

Universitat de València

Factors que més influeixen en l'estat d'ànim

Universitat de València

Clara Albert

NIU: 1530466

Maig del 2020

Disseny d'experiments

Introducció

En aquest treball és realitza un disseny d'un experiment per veure quins són els factors que més influeixen en uns determinats estats d'ànim.

- MOTIVACIÓ:

La motivació d'aquest treball ve donada a la situació viscuda dels últims mesos. Hem estat vivint un confinament i per tant hem passat molt temps a casa.

He estat escoltant moltes vegades dels meus amics, familiars, que estan avorrits, estressats, farts d'estar a casa, que tenen ganes de veure als seus amics... per tant hem va semblar interessant veure quines activitats feien canviar més els estats d'ànim.

Al llarg del temps, ha estat canviat molt la idea, ja que la meva primera era fer un qüestionari, però no s'adaptava al que es demanava al treball. A partir d'allà, s'ha anat "polint" fins trobar un disseny de l'experiment idoni.

- HIPÒTESIS:

El teu estat d'ànim varia en funció si et mantens ocupat (realitzes alguna activitat) o no.

Amb aquest treball pretenc fer el tests d'hipòtesi per comprovar si realment el fer certes activitats afecta, tan positivament com negativament, l'estat d'ànim.

Per fer l'estudi més complert, afegirem alguns factors per detectar les condicions que fan que aquella activitat influeixi més als ànims.

- OBJECTIUS:

L'objectiu principal del treball és comprovar com afecten les diferents activitats en determinats estats d'ànims. Fer fer-ho necessitaré dissenyar un experiments i veure quin és el model més adient, per tant, com a altres objectius tinc posar en pràctica tots els mètodes i models ensenyats durant l'assignatura. A més a més, familiaritzar-me encara més amb l'anàlisi de dades i amb la posterior comprovació de les hipòtesis del model. Finalment, ser capaç de saber redactar un informe amb tota la informació relacionada amb el treball.

1. Material i mètodes

Per fer aquest experiment passarem una pauta de activitats a un grup d'individus durant uns cert número de dies i els preguntarem com perceben el seu estat d'ànim.

Pe tant tindrem en consideració 5 factors:

- Factor A: Període (2 nivells)
- Factor B: Lloc (2 nivells)
- Factor C: Activitats (5 nivells)
- Factor bloc A: Individus (10 nivells)
- Factor bloc B: Dia (5 nivells)

Pels factors blocs, el dia representa el dia que es van realitzar les combinacions. Aquest experiment va durar 5 dies, que és el número de nivells. Per altra banda, el factor individu representa al número de individu que va realitzar aquella combinació. Tenim un total de 10 individus. He intentat fer els individus més homogenis possibles, per tal d'evitar l'efecte de factors no controlables.

Els altres tres factors són els que tindrem més en compte a fer el model. El factor període té com a nivells matí o tarda. És el moment en el qual havien de realitzar l'activitat. Per altra banda, el factor lloc fa referència a l'espai en el qual feien l'acció. Pren com a nivells, exterior o interior de casa. Per últim, tenim el factor activitat. Aquestes són:

- A- 15 minuts d'exercici físic
- B- 15 minuts de lectura
- C- 15 minuts de mirar pelis o series
- D- 15 minuts de parlar amb algun familiar o amic que no estigui convivint amb nosaltres
- E- 15 minuts de feines de la llar

A cada individu se li associarà 10 combinacions diferents tenint en compte tots els factors.

Per a cada unitat experimental, que en tenim un total de 100, s'han recollit 6 variables respostes que corresponen a 3 estats d'ànims: estrès, avorriment i felicitat , abans i després de realitzar l'activitat corresponent.

Les fonts de variació a estudiar són els 5 factors per separat, les interaccions entre els factors que no són blocs, l'error i el total.

2. Metodologia estadística

S'han recollit dades de 10 individus, on cada un tenia 10 combinacions aleatòries. Cada individu va realitzar dues combinacions al dia. D'aquestes dues, una es realitzava al matí i l'altra a la tarda.

Per fer-ho de la manera més aleatòria possible vaig fer les combinacions amb R. Per a cada individu, vaig fer dues seqüències de les 5 activitats sense repeticions. Vaig fer-ne dues, ja que vaig associar una al matí i l'altra a la tarda. Per una altra banda, vaig fer unes altres dues seqüències del factor B (exterior – interior) on també corresponien al matí o a la tarda. Per tant, vaig assignar cada el tractament a una unitat experimental. Aquesta aleatorització augmenta la suposició entre la independència de les observacions i dels errors.

