecte de la mitjana global
- $X_i - \mu_i$: la desviació del valor de X_i en un subjecte respecte de la mitjana del seu nivell
- $\mu_i - \mu$: desviació de la mitjana del nivell i-èsim respecte de la mitjana global

Identitat de les sumes de quadrats

Teorema

$SS_{Total} = SS_{Tr} + SS_E$

on:

- $SS_{Total} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2$; és la Suma Total de Quadrats i representa la variabilitat global de la mostra
- $SS_{Tr} = \sum_{i=1}^{k} n_i (\overline{X}_i - \overline{X})^2$; és la Suma de Quadrats dels Tractaments i representa la variabilitat de les mitjanes
- $SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2$; és la Suma de Quadrats dels Residus o dels Errors i representa la variabilitat dins les mostres

Identitat de les sumes de quadrats

Teorema

$SS_{Total} = SS_{Tr} + SS_E$

on:

- $SS_{Total} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2$; és la Suma Total de Quadrats i representa la variabilitat global de la mostra
- $SS_{Tr} = \sum_{i=1}^{k} n_i (\overline{X}_i - \overline{X})^2$; és la Suma de Quadrats dels Tractaments i representa la variabilitat de les mitjanes
- $SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2$; és la Suma de Quadrats dels Residus o dels Errors i representa la variabilitat dins les mostres

  La variabilitat total de la mostra descompon en la suma de les variabilitats de les mostres de cada nivell més la variabilitat de les mitjanes

Exemple

- $SS_{Total} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2$

```
> SSTotal=sum((IB-Xb)^2)
> SSTotal
[1] 69.42969
```

- $SS_{Tr} = \sum_{i=1}^{k} n_i (\overline{X}_i - \overline{X})^2$

```
> SSTr=sum(table(TS)*(Xb.i[,2]-Xb)^2)
> SSTr
[1] 19.58594
```

- $SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2$

```
> SSE=sum((IB-Xb.i[,2])^2)
> SSE
[1] 49.84375
```

Exemple

- $SS_{Total} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2$

```
> SSTotal=sum((IB-Xb)^2)
> SSTotal
[1] 69.42969
```

- $SS_{Tr} = \sum_{i=1}^{k} n_i(\overline{X}_i - \overline{X})^2$

```
> SSTr=sum(table(TS)*(Xb.i[,2]-Xb)^2)
> SSTr
[1] 19.58594
```

- $SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2$

```
> SSE=sum((IB-Xb.i[,2])^2)
> SSE
[1] 49.84375
```

```
> SSTr+SSE
[1] 69.42969
```

Contrast

$SS_{Total} = SS_{Tr} + SS_E$

Rebutjarem H_0 si SS_{Total} és prou més gran que SS_E

Contrast

$SS_{Total} = SS_{Tr} + SS_E$

Rebutjarem H_0 si SS_{Total} és prou més gran que SS_E

Rebutjarem H_0 si SS_{Tr} és prou més gran que SS_E

Contrast

$SS_{Total} = SS_{Tr} + SS_E$

Rebutjarem H_0 si SS_{Total} és prou més gran que SS_E

Rebutjarem H_0 si SS_{Tr} és prou més gran que SS_E

Per mesurar-ho emprarem els estadístics següents:

- Quadrat mitjà dels tractaments:

$$MS_{Tr} = \frac{SS_{Tr}}{k-1}$$

- Quadrat mitjà dels residus (dels errors, residual):

$$MS_E = \frac{SS_E}{N-k}$$

i compararem MS_{Tr} amb MS_E

Contrast

Si se satisfan les condicions necessàries per fer una ANOVA d'1 via:

- $E(MS_{Tr}) = \sigma^2 + \sum_{i=1}^{k} \frac{n_i(\mu_i - \mu)^2}{k-1}$
- $E(MS_E) = \sigma^2$

En particular, MS_E és estimador no esbiaixat de la variància comuna σ^2

Contrast

Si se satisfan les condicions necessàries per fer una ANOVA d'1 via:

- $E(MS_{Tr}) = \sigma^2 + \sum_{i=1}^{k} \frac{n_i(\mu_i - \mu)^2}{k-1}$
- $E(MS_E) = \sigma^2$

