UNITATS 12, 13 i 14: COEFICIENTS DE CORRELACIÓ I INTRODUCCIÓ ALS MODELS ESTADÍSTICS.

Situació 1

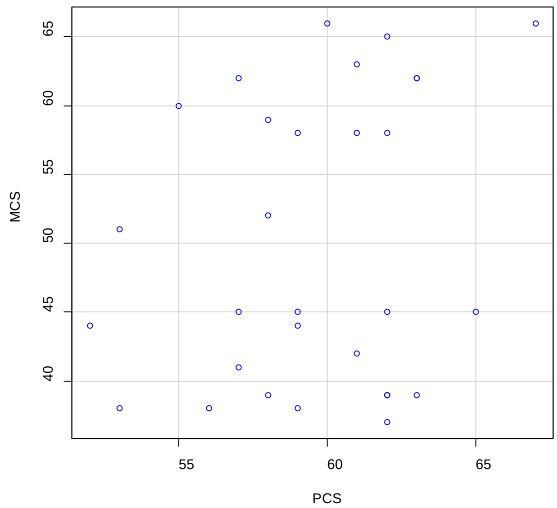
1. ¿Existeix relació estadísticament significativa entre el component físic i el component mental de l'escala de QVRS?

1er. Plantejament hipòtesi nul·la:

H_0: $\rho_{PCS\text{-}MCS} = 0$

2n. Càlcul estadístic de contrasts: Coeficient de correlació lineal de *Pearson*

Es vol determinar si existeix relació entre dos variables mesurades en escala d'interval, per tant es pot utilitzar el coeficient de correlació lineal de *Pearson*. El primer pas, consistirà en construir el núvol de punt, per determinar visualment si la relació pot ser lineal:



A partir del núvol de punts, dona la sensació que no existeix relació entre les dues variables, de totes maneres quantificarem la possible relació lineal a partir del coeficient de correlació lineal de *Pearson* tal com es mostra a continuació:

ID	PCS	MCS	(PCS-$\bar{x}$)	(MCS-$\bar{y}$)	(PCS-$\bar{x}$)^2	(MCS-$\bar{y}$)^2	(PCS-$\bar{x}$)·(MCS-$\bar{y}$)
1	65	45	5,47	-5	29,88	25	-27,33
2	63	39	3,47	-11	12,02	121	-38,13
3	59	38	-0,53	-12	0,28	144	6,40
4	58	52	-1,53	2	2,35	4	-3,07
5	62	65	2,47	15	6,08	225	37,00

ID	PCS	MCS	(PCS-$\bar{x}$)	(MCS-$\bar{y}$)	(PCS-$\bar{x}$)^2	(MCS-$\bar{y}$)^2	(PCS-$\bar{x}$)·(MCS-$\bar{y}$)
6	57	45	-2,53	-5	6,42	25	12,67
7	53	51	-6,53	1	42,68	1	-6,53
8	52	44	-7,53	-6	56,75	36	45,20
9	62	39	2,47	-11	6,08	121	-27,13
10	63	62	3,47	12	12,02	144	41,60
11	61	63	1,47	13	2,15	169	19,07
12	62	45	2,47	-5	6,08	25	-12,33
13	55	60	-4,53	10	20,55	100	-45,33
14	62	37	2,47	-13	6,08	169	-32,07
15	59	44	-0,53	-6	0,28	36	3,20
16	58	59	-1,53	9	2,35	81	-13,80
17	53	38	-6,53	-12	42,68	144	78,40
18	57	41	-2,53	-9	6,42	81	22,80
19	61	58	1,47	8	2,15	64	11,73
20	57	62	-2,53	12	6,42	144	-30,40
21	62	39	2,47	-11	6,08	121	-27,13
22	56	38	-3,53	-12	12,48	144	42,40
23	59	58	-0,53	8	0,28	64	-4,27
24	67	66	7,47	16	55,75	256	119,47
25	63	62	3,47	12	12,02	144	41,60
26	59	45	-0,53	-5	0,28	25	2,67
27	61	42	1,47	-8	2,15	64	-11,73
28	58	39	-1,53	-11	2,35	121	16,87
29	62	58	2,47	8	6,08	64	19,73
30	60	66	0,47	16	0,22	256	7,47
Total	1786	1500	0,00	0	367,47	3118	249,00

PCS: $\bar{x}$ = 59,53; $s = \sqrt{367{,}47/29} = 3{,}56$

MCS: $\bar{y}$ = 50; $s = \sqrt{3118/29} = 10{,}37$

$$s_{PCS-MCS} = 249{,}00/29 = 8{,}59$$

$$r_{PCS-MCS} = \frac{s_{PCS-MCS}}{s_{PCS} \cdot s_{MCS}} = \frac{8{,}59}{3{,}56 \cdot 10{,}37} = 0{,}232623$$

3r. Prendre Decisió

Per prendre la decisió estadística podem transformar el valor del coeficient de correlació de *Pearson* a un valor de F de *Snedecor* i utilitzar aquest model de probababilitat per trobar el valor de significació associat:

$$F = \frac{r^2}{(1 - r^2)/(n-2)} = \frac{0{,}054113}{(1 - 0{,}05413)/28} = 1{,}601856$$

$\nu_{num} = 1$; $\nu_{den} = n - 2 = 30 - 2 = 28$

Llistat R:

```
> pf(c(1.601856), df1=1, df2=28, lower.tail=FALSE)
[1] 0.2160709
```

$P_{bilateral} = 0{,}2160709 \rightarrow$ No es rebutja H_0

4rt. Conclusions

A partir de les nostres dades no es pot afirmar que existeixi relació lineal estadísticament significativa entre les puntuacions en el component PCS i el component MCS de l'escala QVRS, donat el coeficient de correlació de *Pearson* té un valor de 0,232623 que transformat a *F* té un valor de 1,601856 que amb 1 grau de llibertat del numerador i 28 del denominador té un grau de significació de 0,2160709, valor lo suficientment gran com per a no rebutjar la hipòtesi nul·la.