Es van fer de manera aleatòria, per evitar l'efecte període i evitar l'efecte de factors no controlables. En aquest disseny, aquests factors poden ser el temps meteorològic o el dia de la setmana. S'acostuma a dir que els dilluns la gent està més trista i també amb un dia nuvolat.

A continuació, hi ha un fragment de R amb les funcions que vaig utilitzar, on les sortides d'aquestes corresponen a la seqüència del primer individu.

```
sample(c("A", "B", "C", "D", "E"), 5, replace=F)
## [1] "E" "C" "B" "A" "D"
sample(c("A", "B", "C", "D", "E"), 5, replace=F)
## [1] "C" "B" "E" "D" "A"
sample(c('ext', 'int'), 5, replace=T, prob=c(0.5, 0.5))
## [1] "ext" "ext" "int" "int" "ext"
sample(c('ext', 'int'), 5, replace=T, prob=c(0.5, 0.5))
## [1] "int" "ext" "int" "ext" "ext"
```

Una vegada fetes totes les combinacions per a cada individu, se'ls hi va enviar als corresponents individus. Aquests van anar apuntant en una escala del 0 al 10 els corresponents estats d'ànims abans i després de fer cada tractament.

Per aquest model, el disseny experimental adient serà un disseny de tres factors i dos blocs completament aleatoritzats. En aquest cas, aquest disseny és incomplet ja que per a cada individu no tenim totes les combinacions possibles.

Per fer l'anàlisi del disseny s'utilitzarà el software R. Passos que seguiré:

1. Editar les dades com per exemple restar després amb abans per tenir una variable resposta per a cada estat d'ànim.
2. Avaluar el factor bloc. Es realitzaran dos models, un amb el factor bloc i l'altre sense tenir-lo en compte. Si el factor bloc no és significatiu, faré la taula ANOVA amb el model sense aquest factor.
3. Analitzar el model escollit. En aquest model tindrem en compte tots els factors i les possibles interaccions.
4. Fer les mitjanes pel factor activitat, ja que és el factor principal.

Abans hem esmentat les interaccions. Les que, a priori, poden semblar més interessants són: factor A amb factor C, factor B amb factor C i factor A amb factor B i Factor C. Estudiar la interacció entre el factor A amb el factor B no la trobo gaire interessant, ja que el que volem avaluar principalment és l'efecte de l'activitat i en aquest últim cas no el tenim en compte.

Les hipòtesis que ens interessarà testar són:

- Test hipòtesis pel factor A: $H^A_0: \alpha_1 = \dots = \alpha_a$ (no hi ha efecte del factor A)
- Test hipòtesis pel factor B: $H^B_0: \beta_1 = \dots = \beta_b$ (no hi ha efecte del factor B)
- Test hipòtesis pel factor C: $H^C_0: \gamma_1 = \dots = \gamma_c$ (no hi ha efecte del factor C)
- Test hipòtesis per la interacció A-B: $H^{AB}_0: (\alpha\beta)_{ij} = 0$
- Test hipòtesis per la interacció A-C: $H^{AC}_0: (\alpha\gamma)_{ik} = 0$
- Test hipòtesis per la interacció B-C: $H^{BC}_0: (\beta\gamma)_{jk} = 0$
- Test hipòtesis per la interacció A-B-C: $H^{ABC}_0: (\alpha\beta\gamma)_{ijk} = 0$
- Test hipòtesis per al factor bloc individu: $H^{ind}_0: no\ hi\ ha\ efecte\ bloc\ individu$
- Test hipòtesis per al factor bloc dia: $H^{dia}_0: no\ hi\ ha\ efecte\ bloc\ dia$

