

PEC4: contrastes de Hipótesis.

Juan Luis Acebal Rico

```
load("Data_PEC4.Rdata")  
ls()
```

```
## [1] "dat"
```

La enfermedad pulmonar obstructiva crónica (EPOC) es una enfermedad pulmonar común que reduce el flujo de aire y causa problemas respiratorios. Un grupo de investigación decide realizar un estudio para ampliar el conocimiento sobre esta patología. Con tal fin, crearon una cohorte de 150 pacientes con EPOC en la que registraron 14 variables con información clínica de los pacientes. Los datos están basados en un remutreo de la base de datos procedente de Kaggle.

Las variables que se encuentran en el dataset son las siguientes:

- *AGE*: edad del paciente.
- *gender*: genero del paciente (Femenino/Masculino).
- *Diabetes*: diabetes (Yes/No).
- *hypertension*: hipertensión (Yes/No).
- *PackHistory*: carga tabáquica (paquetes-año).
- *COPDSEVERITY*: categoría de severidad del EPOC.
- *CAT*: puntiacópn del cuestionario COPD Assessment Test.
- *MWT1Best*: prueba de marcha de 6 minutos (metros).
- *FEV1PRED*: volumen Espiratorio Forzado de la espirometría (%).
- *FVCPRED*: capacidad vital forzada (%).
- *HAD*: puntuación total del cuestionario de ansiedad y depresión HAD.
- *SGRQ*: puntuación del cuestionario de calidad de vida Saint George's Respiratory.
- *FEV1PRED_post*: volumen Espiratorio Forzado de la espirometría (%) post seis meses de terapia fisioterápéutica.

Os puede ser útil consultar el siguiente material:

- Apuntes de contraste de hipótesis
- Apuntes de contraste de dos muestras

Hay que entregar la práctica en fichero pdf o html. Se recomienda generar el informe con Rmarkdown que genera automáticamente el html/pdf a entregar. Se puede utilizar el fichero .Rmd, que disponéis en la PEC, como plantilla para resolver los ejercicios. Esta PEC debe realizarse de forma estrictamente individual, quedando del todo prohibido el uso de herramientas de IA. Cualquier indicio de copia será penalizado con un suspenso (D) por todas las partes implicadas y la posible evaluación negativa de la asignatura de forma íntegra.

Problema 1

- a) Inicialmente, interesa testear si la cohorte es representativa de la población de pacientes con EPOC. Estudios previos indican que la edad media de esta población es de 73 años. Realiza el contraste de hipótesis sobre si la media de edad de nuestra cohorte es diferente a la media de edad teórica. Plantea en contraste (H_0 y H_1). ¿Podemos rechazar la hipótesis nula de igualdad a un nivel de significación del 0.05? Crees que las diferencias son relevantes clínicamente?

```
head(dat)
```

```
##   AGE    gender Diabetes hypertension PackHistory COPDSEVERITY CAT MWT1Best
## 1  82 Masculino      No           No          45      SEVERE   10     369
## 2  63 Femenino      No           No          20    MODERATE    7     489
## 3  71 Masculino      No          Yes          20      MILD    16     423
## 4  74 Femenino      No           No          75      MILD    24     270
## 5  70 Femenino      No           No          36      MILD     5     459
## 6  67 Masculino      No           No          36    MODERATE   23     400
##  FEV1PRED FVCPRED HAD  SGRQ FEV1PRED_post
## 1      46      73   6 37.04      52.77846
## 2      54     110   2 17.95      63.66380
## 3     102      90   4 38.57     130.12021
## 4      83     102  19 45.00     101.70246
## 5      90     125   0 10.92     109.37212
## 6      62      84  10 50.03      71.59996
```

```
summary(dat)
```

```
##      AGE              gender      Diabetes      hypertension
## Min.   :44.00  Length:150      Length:150      Length:150
## 1st Qu.:65.00  Class :character  Class :character  Class :character
## Median :71.00  Mode  :character  Mode  :character  Mode  :character
## Mean   :69.23
## 3rd Qu.:75.00
## Max.   :83.00
##
##  PackHistory      COPDSEVERITY      CAT      MWT1Best
## Min.   :   3.00  Length:150      Min.   :   3.00  Min.   :210.0
## 1st Qu.: 27.00  Class :character  1st Qu.:12.00  1st Qu.:336.0
## Median : 36.75  Mode  :character  Median :18.00  Median :420.0
## Mean   : 39.62      Mean   :17.65  Mean   :410.7
## 3rd Qu.: 50.00      3rd Qu.:23.75  3rd Qu.:480.0
## Max.   :109.00      Max.   :32.00  Max.   :699.0
##
##                                     NA's   :1
##  FEV1PRED      FVCPRED      HAD      SGRQ
## Min.   :   3.29  Min.   : 27.00  Min.   :   0.00  Min.   :   2.00
## 1st Qu.: 42.00  1st Qu.: 71.00  1st Qu.:   6.00  1st Qu.:28.60
## Median : 60.50  Median : 84.00  Median :10.00  Median :37.38
## Mean   : 58.20  Mean   : 85.41  Mean   :11.19  Mean   :41.26
## 3rd Qu.: 75.00  3rd Qu.:102.75  3rd Qu.:14.75  3rd Qu.:56.20
## Max.   :102.00  Max.   :125.00  Max.   :56.20  Max.   :77.44
##
## FEV1PRED_post
```

```
## Min.    : 3.541
## 1st Qu.: 50.526
## Median : 73.215
## Mean    : 72.139
## 3rd Qu.: 92.487
## Max.    :140.887
##
```

```
t_test_age <- t.test(dat$AGE, mu = 73, alternative = "two.sided")
t_test_age
```

```
##
## One Sample t-test
##
## data:  dat$AGE
## t = -5.6925, df = 149, p-value = 6.448e-08
## alternative hypothesis: true mean is not equal to 73
## 95 percent confidence interval:
##  67.91684 70.53649
## sample estimates:
## mean of x
##  69.22667
```