Llistats R:

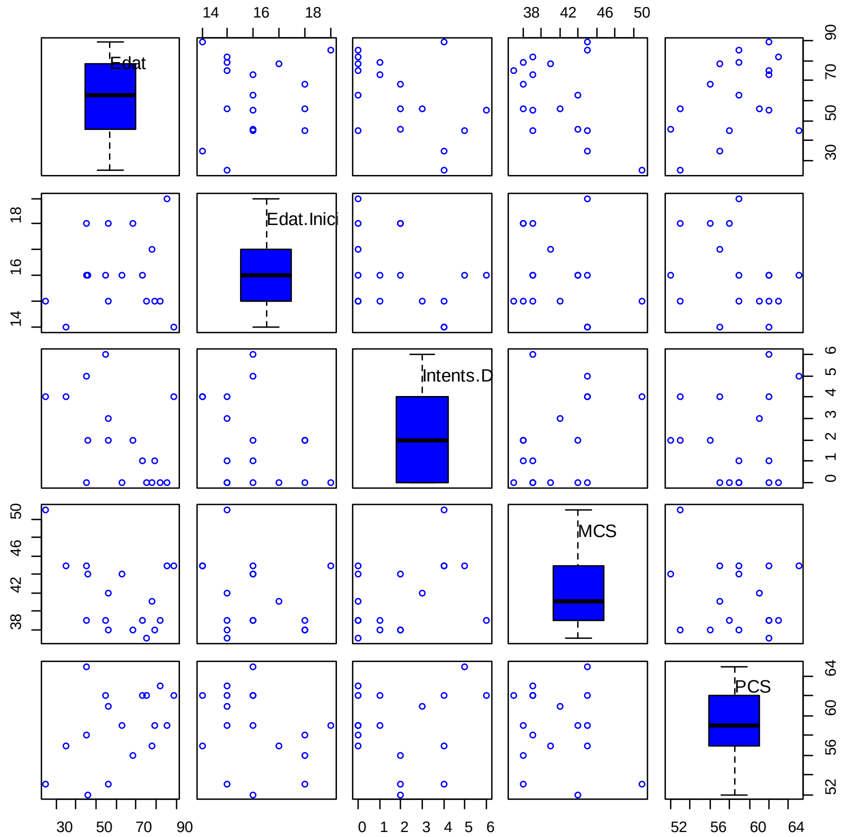
```
> with(Situació1, cor.test(MCS, PCS, alternative="two.sided", method="pearson"))

  Pearson's product-moment correlation

data:  MCS and PCS
t = 1.2656, df = 28, p-value = 0.2161
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1393226  0.5470459
sample estimates:
      cor
0.2326228
```

2. Estableixi la matriu de correlacions de totes les variables quantitatives mesurades com a mínim en escala d'interval de la nostra base de dades. Determini entre quines variables existeix relació estadísticament significativa.

1er. Plantejament hipòtesi nul·la:

H_0: $\rho_{xy} = 0$

2n. Càlcul estadístic de contrasts: Coeficient de correlació lineal de *Pearson*



En els diferents núvols de punts de les relacions dos a dos entre les cinc variables mesurades com a mínim en escala d'interval no sembla detectar-se un patró de relació lineal en cap de les 10 relacions, tot i així obtenim la matriu de correlacions lineal de *Pearson* per quantificar la possible relació:

Llistat R:

```
> 
rcorr.adjust(Situació1[,c("Edat","Edat.Inici.Fumar","Intents.Deixar","MCS","PCS")]
, type="pearson", use="pairwise.complete")

 Pearson correlations:
                   Edat Edat.Inici.Fumar Intents.Deixar     MCS     PCS
Edat             1.0000           0.0995        -0.5002 -0.2882  0.1918
Edat.Inici.Fumar 0.0995           1.0000        -0.3806 -0.2514 -0.2989
Intents.Deixar  -0.5002          -0.3806         1.0000  0.3926  0.0906
MCS             -0.2882          -0.2514         0.3926  1.0000  0.2326
PCS              0.1918          -0.2989         0.0906  0.2326  1.0000

 Number of observations:
                 Edat Edat.Inici.Fumar Intents.Deixar MCS PCS
Edat               30               17             17  30  30
Edat.Inici.Fumar   17               17             17  17  17
Intents.Deixar     17               17             17  17  17
MCS                30               17             17  30  30
PCS                30               17             17  30  30

 Pairwise two-sided p-values:
                 Edat   Edat.Inici.Fumar Intents.Deixar MCS    PCS
Edat                    0.7040           0.0409         0.1225 0.3098
Edat.Inici.Fumar 0.7040                  0.1317         0.3305 0.2438
```

```
Intents.Deixar   0.0409 0.1317                        0.1190 0.7296
MCS              0.1225 0.3305           0.1190              0.2161
PCS              0.3098 0.2438           0.7296        0.2161

 Adjusted p-values (Holm's method)
                 Edat   Edat.Inici.Fumar Intents.Deixar MCS    PCS
Edat                    1.0000           0.4087         1.0000 1.0000
Edat.Inici.Fumar 1.0000                  1.0000         1.0000 1.0000
Intents.Deixar   0.4087 1.0000                          1.0000 1.0000
MCS              1.0000 1.0000           1.0000                1.0000
PCS              1.0000 1.0000           1.0000         1.0000
```