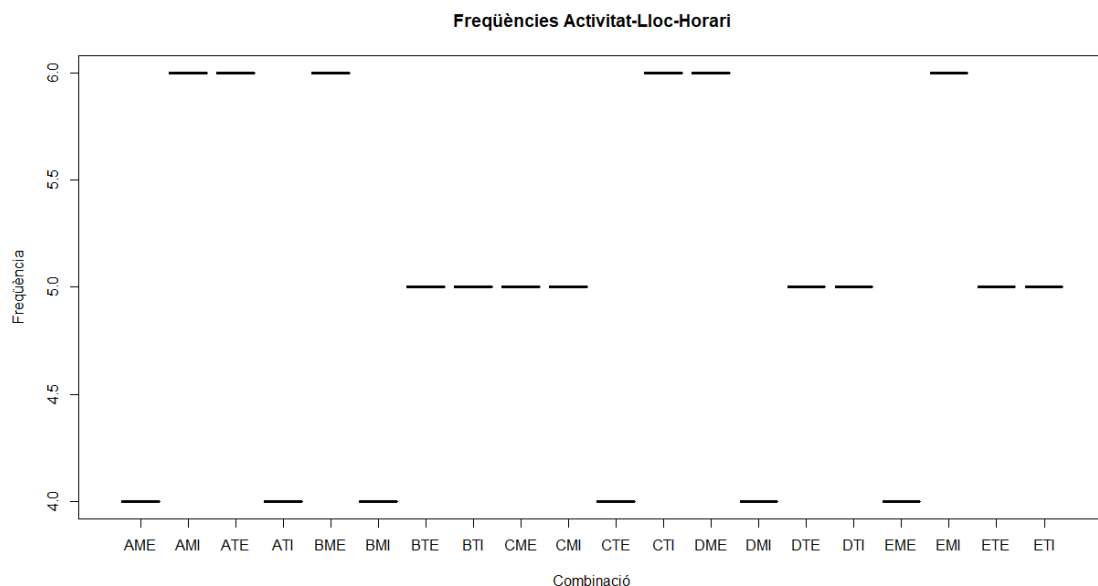
Finalment, ja tenim el disseny construït amb les hipòtesis que ens interessen contrastar. Per analitzar les dades farem servir la metodologia estadística de l'Anàlisi de la Variància, és a dir, les taules ANOVA. En aquest cas, el nivell de significació que utilitzarem serà l'estàndard; 0.05.

3. Validació de les dades i anàlisi descriptiva

En aquest punt tractarem de fer una validació de les dades completa amb l'ajuda de l'anàlisi descriptiva.

Al fitxer R podem observar totes les comprovacions de si la meua base de dades està correcte. Primer de tot, validem que no hi ha cap dada faltant. A més a més, al ser una base de dades relativament petita, és fàcil revisar-les visualment. Veiem que no hi ha cap anomalia o error. També comprovem que totes les variables es troben entre els límits establerts, com per exemple, que els valors de les variables respostes es troben entre 0 i 10.

Convé comprovar que tenim un número semblant d'observacions per a cada combinació o tractament. Per això fem sub-taules amb les condicions que volem i després les representem amb un gràfic.



En aquest gràfic observem que les freqüències per a totes les combinacions possibles tenint en compte els tres factors principals: activitat, lloc i horari, són semblants. Això vol dir que més o menys, hem avaluat el mateix número de cops cada combinació. No dona exacte, ja que s'han aleatoritzat les combinacions.

A més a més, una altra validació important és que per a cada dia, tinguem més a menys un número semblant d'activitats realitzades i hi estiguin present totes. Al document d'R, es pot observar que és així, on s'utilitza la funció *group_by()* i després *summarise()*.

Aprofitem la funció *summary()* que ens proporciona R, amb els estadístics descriptius més bàsic i importants.

summary(dades)

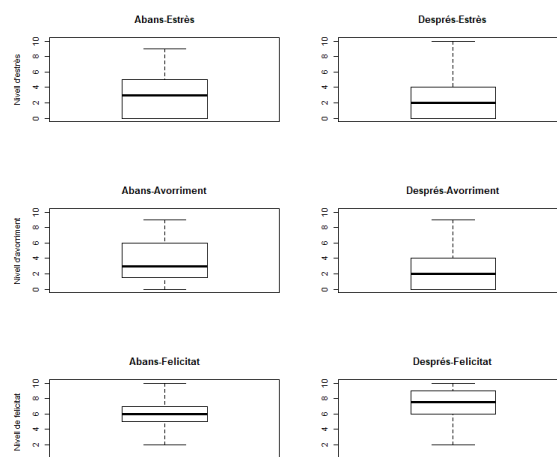
```
##      Dia      Individu      Horari      Lloc      Activitat
## Min.   :1    Min.   : 1.0    Min.   :1.0    Min.   :1.0    Length:100
## 1st Qu.:2    1st Qu.: 3.0    1st Qu.:1.0    1st Qu.:1.0    Class :character
## Median :3    Median : 5.5    Median :1.5    Median :1.5    Mode  :character
## Mean   :3    Mean   : 5.5    Mean   :1.5    Mean   :1.5
## 3rd Qu.:4    3rd Qu.: 8.0    3rd Qu.:2.0    3rd Qu.:2.0
## Max.   :5    Max.   :10.0    Max.   :2.0    Max.   :2.0

## Abans - Estrés Després - Estrés Abans - Avorriment Després - Avorriment
## Min.   :0.00    Min.   : 0.00    Min.   :0.00    Min.   :0.00
## 1st Qu.:0.00    1st Qu.: 0.00    1st Qu.:1.75    1st Qu.:0.00
## Median :3.00    Median : 2.00    Median :3.00    Median :2.00
## Mean   :3.24    Mean   : 2.17    Mean   :3.40    Mean   :2.41
## 3rd Qu.:5.00    3rd Qu.: 4.00    3rd Qu.:6.00    3rd Qu.:4.00
## Max.   :9.00    Max.   :10.00    Max.   :9.00    Max.   :9.00

## Abans - Felicitat Després - Felicitat
## Min.   : 0.00    Min.   : 0.00
## 1st Qu.: 5.00    1st Qu.: 6.00
## Median : 6.00    Median : 7.50
## Mean   : 6.12    Mean   : 7.27
## 3rd Qu.: 7.00    3rd Qu.: 9.00
## Max.   :10.00    Max.   :10.00
```