En particular, MS_E és estimador no esbiaixat de la variància comuna σ^2

Si $H_0 : \mu_1 = \cdots = \mu_k (= \mu)$ és certa,

$$\sum_{i=1}^{k} \frac{n_i(\mu_i - \mu)^2}{k-1} = 0,$$

i si H_0 no és certa, aquesta quantitat és > 0

Contrast

Per tant

- si H_0 és certa, $E(MS_E) = E(MS_{Tr})$ i hauríem d'esperar que aquests dos estadístics tinguessin valors propers: que

$$\frac{MS_{Tr}}{MS_E} \approx 1$$

- si H_0 és falsa, $E(MS_{Tr}) > E(MS_E)$ i hauríem d'esperar que

$$\frac{MS_{Tr}}{MS_E} > 1$$

Contrast

Prenem com a estadístic de contrast el quocient

$$F = \frac{MS_{Tr}}{MS_E}$$

Si H_0 és certa (i se satisfan les condicions per fer una ANOVA):

- la seva distribució és $F_{k-1,N-k}$ (F de Fisher-Snedecor amb $k-1$ i $N-k$ graus de llibertat)
- el seu valor serà proper a 1

A més, si $k = 2$, F és igual al quadrat de l'estadístic del test t de 2 mostres independents i variàncies iguals

Contrast

Prenem com a estadístic de contrast el quocient

$$F = \frac{MS_{Tr}}{MS_E}$$

Si H_0 és certa (i se satisfan les condicions per fer una ANOVA):

- la seva distribució és $F_{k-1,N-k}$ (F de Fisher-Snedecor amb $k-1$ i $N-k$ graus de llibertat)
- el seu valor serà proper a 1

A més, si $k = 2$, F és igual al quadrat de l'estadístic del test t de 2 mostres independents i variàncies iguals

Rebutjarem la hipòtesi nul·la si F és molt gran:

$$\text{p-valor} = P(F_{k-1,N-k} \geqslant F)$$

Contrast

- Calculam les sumes de quadrats

$$SS_{Tr},\ SS_E$$

- Calculam

$$MS_{Tr} = \frac{SS_{Tr}}{k-1},\ MS_E = \frac{SS_E}{N-k},\ F = \frac{MS_{Tr}}{MS_E}$$

- Calculam el p-valor

$$P(F_{k-1,N-k} \geqslant F)$$

- Si el p-valor és més petit que el nivell de significació α rebutjam H_0 i concloem que no totes les mitjanes són iguals. En cas contrari, acceptam H_0.

Exemple

- Ja sabem que $N = 32$, $k = 4$, $SS_{Total} = 69.43$, $SS_{Tr} = 19.586$ i $SS_E = 49.844$.

- Els quadrats mitjans són:

```
> MSTr=SSTr/(k-1); MSTr
[1] 6.528646
> MSE=SSE/(N-k); MSE
[1] 1.780134
```

Exemple

- Ja sabem que $N = 32$, $k = 4$, $SS_{Total} = 69.43$, $SS_{Tr} = 19.586$ i $SS_E = 49.844$.

- Els quadrats mitjans són:

```
> MSTr=SSTr/(k-1); MSTr
[1] 6.528646
> MSE=SSE/(N-k); MSE
[1] 1.780134
```

- L'estadístic de contrast F val:

```
> EstF=MSTr/MSE; EstF
[1] 3.667503
```

Exemple

- El p-valor $P(F_{k-1,N-k} \geqslant F)$ val

```
> 1-pf(EstF,k-1,N-k)
[1] 0.02399446
```

Exemple

- El p-valor $P(F_{k-1,N-k} \geqslant F)$ val

```
> 1-pf(EstF,k-1,N-k)
[1] 0.02399446
```

Conclusió: Hem obtingut evidència estadística que els nivells de benevolència mitjans per als diferents tipus de somriure no són tots iguals (ANOVA, p-valor 0.024)

Exemple

- El p-valor $P(F_{k-1,N-k} \geqslant F)$ val

```
> 1-pf(EstF,k-1,N-k)
[1] 0.02399446
```

Conclusió: Hem obtingut evidència estadística que els nivells de benevolència mitjans per als diferents tipus de somriure no són tots iguals (ANOVA, p-valor 0.024)

