

Diseño de Experimentos AA

Jesus Emmanuel Ramos Davila

2023-03-13

Para esta seccion usaremos nuestros datos de Ataques al corazon para realizar un Diseño de experimentos. Esta seccion incluirea **ANOVA para diseños factoriales con 2 factores** , en esta seccion se usaran **efectos fijos** los cuales nosotros tenemos el control de los niveles (caso contrario a efectos aleatorios).

Paso 1: Carga de datos.

```
library(readr)
heart <- read_csv("heart.csv")

## Rows: 303 Columns: 14
## -- Column specification -----
## Delimiter: ","
## dbl (14): age, sex, cp, trtbps, chol, fbs, restecg, thalachh, exng, oldpeak,...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(heart)

## # A tibble: 6 x 14
##   age  sex  cp trtbps  chol  fbs restecg thala~1 exng oldpeak  slp  caa
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1   63    1    3   145   233    1     0     150    0    2.3    0    0
## 2   37    1    2   130   250    0     1     187    0    3.5    0    0
## 3   41    0    1   130   204    0     0     172    0    1.4    2    0
## 4   56    1    1   120   236    0     1     178    0    0.8    2    0
## 5   57    0    0   120   354    0     1     163    1    0.6    2    0
## 6   57    1    0   140   192    0     1     148    0    0.4    1    0
## # ... with 2 more variables: thall <dbl>, output <dbl>, and abbreviated
## #   variable name 1: thalachh
```

Para este diseno de experimentos se estudiara la relacion entre la variable chol (Colesterol total en sangre) y dos variables slp (slope) y (Tipo de angina de pecho) cp.

Nuestro interes es saber si a un mayor grado de colesterol en sangre y edad existe un cambio en el tipo de angina de pecho reportada.

Antes de realizar el experimento , debemos de revisar si el diseño es **balanceado**, esto quiere decir si todas las muestras son del mismo tamaño.

```
heart$cp <- as.factor(heart$cp)
heart$slp <- as.factor(heart$slp)
model<-lm(chol ~ cp * slp,data=heart)
replications(model,data=heart)
```

```
## $cp
## cp
##    0    1    2    3
## 143  50  87  23
##
## $slp
## slp
##    0    1    2
##   21 140 142
##
## $'cp:slp'
##      slp
## cp    0    1    2
##    0 11 84 48
##    1  2 12 36
##    2  5 33 49
##    3  3 11  9
```

```
(isBalanced=!is.list(replications(model, data=heart)))
```

```
## [1] FALSE
```

Observaciones: En este conjunto de datos los tipos de angina de pecho se encuentran desbalanceados.

Diseno de experimento ANOVA

Aunque sabemos que nuestro conjunto de datos no es **balanceado**. Revisemos con anova si existe un interaccion entre variable de respuesta e interacciones entre las variables dependientes.

```
heart.aov = aov(chol ~ cp * slp, data = heart)
summary(heart.aov)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## cp              3   5001   1666.9    0.607  0.611
## slp             2   2249   1124.5    0.410  0.664
## cp:slp          6   5412    902.1    0.329  0.922
## Residuals     291 798639   2744.5
```

Resultados: Se observo que no existe relacion entre el tipo de angina de pecho (cp) y slope (slp) con nuestra variable de respuesta chol (Colesterol), esto en previous analisis de seleccion de caracteristicas y regresiones no se observo alguna relacion entre ellas , tambien no se observo alguna interaccion entre las variables **cp** y **slp** , lo cual aunque se podria asumir que dado slp (slope) es un *estudio de electrocardiograma* no existe relacion alguna el tipo de angina de pecho con el resultado de slp.