Per al factor dia veiem que els dies estan distribuïts de forma uniforme en l'interval [1,5] prenent només números naturals. Per altra banda, amb els individus passa el mateix, però l'interval és [1,10] i amb l'horari i el lloc on l'interval per als dos és [1,2]. Tots aquests estadístics ens estan donant la mateixa informació del factor; que estan equilibrats (hi ha el mateix número de dades per a cada nivell del factor).

Per altra banda, els quartils de les variables respostes ens ajuden a començar a tenir una idea dels resultats. En aquest cas, no estem diferenciant per cap factor, però veiem com hi ha certes diferències, tot i que no sabem amb quin nivell de significació ho podem afirmar. En la següent gràfica podem veure de una manera molt més clara de les diferències que parlava. El rang de les dades és de 0 a 10 per a tots els gràfics de caixa, però veiem que els nivell d'estrès i avorriment són més baixos que 5 i els de felicitat més alts que 5. La felicitat és la variable menys dispersa, però és l'única amb valors anòmals.



4. Anàlisi principal

Una vegada feta la validació de les dades, passarem a fer el disseny corresponent. Primer de tot, realitzarem una transformació de les variables respostes. Ara en tenim 6 en total, dues per a cada estat d'ànim. El que farà serà convertir-les en 3 variables respostes en total (una per a cada estat). Per fer-ho, restaré l'avaluació de després menys la d'abans.

Els nous valors de les variables respostes podran prendre valors negatius i seran entre -10 i 10, en els casos més extrems. La interpretació de les noves variables aquest cop canvia i serà: si una variable resposta pren un valor negatiu significa que aquell estat d'ànim en contret a disminuït després de fer l'activitat comparat amb abans de fer-la. Si pel contrari, el valor és positiu vol dir que el teu nivell ha augmentat. En cas de ser zero, s'ha quedat igual.

Per a l'estrès el millor són valors negatius, igual que per a l'avorriment, ja que significarà que després de fer l'activitat han augmentat. Aquestes dos actituds prenen una connotació negativa, per tant com menys estressat o avorrit et sentis millor. En canvi, per la felicitat el millor són els valors negatius (té una connotació positiva), ja que haurà dit que després de fer l'activitat ha estat més feliç.

Per poder aplicar els resultats de la taula ANOVA d'una manera vàlida s'han de complir tres requisits:

1. Independència: Les k mostres són independents.
2. Normalitat: $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \dots, k$
3. Homoscedasticitat: Igualtat de variàncies $\rightarrow \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$

Per tant, abans de realitzar l'anàlisi de les dades haurem de comprovar que els requisits acabats d'esmentar es compleixen.