> *Alerta! Només concloem que no totes les mitjanes són iguals: no que totes les mitjanes són diferents. No és el mateix!*

Taula ANOVA

Un contrast ANOVA es resumeix en una taula ANOVA:

Origen Variació	Graus llibertat	Sumes de quadrats	Quadrats mitjans	Estadístic de contrast	p-valor
Nivells	$k-1$	SS_{Tr}	MS_{Tr}	F	p-valor
Residus	$N-k$	SS_E	MS_E		

Taula ANOVA

Un contrast ANOVA es resumeix en una taula ANOVA:

Origen Variació	Graus llibertat	Sumes de quadrats	Quadrats mitjans	Estadístic de contrast	p-valor
Nivells	$k-1$	SS_{Tr}	MS_{Tr}	F	p-valor
Residus	$N-k$	SS_E	MS_E		

Exemple:

Origen Variació	Graus llibertat	Sumes de quadrats	Quadrats mitjans	Estadístic de contrast	p-valor
Nivells	3	19.586	6.53	3.67	0.024
Residus	28	49.844	1.78		

Amb R

```
> summary(aov(IB~TS,data=Benev))
            Df Sum Sq Mean Sq F value Pr(>F)
TS           3  19.59   6.529   3.668  0.024 *
Residuals   28  49.84   1.780
------
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05
    ‘.’ 0.1 ‘ ’ 1
```

```
> summary(aov(Benev$IB~Benev$TS))
            Df Sum Sq Mean Sq F value Pr(>F)
Benev$TS     3  19.59   6.529   3.668  0.024 *
Residuals   28  49.84   1.780
---------
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05
    ‘.’ 0.1 ‘ ’ 1
```

El valor de `Pr(> F)` és el p-valor del contrast

Podíem aplicar una ANOVA d'1 via?

Per poder fer una ANOVA d'1 via cal que:

1. Les mostres de cada nivell són aleatòries simples i independents i $N \geqslant k+1$

Podíem aplicar una ANOVA d'1 via?

Per poder fer una ANOVA d'1 via cal que:

1. Les mostres de cada nivell són aleatòries simples i independents i $N \geqslant k+1$
2. La població definida per cada nivell ha de ser normal.

Podíem aplicar una ANOVA d'1 via?

Per poder fer una ANOVA d'1 via cal que:

1. Les mostres de cada nivell són aleatòries simples i independents i $N \geqslant k+1$
2. La població definida per cada nivell ha de ser normal. $\Rightarrow$ Contrastos de normalitat

Podíem aplicar una ANOVA d'1 via?

Per poder fer una ANOVA d'1 via cal que:

1. Les mostres de cada nivell són aleatòries simples i independents i $N \geqslant k+1$
2. La població definida per cada nivell ha de ser normal. $\Rightarrow$ Contrastos de normalitat
3. Totes aquestes poblacions han de tenir la mateixa variància

Podíem aplicar una ANOVA d'1 via?

Per poder fer una ANOVA d'1 via cal que:

1. Les mostres de cada nivell són aleatòries simples i independents i $N \geqslant k+1$
2. La població definida per cada nivell ha de ser normal. $\Rightarrow$ Contrastos de normalitat
3. Totes aquestes poblacions han de tenir la mateixa variància $\Rightarrow$ Contrast d'homocedasticitat

Podíem aplicar una ANOVA d'1 via?

Contrastos de normalitat: Triau el més adient en funció de les dades

```
> shapiro.test(IB[TS=="S"])$p.value
[1] 0.7357702
> shapiro.test(IB[TS=="F"])$p.value
[1] 0.4697832
> shapiro.test(IB[TS=="C"])$p.value
[1] 0.7704917
> shapiro.test(IB[TS=="N"])$p.value
[1] 0.03766128
```

Podíem aplicar una ANOVA d'1 via?

Contrast d'homocedasticitat:

- Si les vv.aa. són normals, recomanam el test de Bartlett: funció `bartlett.test`

```
> bartlett.test(IB~TS,data=Benev)

  Bartlett test of homogeneity of variances

data:  IB by TS
Bartlett's K-squared = 2.5933, df = 3, p-value =
    0.4587
```

Podem acceptar que tenen les variàncies iguals

Podíem aplicar una ANOVA d'1 via?