3r. Prendre Decisió

Com es pot veure en el llistat de R el grau de significació corresponent a la correlació entre edat i intens deixar de fumar és de 0,0409, que ens pot portar a rebutjar la hipòtesi nul·la, tota la resta són superiors, oscil·len entre 0,1190 i 0,7296 i per tant ens porten al no rebuig de la hipòtesi nul·la. De totes maneres, si tenim en compte el grau de significació un cop aplicat el mètode de *Holm* per corregir per múltiples contrastos, per cap dels 10 coeficients de correlació s'arriba a la significació estadística (graus de significació entre 0,4087 i 1, el que ens porta a no rebutjar la hipòtesi nul·la).

4rt. Conclusions

Com es pot veure en el llistat de R, cap dels 10 coeficients de correlació de *Pearson* arriba a la significació estadística un cop corregit pel mètode de *Holm* el grau de significació, ja que la p més petita és de 0,4087 corresponent al coeficient de correlació entre l'edat i els intents de deixar de fumar. Per tant, podem afirmar que cap dels 10 coeficients de correlació de *Pearson* obtinguts arriba a la significació estadística i per tant no hi ha relació lineal entre les dues variables implicades.

3. Determini quina és la millor variable per predir els valors en el component mental de l'escala de QVRS, raoni la seva resposta. Construeixi el model de regressió, valori la seva validesa i en cas de que el model sigui adequat faci la predicció puntual i per interval d'una persona que té un valor de 60 en aquesta variable predictiva.

Com acabem de veure en l'exercici 2, cap de les variables quantitatives mesurades com a mínim en escala d'interval està estadísticament relacionada amb la puntuació en el component mental de l'escala QVRS, per tant cap d'aquestes variables serà una bona predictora i no podem construir cap model de regressió a partir de les variables de que es disposa.

Situació 2

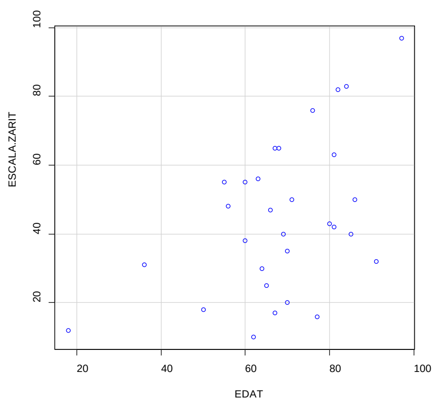
4. ¿Existeix relació estadísticament significativa entre l'edat dels cuidadors familiars i la seva puntuació en l'escala Zarit per la mesura del dol anticipat?

1er. Plantejament hipòtesi nul·la:

H_0: $\rho_{\text{Edat-Zarit}} = 0$

2n. Càlcul estadístic de contrasts: Coeficient de correlació lineal de *Pearson*

El primer pas, consistirà en construir el núvol de punt, per determinar visualment si la relació pot ser lineal:



A partir del núvol de punts, dona la sensació que pot existir una relació lineal entre les dues variables, per tant quantificarem aquesta relació a partir del coeficient de correlació lineal de *Pearson* tal com es mostra a continuació:

Llistat R:

```
> with(Situació2, cor.test(EDAT, ESCALA.ZARIT, alternative="two.sided",
method="pearson"))

  Pearson's product-moment correlation

data:  EDAT and ESCALA.ZARIT
t = 2.9949, df = 28, p-value = 0.005689
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1608355 0.7243015
sample estimates:
      cor
0.4925636
```

Com es pot veure en el llistat de R, el valor del coeficient de correlació de *Pearson* entre les dues variables és de 0,4925636.

3r. Prendre Decisió

$P_{\text{bilateral}} = 0{,}0056689 \rightarrow$ Es rebutja H_0

4rt. Grandària de l'efecte: coeficient de determinació

$r^2 = 0,4925636^2 = 0,242619$

5è. Conclusions

Com es pot veure en el llistat de R, un cop transformat el coeficient de correlació de *Pearson*, 0,4925636, a un valor de t de Student (2,9949) amb 28 graus de llibertat té un grau de significació bilateral associat de 0,005689, valor lo suficientment petit com per a rebutjar la hipòtesi nul·la, i per tant podem afirmar que existeix relació lineal directa estadísticament significativa entre l'edat dels cuidadors i la seva puntuació en l'escala de dol anticipat de Zarit. A més a més, podem afirmar que la intensitat de la relació és mitjana, donat que el valor del coeficient de determinació és igual a 0,242619.

5. ¿Les nostres dades confirmen la hipòtesi que existeix relació entre les puntuacions en l'escala Zarit per la mesura del dol anticipat i la puntuació en l'escala de severitat de l'estat del malalt (GDS)?

1er. Plantejament hipòtesi nul·la:

H_0: $\rho_{Zarit\text{-}GDS} = 0$

2n. Càlcul estadístic de contrasts: Coeficient de correlació de *Spearman*

Com que la puntuació en l'escala de severitat de l'estat del malalt (GDS) és una variable ordinal, per determinar si existeix relació entre aquestes puntuacions i les puntuacions en l'escala de dol anticipat de Zarit obtindrem el coeficient de correlació de *Spearman*.