- La independència l'haurem de garantir nosaltres amb una adequada presa de dades. En el nostre cas, comprovarem aquesta independència al realitzar el model per fer l'ANOVA.
- La normalitat la contrastarem mitjançant el test Kolmogorov-Smirnov i amb els següents gràfics.

```
mod.estres=lm(estres~act+lloc+periode,data=data)
ks.test(mod.estres$residuals,"pnorm",mean(mod.estres$residuals),sd(mod
.estres$residuals))
## One-sample Kolmogorov-Smirnov test
##
## data:  mod.estres$residuals
```



```
## D = 0.085004, p-value = 0.4653
## alternative hypothesis: two-sided

mod.avo=lm(avorriment~act+lloc+periode,data=data)
ks.test(mod.avo$residuals,"pnorm",mean(mod.avo$residuals),sd(mod.avo$residuals))
## One-sample Kolmogorov-Smirnov test
##
## data: mod.avo$residuals
## D = 0.075373, p-value = 0.6209
## alternative hypothesis: two-sided

mod.feli=lm(felicitat~act+lloc+periode,data=data)
ks.test(mod.feli$residuals,"pnorm",mean(mod.feli$residuals),sd(mod.feli$residuals))
## One-sample Kolmogorov-Smirnov test
##
## data: mod.feli$residuals
## D = 0.11677, p-value = 0.1308
## alternative hypothesis: two-sided
```

Per fer el test, hem tingut en compte el model sense interaccions ni blocs, només tenint en compte els factors activitat, lloc i període. Per a cada variable resposta hem fet un test diferent. En tots els casos els tests ens donen un p-valor superior a 0.05, per tant podem acceptar la hipòtesi nul·la de normalitat de les dades.

- L'homoscedasticitat també cal contrastar-la amb un test apropiat. En aquest cas aplicarem el test Breusch-Pagan.

```
bptest(mod.estres)
## studentized Breusch-Pagan test
##
## data: mod.estres
## BP = 6.4247, df = 6, p-value = 0.3773

bptest(mod.avo)
## studentized Breusch-Pagan test
##
## data: mod.avo
## BP = 1.1121, df = 6, p-value = 0.981

bptest(mod.feli)
## studentized Breusch-Pagan test
##
## data: mod.feli
## BP = 6.4259, df = 6, p-value = 0.3772
```

Continuem amb els models creats per la normalitat. En aquest cas tornem a acceptar les hipòtesis nul·les d'igualtat de variàncies.

Per tant, amb els resultats dels requisits podem aplicar la taula ANOVA a les nostres dades, ja que són independents, normals i amb variàncies iguals.

Realitzarem 3 taules ANOVA, una per a cada variable resposta. És tindran en compte tots els factors i totes les interaccions i el factor bloc dia el posarem com a factor fixe, ja que no passa res si el posem d'aquesta manera al model. Tot i tractant-lo com a fixe ara, no tindrem en compte les seves interaccions amb els altres factors.

Anàlisi de la variable resposta estrès

Realitzarem el model tenint en compte l'individu com a factor bloc i sense tenir-lo en compte. Els resultats amb R de la taula ANOVA dels dos models són els següents:

```
m1=lm(estres~act*lloc*horari, data=data)
m1.1=lmer(estres~act*lloc*horari+dia+(1|ind), data=data)
anova(m1)

## Analysis of Variance Table
##
## Response: stres
##              Df Sum Sq Mean Sq F value Pr(>F)
## act           4  11.060   2.7650   1.2765 0.2862
## lloc          1   3.017   3.0170   1.3929 0.2414
## horari        1   0.250   0.2500   0.1154 0.7349
## act:lloc      4   9.658   2.4145   1.1147 0.3555
## act:horari    4   9.090   2.2726   1.0492 0.3873
## lloc:horari   1   1.568   1.5676   0.7237 0.3975
## act:lloc:horari 4  10.584   2.6459   1.2215 0.3083
## Residuals    80 173.283   2.1660

anova(m1.1)

## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## act           8.4478  2.11196     4  77.145   1.1073 0.3592
## lloc          3.1574  3.15743     1  73.643   1.6555 0.2022
## horari        0.8804  0.88043     1  74.755   0.4616 0.4990
## dia           0.0063  0.00626     1  71.248   0.0033 0.9545
## act:lloc      6.2767  1.56917     4  77.285   0.8227 0.5146
## act:horari    10.2518  2.56295     4  77.105   1.3438 0.2614
## lloc:horari   1.2287  1.22870     1  75.218   0.6442 0.4247
## act:lloc:horari 8.9890  2.24726     4  77.032   1.1783 0.3270
```

El primer model, que anomenem m1, no tenim en compte el factor bloc individu i en el segon model, que anomenem m1.1, si que el tenim en compte. Veiem que els p-valors entre un model i l'altre varien, per tant hem de tenir en compte el factor bloc.

Una vegada ja sabem quin model és el més adient per a la variable resposta estres, veiem que tots els p-valors són més grans que 0.05 (el nivell de significació escollit). Per tant, podem afirmar amb un 95% de confiança, que les mitjanes són iguals per tots els tests d'hipòtesis esmentats a l'apartat 2.