Contrast d'homocedasticitat:

- Si les vv.aa. no són normals, recomanam el test de Fligner-Killeen: funció `fligner.test`

```
> fligner.test(IB~TS,data=Benev)

  Fligner-Killeen test of homogeneity of
   variances

data:  IB by TS
Fligner-Killeen:med chi-squared = 3.4352,
df = 3, p-value = 0.3293
```

Podem acceptar que tenen les variàncies iguals

Contrast no paramètric

Si no podem emprar una ANOVA d'1 via, hem d'emprar un test no paramètric

El més popular es el test de Kruskal-Wallis, que generalitza el test de Mann-Whitney a més de 2 poblacions igual que ANOVA generalitza el test t

```
> kruskal.test(IB~TS,data=Benev)

  Kruskal-Wallis rank sum test

data:  IB by TS
Kruskal-Wallis chi-squared = 7.8813, df = 3,
p-value = 0.04853
```

Conclusió: Hem obtingut evidència estadística que els nivells de benevolència mitjans per als diferents tipus de somriure no són tots iguals (test de Kruskal-Wallis, p-valor 0.0485)

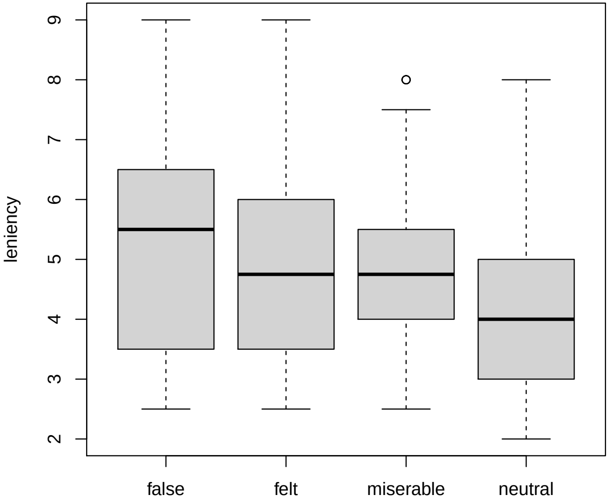
Exemple complet

Les dades completes de l'experiment del somriure són a https://raw.githubusercontent.com/AprendeR-UIB/MatesIIAD/master/dades/smiles.txt

```
> Benev.compl=read.table("https://raw.
   githubusercontent.com/AprendeR-UIB/MatesIIAD/
   master/dades/smiles.txt",header=TRUE)
> str(Benev.compl)
'data.frame': 136 obs. of  2 variables:
 $ smile   : chr  "felt" "felt" "felt" "felt"
   ...
 $ leniency: num  7 3 6 4.5 3.5 4 3 3 3.5 4.5
   ...
```

Exemple complet

```
> boxplot(leniency~smile,data=Benev.compl)
```



Exemple complet

```
> shapiro.test(Benev.compl$leniency[Benev.compl$
    smile=="felt"])$p.value
[1] 0.06464362
> shapiro.test(Benev.compl$leniency[Benev.compl$
    smile=="false"])$p.value
[1] 0.14510337
> shapiro.test(Benev.compl$leniency[Benev.compl$
    smile=="miserable"])$p.value
[1] 0.02671712
> shapiro.test(Benev.compl$leniency[Benev.compl$
    smile=="neutral"])$p.value
[1] 0.07334684
```

Exemple complet

```
> fligner.test(leniency~smile,data=Benev.compl)

	Fligner-Killeen test of homogeneity of
	 variances

data:  leniency by smile
Fligner-Killeen:med chi-squared = 3.1294,
df = 3, p-value = 0.3721
```

Exemple complet

```
> summary(aov(leniency~smile,data=Benev.compl))
             Df Sum Sq Mean Sq F value Pr(>F)
smile         3   27.5   9.178   3.465 0.0182 *
Residuals   132  349.7   2.649
```

Conclusió: Hem obtingut evidència estadística que els nivells de benevolència mitjans per als diferents tipus de somriure no són tots iguals (ANOVA, p-valor 0.018)

Exemple complet

```
> kruskal.test(leniency~smile,data=Benev.compl)

  Kruskal-Wallis rank sum test

data:  leniency by smile
Kruskal-Wallis chi-squared = 9.1747, df = 3,
p-value = 0.02706
```

Conclusió: Hem obtingut evidència estadística que els nivells de benevolència mitjans per als diferents tipus de somriure no són tots iguals (test de Kruskal-Wallis, p-valor 0.027)

I ara què?