Llistat R:

```
> with(Situació2, cor.test(ESCALA.ZARIT, GDS, alternative="two.sided", method="spearman"))

  Spearman's rank correlation rho

data:  ESCALA.ZARIT and GDS
S = 5477.5, p-value = 0.2459
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.2185687
```

Com es pot veure en el llistat de R, el valor del coeficient de correlació de *Spearman* entre les dues variables és de -0,2185687.

3r. Prendre Decisió

$P_{bilateral}$ = 0,2459 → No es rebutja H_0

5è. Conclusions

Com es pot veure en el llistat de R, el coeficient de correlació de *Spearman* té un valor de -0,2185687, que té un grau de significació bilateral associat de 0,2459, valor lo suficientment gran com per a no rebutjar la hipòtesi nul·la, i per tant podem afirmar que no existeix relació estadísticament significativa entre la puntuació en l'escala de dol anticipat de Zarit i la puntuació en l'escala de severitat de l'estat del malalt (GDS).

6. Determini quina és la millor variable per predir els valors en la puntuació de l'escala Zarit per mesurar el dol anticipat, raoni la seva resposta. Construeixi el model de regressió, valori la seva validesa.

El primer pas consisteix en determinar quina variable pot ser el millor predictor de les puntuacions en l'escala de dol anticipat de Zarit, i per això obtenim les correlacions de *Pearson* entre les puntuacions en l'escala de Zarit i la resta de variables quantitatives mesures com a mínim en escala d'interval de la base de dades proporcionada:

Variable	r	p
Edat cuidador	0,4926	0,0057
Episodis ansiosos abans intervenció	-0,2378	0,2057
Episodis ansiosos després intervenció	-0, 0690	0,7492

Com es pot veure, de les tres variables, l'única que presenta una relació estadísticament significativa amb les puntuacions de Zarit és l'edat del cuidador, per tant aquesta és la variable que es té en compte per generar el model de regressió lineal simple:

1er. Especificació del model:

Zarit = $b_0 + b_1 \cdot$ Edat cuidador + e_i

2n. Estimació dels coeficients:

Llistat R:

```
> RegModel.1 <- lm(ESCALA.ZARIT~EDAT, data=Situació2)
> summary(RegModel.1)

Call:
lm(formula = ESCALA.ZARIT ~ EDAT, data = Situació2)
```

```
Residuals:
   Min     1Q Median     3Q    Max
-34.49 -13.36   0.60  15.92  32.78

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.3778    16.1291  -0.147  0.88386
EDAT          0.6866     0.2293   2.995  0.00569 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.79 on 28 degrees of freedom
Multiple R-squared:  0.2426,	Adjusted R-squared:  0.2156
F-statistic: 8.969 on 1 and 28 DF,  p-value: 0.005689
```

Com es pot veure a partir del llistat de R, el valor de la constant i del coeficient de regressió són:

$$b_0 = -2{,}3778$$

$$b_1 = 0{,}6866$$

Per tant, l'equació de regressió simple obtinguda és la següent:

$$\widehat{Zarit} = \; -2{,}3778 + 0{,}6866 \cdot \; Edat\ \widehat{cuidadors}$$

3r. Validació del model

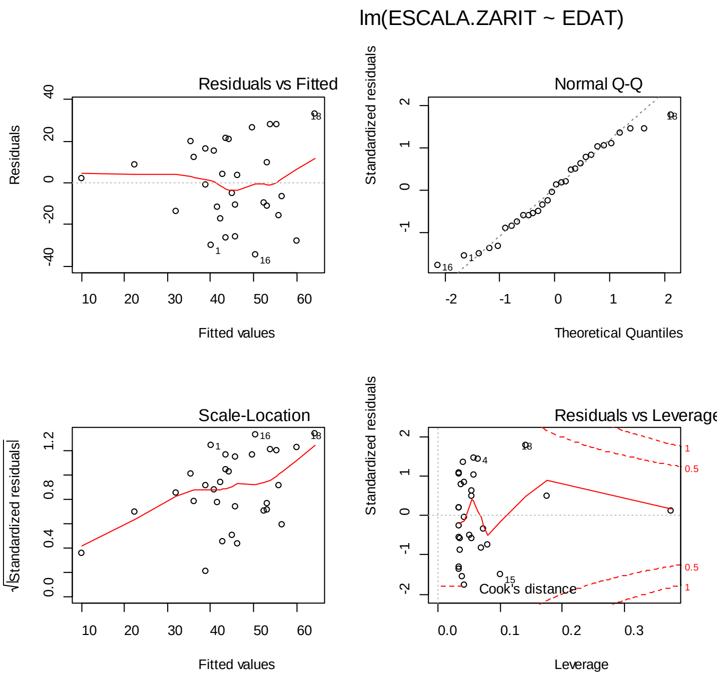
Com es pot veure en el llistat de R, el valor de l'estadístic F de Snedecor de l'anàlisi de variància de la regressió és de 8,969 que amb un grau de llibertat del numerador i 28 del denominador té un grau de significació de 0,005689, el que indica que la proporció de variabilitat de les puntuacions de l'escala Zarit explicada per l'equació de regressió és estadísticament superior a la no explicada, per tant el model és vàlid. A la mateixa conclusió s'arriba si ens fixem en la significació del coeficient de regressió ($b_1 = 0{,}6866$; $t(28) = 2{,}995$; $p = 0{,}00569$), aspecte totalment lògic donat que es tracta d'un model de regressió lineal simple. A més a més, el valor del coeficient de determinació ajustat és de 0,2156, el que indica que a partir de l'equació de regressió obtinguda, s'aconsegueix explicar un 21,56% del total de variabilitat de les puntuacions de Zarit.