Tot i això, el factor dia té un p-valor del 0.9545, per tant significa que les mitjanes són pràcticament iguals i és un factor que podríem treure del model (com més petit el p-valor és, el model és més eficient).

Apliquem la funció `emmeans(m1, ~act)` que ens proporciona R i ens surten les mitjanes per cada activitat. Abans hem dit que els millors valors per al estrès són els negatius, per tant, en mitjana l'activitat A, exercici físic, és la que provoca un nivell d'estrès més inferior al acabar que al començar. En canvi, mirar pel·lícules o series, també redueix el nivell d'estrès, però és l'activitat que ho fa menys.

```
emmeans(m1, ~act)
```

```
## NOTE: Results may be misleading due to involvement in interactions

##   act emmean    SE df lower.CL upper.CL
##   A   -1.562 0.336 80    -2.23   -0.8941
##   B   -0.954 0.333 80    -1.62   -0.2925
##   C   -0.613 0.333 80    -1.27    0.0492
##   D   -1.371 0.333 80    -2.03   -0.7091
##   E   -1.012 0.333 80    -1.67   -0.3508
##
## Results are averaged over the levels of: lloc, horari
## Confidence level used: 0.95
```

Tot i que hem rebutjat totes les hipòtesis nul·les dels factors i de les interaccions, he volgut veure les mitjanes per saber quina és la millor activitat tot i que les diferències no siguin significatives.

A més a més, podem comparar les mitjanes de l'activitat tenint en compte els factors lloc i horari amb la funció `emmeans(m1, ~act+lloc+horari)` (veure els resultats al arxiu R). Tot i que les diferències no són significatives, la mitjana més baixa de totes correspon a l'activitat A feta a l'exterior i al matí.

Anàlisi de la variable resposta avorriment

Com he fet amb l'altra variable resposta, també proposem dos models inicials.

```
m2=lm(avo~act*lloc*horari+dia, data=data)
m2.1=lmer(avo~act*lloc*horari+dia+(1|ind), data=data)
```

```
anova(m2)
```

```
## Analysis of Variance Table
##
## Response: avo
##
##          Df Sum Sq Mean Sq F value    Pr(>F)
## act         4   73.34  18.3350   3.1215 0.01942 *
## lloc         1    5.57   5.5670   0.9478 0.33326
## horari       1    2.25   2.2500   0.3831 0.53775
## dia          1    8.25   8.2502   1.4046 0.23951
## act:lloc      4    8.25   2.0637   0.3513 0.84237
## act:horari    4   17.68   4.4203   0.7525 0.55928
## lloc:horari   1    0.60   0.6044   0.1029 0.74923
## act:lloc:horari 4   31.02   7.7542   1.3201 0.26976
## Residuals    79 464.03   5.8737
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(m2.1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##
##          Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## act        28.0586   7.0146     4  78.327   1.2728 0.2878
## lloc         1.6040   1.6040     1  74.952   0.2910 0.5912
## horari       0.0008   0.0008     1  76.313   0.0001 0.9904
## dia          6.1847   6.1847     1  71.644   1.1222 0.2930
## act:lloc     25.5526   6.3881     4  77.807   1.1592 0.3354
## act:horari   24.0115   6.0029     4  78.207   1.0892 0.3677
## lloc:horari   0.3020   0.3020     1  76.858   0.0548 0.8155
## act:lloc:horari 22.4741   5.6185     4  77.203   1.0195 0.4027
```

En aquest cas si que existeixen diferències entre un model i l'altre, sobretot en el factor activitat. Amb el model sense tenir en compte el factor bloc, rebutgem la hipòtesi d'igualtat de mitjanes i, pel contrari, amb el model amb bloc, l'activitat no és significativa.

Per tant, estem rebutjant la hipòtesi nul·la H_0^{ind} i si que hi ha efecte individu i en conseqüència l'efecte bloc ha estat eficient en aquest model.

En aquest cas, el factor que menys explicació té en el model és l'horari, per tant fer-ho al matí o a la tarda no crea diferències al avorriment. La interacció lloc i horari també té un p-valor molt gran, per tant no serà gaire significatiu.

```
emmeans(m2, ~act)
```

```
## NOTE: Results may be misleading due to involvement in interactions
##
##   act emmean      SE df lower.CL upper.CL
## A   -1.293 0.557 79    -2.401  -0.1839
## B   -1.223 0.548 79    -2.314  -0.1330
## C   -1.183 0.548 79    -2.274  -0.0919
## D   -1.804 0.554 79    -2.907  -0.7006
```

```
## E      0.562 0.551 79    -0.536    1.6593
##
## Results are averaged over the levels of: lloc, horari
## Confidence level used: 0.95
```

Al mostrar totes les mitjanes depenent de quina activitat realitzis veiem que les mitjanes són més o menys molt semblants, on el valor més petit, per tant el millor valor, és parlar amb un amic o familiar, menys la mitjana de l'activitat E, feines de la llar, que és inclús positiva. Per tant, ens esta indicant que els individus se sentien més avorrits després de fer les feines de la llar que abans.