I ara què?

Si hem rebutjat la hipòtesi nul·la $H_0 : \mu_1 = \cdots = \mu_k$, podem demanar-nos quins són els nivells diferents

Es pot fer de diverses maneres, aquí veurem la més “òbvia”: comparar totes les parelles de mitjanes per mitjà de tests t

Comparacions posteriors per parelles

Es realitzen els $\binom{k}{2}$ contrastos de 2 mitjanes amb mostres independents i variàncies iguals

$$\left.\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_1 & : \mu_i \neq \mu_j \end{array}\right\}$$

Comparacions posteriors per parelles

Es realitzen els $\binom{k}{2}$ contrastos de 2 mitjanes amb mostres independents i variàncies iguals

$$\left.\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_1 & : \mu_i \neq \mu_j \end{array}\right\}$$

Però ara no empram l'estadístic

$$T_{i,j} = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right) \cdot \frac{(n_i-1)\widetilde{S}_i^2 + (n_j-1)\widetilde{S}_j^2}{n_i+n_j-2}}}$$

que, si $\mu_i = \mu_j$, té distribució $t_{n_i+n_j-2}$

Comparacions posteriors per parelles

Es realitzen els $\binom{k}{2}$ contrastos de 2 mitjanes amb mostres independents i variàncies iguals

$$\left.\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_1 & : \mu_i \neq \mu_j \end{array}\right\}$$

Ara empram l'estadístic

$$T_{i,j} = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right) \cdot MS_E}}$$

que, $\mu_i = \mu_j$, té distribució t_{N-k}

Comparacions posteriors per parelles

Es realitzen els $\binom{k}{2}$ contrastos de 2 mitjanes amb mostres independents i variàncies iguals

$$\left.\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_1 & : \mu_i \neq \mu_j \end{array}\right\}$$

Ara empram l'estadístic

$$T_{i,j} = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right) \cdot MS_E}}$$

que, $\mu_i = \mu_j$, té distribució t_{N-k}

El p-valor de cada contrast és $2P(t_{N-k} \geqslant |T_{i,j,0}|)$, on $T_{i,j,0}$ és el valor que pren $T_{i,j}$ a la nostra mostra

Alerta!

Si es realitzen c contrastos a un nivell de significació α, la probabilitat d'algun Error de Tipus I és més gran que α: de fet, és $1-(1-\alpha)^c$

Al nostre exemple, si realitzam els 6 contrastos

$$\mu_s \stackrel{?}{=} \mu_f,\ \mu_s \stackrel{?}{=} \mu_c,\ \mu_s \stackrel{?}{=} \mu_n,\ \mu_f \stackrel{?}{=} \mu_c,\ \mu_f \stackrel{?}{=} \mu_n,\ \mu_c \stackrel{?}{=} \mu_n$$

amb nivell de significació $\alpha = 0.05$, la probabilitat d'Error de Tipus I a qualcun és $1-(1-0.05)^6 \approx 0.265$.

Alerta!

Si es realitzen c contrastos a un nivell de significació α, la probabilitat d'algun Error de Tipus I és més gran que α: de fet, és $1-(1-\alpha)^c$

Al nostre exemple, si realitzam els 6 contrastos

$$\mu_s \stackrel{?}{=} \mu_f,\ \mu_s \stackrel{?}{=} \mu_c,\ \mu_s \stackrel{?}{=} \mu_n,\ \mu_f \stackrel{?}{=} \mu_c,\ \mu_f \stackrel{?}{=} \mu_n,\ \mu_c \stackrel{?}{=} \mu_n$$

amb nivell de significació $\alpha = 0.05$, la probabilitat d'Error de Tipus I a qualcun és $1-(1-0.05)^6 \approx 0.265$.