A més a més per determinar la validesa del model, cal fer una anàlisi dels residuals:

Llistat R:

```
> resettest(ESCALA.ZARIT ~ EDAT, power=2:3, type="regressor", data=Situació2)

  RESET test
data:  ESCALA.ZARIT ~ EDAT
RESET = 0.67119, df1 = 2, df2 = 26, p-value = 0.5197

> bptest(ESCALA.ZARIT ~ EDAT, varformula = ~ fitted.values(RegModel.1),
studentize=FALSE, data=Situació2)

  Breusch-Pagan test
data:  ESCALA.ZARIT ~ EDAT
BP = 2.4047, df = 1, p-value = 0.121
```

```
> dwtest(ESCALA.ZARIT ~ EDAT, alternative="greater", data=Situació2)

 Durbin-Watson test
data:  ESCALA.ZARIT ~ EDAT
DW = 1.8276, p-value = 0.3055
alternative hypothesis: true autocorrelation is greater than 0
```



Com es pot veure en el llistat de R, el test RESET de Ramsey, que avalúa si el model especificat pot ajustar a un model quadràtic o cúbic, ens porta a no rebutjar la hipòtesi nul·la (0,67119, 2 graus de llibertat del numerador i 26 del denominador; p = 0,5179), el que va en la línia de poder afirmar que el model lineal ajustat és adequat. Per altra banda la prova de *Breusch-Pagan* valora si existeix heterocedasticitat entre els errors al llarg del recorregut de valors de la variable criteri, en aquest cas no es rebutja la hipòtesi nul·la (BP = 2,4047; 1 grau de llibertat; p = 0,121), per tant podem considerar que el residual es distribueix de forma homogènia.
Finalment, la prova de *Durwin-Watson* valora si existeix autocorrelació entre els errors (DW = 1,8276; p = 0,3055). En aquest cas no es rebutja la hipòtesi nul·la, per tant podem afirmar que no existeix autocorrelació entre els residuals.

A més a més, s'ha de valorar si el residual es distribueix aleatòriament, seguint un model de la llei normal. Això es pot fer a partir de proves de bondat d'ajuts o bé a nivell gràfic. Com es pot veure a la representació gràfica obtinguda, els residuals es comporten aleatòriament (gràfic *residuals vs fitted*) i seguint el model de la llei normal, amb petits desajustos a les cues (gràfic *normal Q-Q*). D'altra banda, en el gràfic *residuals vs leverage*, podem veure que els residuals *studentitzats* presenten valors entre -2 i +2 t (-,8578 i 1,8654), el que indicaria que els valors dels residuals no són molt elevats, i tampoc existeixen valors influents donat que el valor de la distància de *Cook* en cap cas supera el 0,5, de fet el valor més elevat és de 0,2642 corresponent al subjecte 18.

7. ¿Existeixen diferències estadísticament significatives en la puntuació en l'escala Zarit per mesurar dol anticipat en funció de la classe social del cuidador familiar?

1er. Plantejament hipòtesi nul·la:

H_0: $\mu_{\text{Zarit-Baix}} = \mu_{\text{Zarit-Mig}} = \mu_{\text{Zarit-Alt}} = = \mu_{\text{Zarit-Molt Alt}}$

2n. Càlcul de l'estadístic de contrast: ANOVA de grups independents

2.1. Supòsits i condicions d'aplicació

2.1.1. Assumim que les mostres en els quatre grups de comparació s'han obtingut a l'atzar del total de subjectes de la població

2.1.2. Normalitat

H_0: F(puntuació Zarit) = Fo (puntuació Zarit)

Càlcul estadístic de contrast: Prova de bondat d'ajuts a la normal de *Shapiro-Wilk*

Llistat R:

```
> .norm.test  # Pruebas de normalidad
# Prueba de ajuste a la normal: Shapiro-Wilk :
Situació2$NIVELL.SOCIAL: ALT
  Shapiro-Wilk normality test
data:  Situació2$ESCALA.ZARIT
W = 0.94703, p-value = 0.7162
-----------------------------------------------------
Situació2$NIVELL.SOCIAL: BAIX
  Shapiro-Wilk normality test
data:  Situació2$ESCALA.ZARIT
W = 0.88314, p-value = 0.2018
-----------------------------------------------------
Situació2$NIVELL.SOCIAL: MIG
  Shapiro-Wilk normality test
data:  Situació2$ESCALA.ZARIT
W = 0.94825, p-value = 0.6936
-----------------------------------------------------
Situació2$NIVELL.SOCIAL: MOLT ALT
  Shapiro-Wilk normality test
data:  Situació2$ESCALA.ZARIT
W = 0.98811, p-value = 0.9916
```

Decisió estadística: com es pot veure en el llistat de R, la puntuació en l'escala de dol anticipat de Zarit segueix el model de la llei normal en els quatre grups a comparar. Així l'estadístic *Shapiro-Wilk* de bondat d'ajuts a la llei normal, mostra els següents valors: nivell social baix: W = 0,88314; p = 0,2018; nivell social mig: W = 0,94825; p = 0,6936; nivell social alt: W = 0,94703; p = 0,7162 i nivell social molt alt: W = 0,98811; p = 0,9916. En els quatre casos el grau de significació és gran, el que ens porta a no rebutjar la hipòtesi nul·la i per tant poder afirma que les puntuacions en l'escala de dol anticipat de Zarit segueixen el model de la llei normal en els quatre grups.