En aquest cas, podem tornar a comparar les mitjanes tenint en compte els altres factors, tot i que la funció serà *emmeans(m2,~act+lloc+horari)*. Hi ha una combinació que destaca respecte les altres, que és la lectura durant 15 minuts a l'exterior i pel matí. Totes les altres combinacions tenen mitjanes semblants on es troben a l'interval (-2,2).

Anàlisi de la variable resposta felicitat

Finalment, tenim els dos models per a la felicitat.

```
m3=lm(feli~act*lloc*horari+dia, data=data)
m3.1=lmer(feli~act*lloc*horari+dia+(1|ind), data=data)

## boundary (singular) fit: see ?isSingular

anova(m3)

## Analysis of Variance Table
##
## Response: feli
##
##          Df Sum Sq Mean Sq F value    Pr(>F)
## act         4   50.100  12.5250   3.7985 0.007078 **
## lloc         1    5.058   5.0581   1.5340 0.219187
## horari       1    0.810   0.8100   0.2456 0.621532
## dia         1    0.848   0.8477   0.2571 0.613534
## act:lloc     4   27.356   6.8390   2.0741 0.092045 .
## act:horari   4   16.815   4.2038   1.2749 0.286930
## lloc:horari  1    0.264   0.2636   0.0799 0.778134
## act:lloc:horari 4    1.006   0.2514   0.0762 0.989279
## Residuals   79 260.494   3.2974
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(m3.1)

## Type III Analysis of Variance Table with Satterthwaite's method
##
##          Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## act         8.2325  2.05812      4     79   0.6242 0.6466
```

```
## lloc          1.3432 1.34316      1    79  0.4073 0.5252
## horari        0.0358 0.03576      1    79  0.0108 0.9173
## dia           0.1728 0.17277      1    79  0.0524 0.8195
## act:lloc      3.3709 0.84273      4    79  0.2556 0.9054
## act:horari    1.4350 0.35876      4    79  0.1088 0.9791
## lloc:horari   0.2715 0.27149      1    79  0.0823 0.7749
## act:lloc:horari 1.0056 0.25139      4    79  0.0762 0.9893
```

Ens passa el mateix que amb l'avorriment. Podem afirmar amb un nivell de confiança del 5% que el bloc ha estat eficient. Per tant, ens centrem amb el model que tenim present el bloc.

Com amb totes les altres variables respostes, tots els factors i les interaccions ens donen p-valors més gran que 0.05, i per tant rebutgem totes les hipòtesis nul·les que havíem plantejat al inici menys la hipòtesi nul·la de l'eficiència del factor bloc.

```
emmeans(m3, ~act)
```

```
## NOTE: Results may be misleading due to involvement in interactions

##   act emmean    SE df lower.CL upper.CL
##   A   2.1363 0.417 79  1.30560    2.967
##   B   0.8110 0.410 79 -0.00592    1.628
##   C   1.1830 0.411 79  0.36548    2.001
##   D   1.6018 0.415 79  0.77511    2.429
##   E   0.0154 0.413 79 -0.80700    0.838
##
## Results are averaged over the levels of: lloc, horari
## Confidence level used: 0.95
```

Veiem que totes les mitjanes són positives, per tant hi ha hagut un efecte al nivell de felicitat positiu. El valor més significatiu és per a l'exercici físic, on el canvi és de més de 2 nivells. Per altre costat, hi ha una activitat que gairebé no hi ha canvi en abans de realitzar-la i després, que en aquest cas és fer coses de la llar.

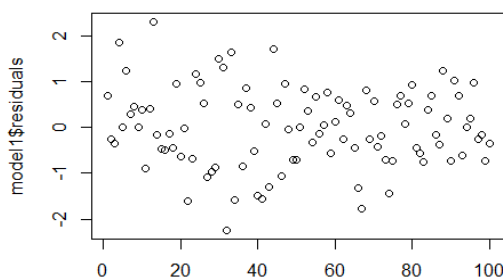
Si tenim en compte els altres factors per calcular les mitjanes, s'observa que l'activitat E és l'única que té més d'una combinació amb mitjana negativa. El contrari passa amb l'activitat A, que manté mitjanes més altes que la resta, especialment al matí.