Haurem de reduir el nivell de significació de cada contrast perquè la probabilitat de cometre algun Error de Tipus I sigui α

O, equivalentement, augmentar (ajustar) el p-valor de cada contrast i comparar-lo amb l'α fixat

Ajust de Bonferroni

És el més popular

Emprant que $1-(1-x)^c \leqslant c \cdot x$, si volem efectuar c contrastos amb nivell de significació (global) α,

- realitzam cada contrast amb nivell de significació α/c (i així el nivell de significació global serà $\leqslant c \cdot \alpha/c = \alpha$)

  o equivalentment

- multiplicam el p-valor de cada contrast per c abans de comparar-lo amb el nivell de significació α

$$p < \alpha/c \Longleftrightarrow c \cdot p < \alpha$$

Ajust de Bonferroni

És el més popular

Emprant que $1 - (1 - x)^c \leqslant c \cdot x$, si volem efectuar c contrastos amb nivell de significació (global) α,

- realitzam cada contrast amb nivell de significació α/c (i així el nivell de significació global serà $\leqslant c \cdot \alpha/c = \alpha$)

  o equivalentment

- multiplicam el p-valor de cada contrast per c abans de comparar-lo amb el nivell de significació α

$$p < \alpha/c \Longleftrightarrow c \cdot p < \alpha$$

Atenció! Els p-valors volen ser probabilitats: si en ajustar-lo passa d'1, s'ha de donar com a p-valor ajustat 1

Ajust de Bonferroni

Al nostre exemple, si realitzam els 6 contrastos, per obtenir un nivell de significació global $\alpha = 0.05$,

- efectuam cada contrast amb nivell de significació $0.05/6 = 0.0083$

  o

- multiplicam cada p-valor per 6

Ajust de Bonferroni

Aplicam la funció `pairwise.t.test` a la variable numérica i el factor, en aquest ordre, i amb el paràmetre `p.adjust.method` per indicar el mètode d'ajust.

Per ara

- `p.adjust.method="none"`: no els ajusta
- `p.adjust.method="bonferroni"`: ajust de Bonferroni

Ajust de Bonferroni

P-valors “reals”, els hauríem de comparar amb $\alpha/6$

```
> pairwise.t.test(Benev$IB, Benev$TS,
  p.adjust.method="none")

Pairwise comparisons using t tests with pooled
    SD

data:  Benev$IB and Benev$TS

  S      F      C
F 0.5173 -      -
C 0.7107 0.7807 -
N 0.0056 0.0265 0.0139

P value adjustment method: none
```

L’únic p-valor per davall de $0.05/6 = 0.0083$ és el de (S,N)

Ajust de Bonferroni

P-valors ajustats segons Bonferroni, els hauríem de comparar amb α

```
> pairwise.t.test(Benev$IB, Benev$TS,
  p.adjust.method="bonferroni")

Pairwise comparisons using t tests with pooled
    SD

data:  Benev$IB and Benev$TS

  S     F     C
F 1.000 -     -
C 1.000 1.000 -
N 0.034 0.159 0.084

P value adjustment method: bonferroni
```

L'únic p-valor per davall de 0.05 és el de (S,N)

Conclusió

Hem obtingut evidència estadística que els índexos mitjans de benevolència quan la foto mostra un somriure sincer i quan la foto és neutra són diferents, i no podem rebutjar la igualtat de cap altra parella d'índexos de benevolència mitjans (ANOVA d'1 via, test posterior per parelles de Bonferroni).

Sincer

Fals

Compungit

Neutre

Ajust de Holm

És més potent que Bonferroni, i el mètode per defecte amb R

1. Siguin $C_1, \ldots, C_c$ els contrastos i $P_1, \ldots, P_c$ els p-valors corresponents
2. Ordenam aquests p-valors en ordre creixent $P_{(1)} \leqslant \cdots \leqslant P_{(c)}$ i reenumeram consistentment els contrastos $C_{(1)}, \ldots, C_{(c)}$
3. Per a cada $j = 1, \ldots, c$, calculam el p-valor ajustat $\widetilde{P}_{(j)} = (c + 1 - j)P_{(j)}$
4. Aleshores rebutjam la hipòtesi nul·la als contrastos $C_{(j)}$ on $\widetilde{P}_{(j)} < \alpha$

Ajust de Holm

Fem a mà el mètode de Holm al nostre exemple:

1) Taula amb els p-valors dels contrastos

Contrast	p-valor
S-F	0.5173
S-C	0.7107
S-N	0.0056
F-C	0.7807
F-N	0.0265
C-N	0.0139

Ajust de Holm

2) Ordenam en ordre creixent del p-valor

Contrast	p-valor
S-N	0.0056
C-N	0.0139
F-N	0.0265
S-F	0.5173
S-C	0.7107
F-C	0.7807

Ajust de Holm

3) Ajustam: multiplicam el j-èsim p-valor per $6+1-j$

Contrast	p-valor	p-valor ajustat
S-N	0.0056	$6 \times 0.0056 = 0.0336$
C-N	0.0139	$5 \times 0.0139 = 0.0695$
F-N	0.0265	$4 \times 0.0265 = 0.1060$
S-F	0.5173	$3 \times 0.5173 = 1.5519$
S-C	0.7107	$2 \times 0.7107 = 1.4214$
F-C	0.7807	$1 \times 0.7807 = 0.7807$

Ajust de Holm

3) Ajustam: multiplicam el j-èsim p-valor per $6 + 1 - j$

Contrast	p-valor	p-valor ajustat
S-N	0.0056	$6 \times 0.0056 = 0.0336$
C-N	0.0139	$5 \times 0.0139 = 0.0695$
F-N	0.0265	$4 \times 0.0265 = 0.1060$
S-F	0.5173	$3 \times 0.5173 = 1$
S-C	0.7107	$2 \times 0.7107 = 1$
F-C	0.7807	$1 \times 0.7807 = 0.7807$

Ajust de Holm

4) Miram els p-valors per davall del nivell de significació α

Contrast	p-valor	p-valor ajustat
S-N	0.0056	$6 \times 0.0056 = 0.0336$
C-N	0.0139	$5 \times 0.0139 = 0.0695$
F-N	0.0265	$4 \times 0.0265 = 0.1060$
S-F	0.5173	$3 \times 0.5173 = 1$
S-C	0.7107	$2 \times 0.7107 = 1$
F-C	0.7807	$1 \times 0.7807 = 0.7807$

Mateixa conclusió que amb Bonferroni: concloem que $\mu_S \neq \mu_N$ i no podem rebutjar que les altres parelles de mitjanes siguin iguals

Ajust de Holm

```
> pairwise.t.test(Benev$IB,Benev$TS)
  #o p.adjust.method="holm"

  Pairwise comparisons using t tests with pooled
    SD

data:  Benev$IB and Benev$TS

  S     F     C
F 1.000 -     -
C 1.000 1.000 -
N 0.034 0.106 0.070
```

Tests posteriors no paramètrics

Si no té sentit emprar tests t amb variàncies iguals, podem fer contrastos per parelles de Mann-Whitney amb

```
pairwise.wilcox.test(..., paired=FALSE,
   p.adjust.method=...)
```

Tests posteriors no paramètrics

Si no té sentit emprar tests t amb variàncies iguals, podem fer contrastos per parelles de Mann-Whitney amb

```
pairwise.wilcox.test(..., paired=FALSE,
   p.adjust.method=...)
```

```
> pairwise.wilcox.test(Benev$IB,Benev$TS,
  paired=FALSE, p.adjust.method="bonferroni")

  Pairwise comparisons using Wilcoxon rank sum
   test with continuity correction

data:  Benev$IB and Benev$TS

  S    F    C
F 1.00 -    -
C 1.00 1.00 -
N 0.12 0.26 0.18

P value adjustment method: bonferroni
```

El problema de les comparacions múltiples

```
> shapiro.test(IB[TS=="S"])$p.value
[1] 0.7357702
> shapiro.test(IB[TS=="F"])$p.value
[1] 0.4697832
> shapiro.test(IB[TS=="C"])$p.value
[1] 0.7704917
> shapiro.test(IB[TS=="N"])$p.value
[1] 0.03766128
```

El problema de les comparacions múltiples

```
> shapiro.test(IB[TS=="S"])$p.value
[1] 0.7357702
> shapiro.test(IB[TS=="F"])$p.value
[1] 0.4697832
> shapiro.test(IB[TS=="C"])$p.value
[1] 0.7704917
> shapiro.test(IB[TS=="N"])$p.value
[1] 0.03766128

> round(4*c(0.7357702,0.4697832,0.7704917,
  0.03766128),3)
[1] 2.943 1.879 3.082 0.151
```