2.1.3. Homocedasticitat

H_0: $\sigma^2_i = \sigma^2_j$

Càlcul estadístic de contrasts: test de Bartlet Box

Llistat R:

```
> with(Situació2, tapply(ESCALA.ZARIT, NIVELL.SOCIAL, var, na.rm=TRUE))
     ALT     BAIX      MIG MOLT ALT 
629.1000 387.9286 507.3571 421.7143 

> bartlett.test(ESCALA.ZARIT ~ NIVELL.SOCIAL, data=Situació2)

  Bartlett test of homogeneity of variances
data:  ESCALA.ZARIT by NIVELL.SOCIAL
Bartlett's K-squared = 0.38917, df = 3, p-value = 0.9425
```

Decisió estadística: com es pot veure en el llistat de R, el valor de la prova de *Bartlet-Box* és de 0,38917 que amb 3 graus de llibertat té un grau de significació de 0,9425, valor suficientment gran com per a no rebutjar hipòtesi nul·la, i per tant poder afirmar que és compleix l'homocedasticitat.

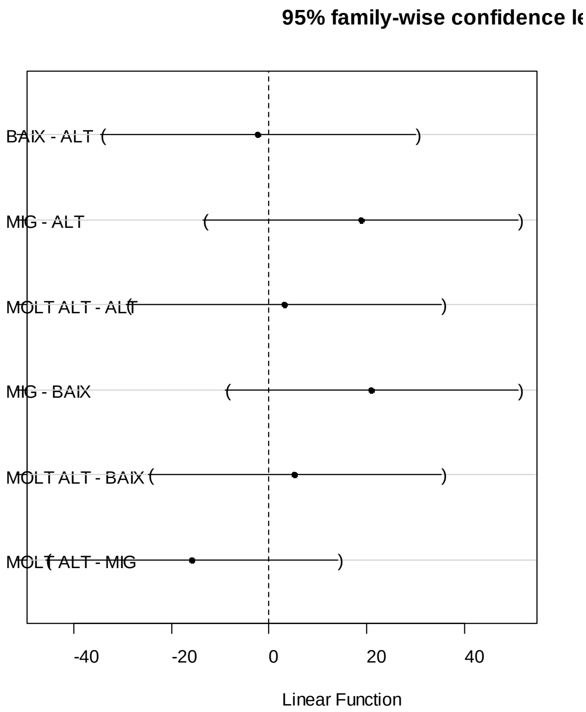
2.2. ANOVA de grups independents

Donat que assumim que les quatres mostres s'han obtingut a l'atzar i s'ha comprovat que les puntuacions en l'escala de dol anticipat de Zarit segueixen el model de la llei normal en els quatre grups a comparar i finalment hi ha homocedasticitat, podem obtenir l'ANOVA de grups independents.

Llistat R:

```
> AnovaModel.2 <- aov(ESCALA.ZARIT ~ NIVELL.SOCIAL, data=Situació2)
> summary(AnovaModel.2)
              Df Sum Sq Mean Sq F value Pr(>F)
NIVELL.SOCIAL  3   2114   704.6   1.482  0.243
Residuals     26  12365   475.6               

>       with(Situació2,       numSummary(ESCALA.ZARIT,       groups=NIVELL.SOCIAL, 
statistics=c("mean", "sd")))
          mean       sd data:n
ALT      39.50 25.08187      6
BAIX     37.25 19.69590      8
MIG      58.25 22.52459      8
MOLT ALT 42.50 20.53568      8

> local({
+   .Pairs <- glht(AnovaModel.2, linfct = mcp(NIVELL.SOCIAL = "Tukey"))
+   print(summary(.Pairs)) # pairwise tests
+   print(confint(.Pairs)) # confidence intervals
+   print(cld(.Pairs)) # compact letter display
+   old.oma <- par(oma=c(0,5,0,0))
+   plot(confint(.Pairs))
+   par(old.oma)
+ })
```

```
   Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: Tukey Contrasts
Fit: aov(formula = ESCALA.ZARIT ~ NIVELL.SOCIAL, data = Situació2)
Linear Hypotheses:
                         Estimate Std. Error t value Pr(>|t|)
BAIX - ALT == 0             -2.25      11.78  -0.191    0.997
MIG - ALT == 0              18.75      11.78   1.592    0.400
MOLT ALT - ALT == 0          3.00      11.78   0.255    0.994
MIG - BAIX == 0             21.00      10.90   1.926    0.242
MOLT ALT - BAIX == 0         5.25      10.90   0.481    0.962
MOLT ALT - MIG == 0        -15.75      10.90  -1.444    0.484
(Adjusted p values reported -- single-step method)

   Simultaneous Confidence Intervals
Multiple Comparisons of Means: Tukey Contrasts
Fit: aov(formula = ESCALA.ZARIT ~ NIVELL.SOCIAL, data = Situació2)
Quantile = 2.7399
95% family-wise confidence level
Linear Hypotheses:
                         Estimate lwr      upr
BAIX - ALT == 0           -2.2500 -34.5186  30.0186
MIG - ALT == 0            18.7500 -13.5186  51.0186
MOLT ALT - ALT == 0        3.0000 -29.2686  35.2686
MIG - BAIX == 0           21.0000  -8.8750  50.8750
MOLT ALT - BAIX == 0       5.2500 -24.6250  35.1250
MOLT ALT - MIG == 0      -15.7500 -45.6250  14.1250

     ALT     BAIX      MIG MOLT ALT
     "a"      "a"      "a"      "a"
```