5. Validació de l'anàlisi i comprovació de les hipòtesis del model

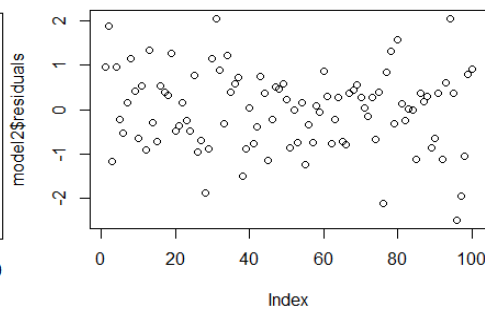
Per fer la validació de l'anàlisi analitzarem els residus d'aquests, per veure si, més o menys, són tots semblants.

Per a tots els models (model per a la variable resposta estrès, avorriment o felicitat) hem rebutjat totes les hipòtesis del model, excepte la hipòtesis de l'efectivitat del bloc individu. Per tant, en els models en factor bloc s'ha de tenir en compte a l'hora de fer la taula ANOVA. Per altra banda, dona igual en quin nivell del factor ens trobem que no hi haurà canvi amb la variable resposta, per tant, tots els factors no tenen cap efecte amb la resposta.

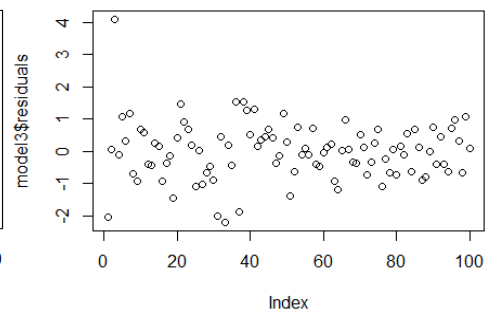
A la gràfica dels residus per als models veiem que no hi ha cap residu anòmal i més o menys tots segueixen una tendència semblant.



Residus model 1



Residus Model 2



Residus Model 3

6. Conclusions

A partir del estudi realitzat i la seva corresponent anàlisi i interpretació de les dades s'obtenen els resultats exposats a continuació.

Podem afirmar amb un 95% de confiança que el tipus d'activitat, el lloc on es realitza i el moment no intervenen al efecte dels estats d'ànim, independentment de quin estiguem valorant.

Per un altre costat, en aquest experiment els individus han estat eficients. Això vol dir que tenim diferències entre uns i els altres. S'hauria de fer un estudi tenint en compte els factors no controlables que poden provocar aquestes diferències.

Tot i això, s'han analitzat les mitjanes de les activitats per poder extreure alguna conclusió més. Les mitjanes per a la felicitat són les que varien més, resultat que no sorprèn, ja que si no tenim en compte el factor bloc, la hipòtesi d'igualtat de mitjanes d'activitats es rebutjava. També ocorria amb l'avorriment, però aquest cas pot ser degut a l'activitat E, que està força llunyana de les altres mitjanes, que són molt semblants.

A més a més, veiem que no hi ha cap efecte significatiu de realitzar l'activitat dins o fora de casa o pel matí i per la tarda. Per tant, podem afirmar que aquests factors no intervenen a la resposta del individu. Tot i així, si veiem els resultats obtinguts a partir de la funció *emmeans()* aplicades a les possibles combinacions entre els factors, s'observa una petita diferència entre el lloc. Per als tres estats d'ànim, és millor fer-ho a l'exterior o a l'interior. En canvi, si mirem les mitjanes tenint en compte l'activitat i l'horari no s'aprecia cap tendència de si es millor fer-ho al matí que a la tarda.

Cal remarcar que totes les diferències que hem esmentat durant el treball i a les conclusions no són significatives, però s'han mirat per tal d'intentar trobar algun patró, diferència per molt petita que sigui.

DISCUSSIÓ:

L'activitat que més redueix l'estrès és 15 minuts d'exercici físic.

L'activitat que més redueix l'avorriment és 15 minuts de parlar amb algú.

L'activitat que més augmenta la felicitat és 15 minuts d'exercici físic.

El moment en que es realitzen no marca cap diferència en el teu estat d'ànim.

Si és realitzen a l'exterior tenen un efecte més positiu en l'estat d'ànim vers si es fan a l'interior.