El problema de les comparacions múltiples

```
> shapiro.test(IB[TS=="S"])$p.value
[1] 0.7357702
> shapiro.test(IB[TS=="F"])$p.value
[1] 0.4697832
> shapiro.test(IB[TS=="C"])$p.value
[1] 0.7704917
> shapiro.test(IB[TS=="N"])$p.value
[1] 0.03766128

> round(4*c(0.7357702,0.4697832,0.7704917,
  0.03766128),3)
[1] 2.943 1.879 3.082 0.151

> round(p.adjust(
  c(0.7357702,0.4697832,0.7704917,0.03766128),
  method="bonferroni"),3)
[1] 1.000 1.000 1.000 0.151
```

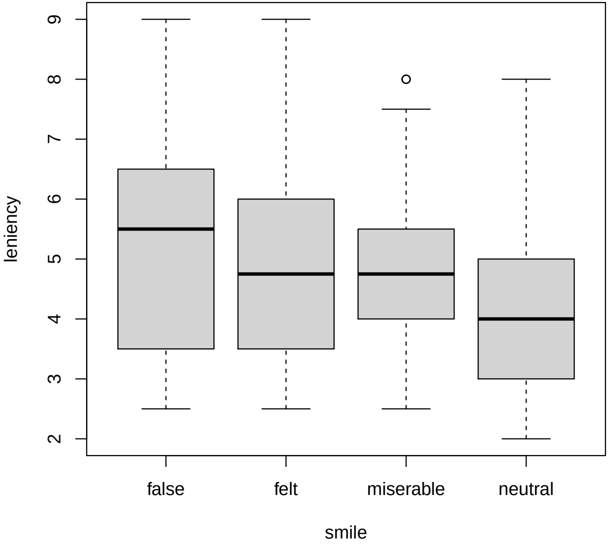
Ajustant per Bonferroni, acceptam que les 4 mostres s'ajusten a variables normals

Exemple complet

Les dades completes de l'experiment del somriure són a https://raw.githubusercontent.com/AprendeR-UIB/MatesIIAD/master/dades/smiles.txt

```
> Benev.compl=read.table("https://raw.
    githubusercontent.com/AprendeR-UIB/MatesIIAD/
    master/dades/smiles.txt",header=TRUE)
> str(Benev.compl)
'data.frame':	136 obs. of  2 variables:
 $ smile   : chr  "felt" "felt" "felt" "felt"
    ...
 $ leniency: num  7 3 6 4.5 3.5 4 3 3 3.5 4.5
    ...
> boxplot(leniency~smile,data=Benev.compl)
```

Exemple complet



Exemple complet

```
> pairwise.t.test(Benev.compl$leniency,Benev.
   compl$smile,p.adjust.method="bonferroni")

  Pairwise comparisons using t tests with pooled
    SD

data:  Benev.compl$leniency and Benev.compl$
   smile

          false felt  miserable
felt      1.000 -     -
miserable 1.000 1.000 -
neutral   0.011 0.278 0.278

P value adjustment method: bonferroni
```

Exemple complet

```
> pairwise.t.test(Benev.compl$leniency,Benev.
    compl$smile)

  Pairwise comparisons using t tests with pooled
    SD

data:  Benev.compl$leniency and Benev.compl$
    smile

          false felt  miserable
felt      0.751 -     -
miserable 0.751 1.000 -
neutral   0.011 0.231 0.231

P value adjustment method: holm
```

Exemple complet

```
> pairwise.wilcox.test(Benev.compl$leniency,
    Benev.compl$smile)

  Pairwise comparisons using Wilcoxon rank sum
    test with continuity correction

data:  Benev.compl$leniency and Benev.compl$
    smile

          false felt  miserable
felt      0.802 -     -
miserable 0.802 0.805 -
neutral   0.032 0.209 0.209

P value adjustment method: holm
Warning messages:
1: In wilcox.test.default(xi, xj, paired =
    paired, ...) :
  cannot compute exact p-value with ties
```