Com es pot veure en el llistat de R, el valor de la F de Snedecor del quadre resum de l'ANOVA de grups independents és igual a 1,482 ($F(3; 26) = 1,482$)

3r. Prendre Decisió

$p\ (F \geq 1,482 \mid H_0) = 0,243 \rightarrow$ No es rebutja hipòtesi nul·la

4rt. Conclusions

A partir de les nostres dades es pot afirmar que no existeixen diferències estadísticament significatives en les puntuacions en l'escala de dol anticipat de Zarit entre els quatre grups de nivell social, donat que el valor de l'estadístic *F* de *Snedecor* de l'ANOVA de grups independents és de 1,482, que amb 3 graus de llibertat del numerador i 26 del denominador té un grau de significació associat de 0,243, valor lo suficientment gran com per a no rebutjar hipòtesi nul·la, per tant no té sentit interpretar els contrastos *a posteriori* perquè no existeixen diferències entre els grups.

8. ¿Les nostres dades confirmen que existeixen diferències estadísticament significatives en el número d'episodis ansiosos en un any abans d'un procés de suport terapèutic en funció de la classe social del cuidador familiar?

1er. Plantejament hipòtesi nul·la:

H_0: $\mu_{\text{Episodis ansiosos inici-Baix}} = \mu_{\text{Episodis ansiosos inici -Mig}} = \mu_{\text{Episodis ansiosos inici -Alt}} = = \mu_{\text{Episodis ansiosos inici -Molt Alt}}$

2n. Càlcul de l'estadístic de contrast: ANOVA de grups independents

2.1. Supòsits i condicions d'aplicació

2.1.1. Assumim que les mostres en els quatre grups de comparació s'han obtingut a l'atzar del total de subjectes de la població

2.1.2. Normalitat

H_0: F(episodis ansiosos inici) = Fo (episodis ansiosos inici)

Càlcul estadístic de contrast: Prova de bondat d'ajuts a la normal de *Shapiro-Wilk*

Llistat R:

```
> .norm.test  # Pruebas de normalidad
# Prueba de ajuste a la normal: Shapiro-Wilk :

Situació2$NIVELL.SOCIAL: ALT
  Shapiro-Wilk normality test
data:  Situació2$EPISODIS_1
W = 0.9067, p-value = 0.415
----------------------------------------------
Situació2$NIVELL.SOCIAL: BAIX
  Shapiro-Wilk normality test
data:  Situació2$EPISODIS_1
W = 0.8765, p-value = 0.1743
----------------------------------------------
Situació2$NIVELL.SOCIAL: MIG
  Shapiro-Wilk normality test
data:  Situació2$EPISODIS_1
W = 0.78017, p-value = 0.0175
----------------------------------------------
```

```
Situació2$NIVELL.SOCIAL: MOLT ALT
  Shapiro-Wilk normality test
data:  Situació2$EPISODIS_1
W = 0.9226, p-value = 0.4513
```

Decisió estadística: com es pot veure en el llistat de R, la el número d'episodis ansiosos patits en un any abans del procés de suport terapèutic segueix el model de la llei normal en tres dels quatre grups a comparar. Així l'estadístic *Shapiro-Wilk* de bondat d'ajuts a la llei normal, mostra els següents valors: nivell social baix: W = 0,8765; p = 0,1743; nivell social mig: W = 0,78017; p = 0,0175; nivell social alt: W = 0,9067; p = 0,415 i nivell social molt alt: W = 0,9226; p = 0,4513. Per nivell social baix, alt i molt alt el grau de significació és gran, el que ens porta a no rebutjar la hipòtesi nul·la i per tant poder afirma que el número d'episodis ansiosos en l'any abans del procés de suport terapèutic segueix el model de la llei normal en aquests tres grups, però no així en el grup de nivell social mig, on el grau de significació és prou petit com per a rebutjar la hipòtesi nul·la i per tant afirmar que en aquest cas la variable no segueix el model de la llei normal.

2.1.3. Homocedasticitat

H_0: $\sigma^2_i = \sigma^2_j$

Càlcul estadístic de contrasts: test de Bartlet Box

Llistat R:

```
> with(Situació2, tapply(EPISODIS_1, NIVELL.SOCIAL, var, na.rm=TRUE))
      ALT      BAIX       MIG  MOLT ALT
 5.866667 10.857143  8.982143  6.982143

> bartlett.test(EPISODIS_1 ~ NIVELL.SOCIAL, data=Situació2)
  Bartlett test of homogeneity of variances
data:  EPISODIS_1 by NIVELL.SOCIAL
Bartlett's K-squared = 0.62544, df = 3, p-value = 0.8906
```

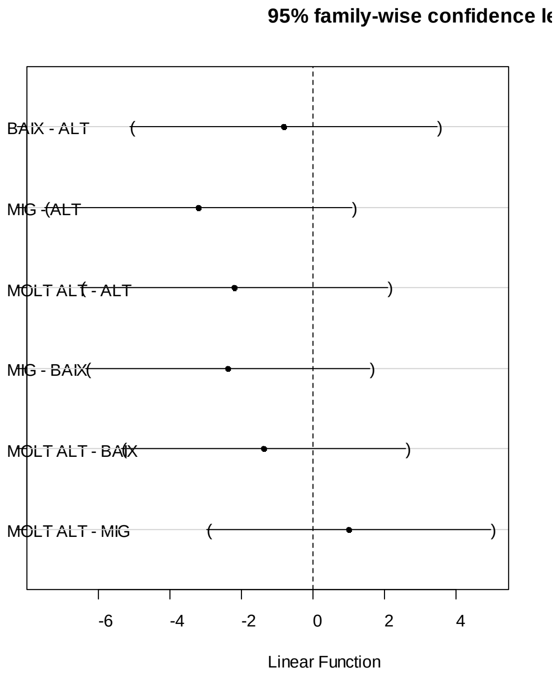
Decisió estadística: com es pot veure en el llistat de R, el valor de la prova de *Bartlet-Box* és de 0,62544 que amb 3 graus de llibertat té un grau de significació de 0,8906, valor suficientment gran com per a no rebutjar hipòtesi nul·la, i per tant poder afirmar que és compleix l'homocedasticitat.

2.2. ANOVA de grups independents

Donat que assumim que les quatres mostres s'han obtingut a l'atzar i s'ha comprovat que el número d'episodis ansiosos segueix el model de la llei normal en tres dels quatre grups a comparar(l'ANOVA és prou robust davant l'incompliment de la normalitat) i finalment hi ha homocedasticitat, podem obtenir l'ANOVA de grups independents.

Llistat R:

```
> AnovaModel.3 <- aov(EPISODIS_1 ~ NIVELL.SOCIAL, data=Situació2)
> summary(AnovaModel.3)
              Df Sum Sq Mean Sq F value Pr(>F)
NIVELL.SOCIAL  3  43.58  14.528    1.74  0.183
Residuals     26 217.08   8.349

> with(Situació2, numSummary(EPISODIS_1, groups=NIVELL.SOCIAL, statistics=c("mean",
"sd")))
             mean       sd data:n
ALT      6.333333 2.422120      6
BAIX     5.500000 3.295018      8
MIG      3.125000 2.997022      8
MOLT ALT 4.125000 2.642374      8

> local({
+   .Pairs <- glht(AnovaModel.3, linfct = mcp(NIVELL.SOCIAL = "Tukey"))
+   print(summary(.Pairs)) # pairwise tests
+   print(confint(.Pairs)) # confidence intervals
+   print(cld(.Pairs)) # compact letter display
+   old.oma <- par(oma=c(0,5,0,0))
+   plot(confint(.Pairs))
+   par(old.oma)
+ })

   Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: Tukey Contrasts
Fit: aov(formula = EPISODIS_1 ~ NIVELL.SOCIAL, data = Situació2)
Linear Hypotheses:
                     Estimate Std. Error t value Pr(>|t|)
BAIX - ALT == 0       -0.8333     1.5605  -0.534    0.950
MIG - ALT == 0        -3.2083     1.5605  -2.056    0.194
MOLT ALT - ALT == 0   -2.2083     1.5605  -1.415    0.501
MIG - BAIX == 0       -2.3750     1.4448  -1.644    0.372
MOLT ALT - BAIX == 0  -1.3750     1.4448  -0.952    0.777
MOLT ALT - MIG == 0    1.0000     1.4448   0.692    0.899
(Adjusted p values reported -- single-step method)

   Simultaneous Confidence Intervals
Multiple Comparisons of Means: Tukey Contrasts
Fit: aov(formula = EPISODIS_1 ~ NIVELL.SOCIAL, data = Situació2)
Quantile = 2.7443
95% family-wise confidence level
Linear Hypotheses:
                     Estimate lwr     upr
BAIX - ALT == 0      -0.8333  -5.1159  3.4493
MIG - ALT == 0       -3.2083  -7.4909  1.0743
MOLT ALT - ALT == 0  -2.2083  -6.4909  2.0743
MIG - BAIX == 0      -2.3750  -6.3399  1.5899
MOLT ALT - BAIX == 0 -1.3750  -5.3399  2.5899
MOLT ALT - MIG == 0   1.0000  -2.9649  4.9649

     ALT     BAIX      MIG MOLT ALT
     "a"      "a"      "a"      "a"
```



Com es pot veure en el llistat de R, el valor de la F de Snedecor del quadre resum de l'ANOVA de grups independents és igual a 1,74 ($F(3; 26) = 1,74$)

3r. Prendre Decisió

$p\ (F \geq 1,74 \mid H_0) = 0,183 \rightarrow$ No es rebutja hipòtesi nul·la

4rt. Conclusions

A partir de les nostres dades es pot afirmar que no existeixen diferències estadísticament significatives en el número d'episodis ansiosos patits en l'any previ al procés de suport terapèutic entre els quatre grups de nivell social, donat que el valor de l'estadístic *F* de *Snedecor* de l'ANOVA de grups independents és de 1,74, que amb 3 graus de llibertat del numerador i 26 del denominador té un grau de significació associat de 0,183, valor lo suficientment gran com per a no rebutjar hipòtesi nul·la, per tant no té sentit interpretar els contrastos *a posteriori* perquè no existeixen diferències entre els grups.