Creo que hay evidencia significativa para decir que, la media de edad es diferente a 73 años. El test nos dice 69,226 años. Aunque la diferencia es pequeña, parece relevante clínica y estadísticamente, ya que la edad es un factor de riesgo importante en la EPOC. Esto se traduce en 3,77 años, que debemos de entender el contexto clínico para valorar si es relevante o no en la progresión de la enfermedad y en la respuesta a los tratamientos.

- b) En estudios previos, se ha estimado que alrededor del 30% de los pacientes con EPOC son mujeres. Queremos evaluar si la proporción de mujeres en nuestra cohorte es significativamente diferente a este valor teórico. Plantea y realiza el contraste de hipótesis para evaluar si la proporción de mujeres en nuestra cohorte es diferente a 30% teórico (H_0 y H_1). ¿Podemos rechazar la hipótesis nula de igualdad a un nivel de significación de 0.05? Comenta los resultados.

```
prop_mujeres <- mean(dat$gender == "Femenino", na.rm = TRUE)
x <- sum(dat$gender == "Femenino", na.rm = TRUE)
n <- sum(!is.na(dat$gender))
prop_test <- prop.test(x, n, p = 0.30, alternative = "two.sided", correct = FALSE)

prop_test
```

```
##
## 1-sample proportions test without continuity correction
##
## data:  x out of n, null probability 0.3
## X-squared = 0.12698, df = 1, p-value = 0.7216
## alternative hypothesis: true p is not equal to 0.3
## 95 percent confidence interval:
##  0.2445491 0.3914398
## sample estimates:
##          p
## 0.3133333
```

Parece que es ligeramente superior a la proporción teórica del 30% de mujeres en la población de pacientes con EPOC. El test ha proporcionado un valor de 31,33% y un p value de 0.7216, no rechazamos la hipótesis nula, por tanto y no podemos decir que la proporción de mujeres en nuestra cohorte sea diferente a la proporción teórica del 30%.

- c) Se desea comparar las medias del volumen espiratorio forzado predicho (FEV1PRED) entre los pacientes con alta carga tabáquica (PackHistory>40) y baja-media carga tabáquica (PackHistory ≤ 40 paquetes-año). ¿Existen diferencias significativas en las medias de FEV1PRED entre estos dos grupos de pacientes? Plantea y evalúa el contraste de hipótesis adecuado con un nivel de significación de 0.1 y comenta los resultados.

```
grupo_alta <- subset(dat, PackHistory > 40)$FEV1PRED
grupo_baja <- subset(dat, PackHistory <= 40)$FEV1PRED
```

```
shapiro.test(grupo_alta)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  grupo_alta
## W = 0.96555, p-value = 0.1158
```

```
shapiro.test(grupo_baja)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  grupo_baja
## W = 0.97017, p-value = 0.02878
```

```
var_alta <- var(grupo_alta, na.rm = TRUE)
var_baja <- var(grupo_baja, na.rm = TRUE)
```

```
t_test_fev1 <- t.test(grupo_alta, grupo_baja, alternative = "two.sided", var.equal = FALSE, conf.level = 0.1)
```

```
t_test_fev1
```

```
##
##  Welch Two Sample t-test
##
## data:  grupo_alta and grupo_baja
## t = -2.0828, df = 123.85, p-value = 0.03933
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
##  -14.116840 -1.606299
## sample estimates:
## mean of x mean of y
##  53.22127  61.08284
```

Existen diferencias significativas en las medias de FEV1PRED entre los dos grupos de pacientes. La media de FEV1PRED en el grupo de alta carga tabáquica es 53,22 y en el grupo de baja-media carga tabáquica

es 61,08. La diferencia entre las medias es de 7,86. El p value del test es de 0.03933, por lo que rechazamos la hipótesis nula de igualdad de medias entre los dos grupos, se ha usado un nivel de confianza del 90%. Es relevante aunque debe ser interpretado en el contexto de la enfermedad, de la prevención y respuesta a los tratamientos.

- d) Estudios previos hipotetizan que la proporción de pacientes con severos o muy severos es mayor en mujeres que en hombres. ¿Qué contraste plantearías para resolver la hipótesis? Presenta el planteamiento teórico del contraste, resuelve el contraste y comenta los resultados. Nota: el nivel de significación del contraste es 0.05 y sin corrección de continuidad de Yates.

```
dat$Severo <- ifelse(dat$COPDSEVERITY %in% c("SEVERE", "VERY SEVERE"), 1, 0)
tabla_severidad <- table(dat$gender, dat$Severo)

tabla_severidad

##
##           0  1
## Femenino 33 14
## Masculino 67 36

prop_mujeres_severo <- mean(dat$Severo[dat$gender == "Femenino"], na.rm = TRUE)
prop_hombres_severo <- mean(dat$Severo[dat$gender == "Masculino"], na.rm = TRUE)
x1 <- sum(dat$Severo[dat$gender == "Femenino"], na.rm = TRUE)
n1 <- sum(dat$gender == "Femenino", na.rm = TRUE)

x2 <- sum(dat$Severo[dat$gender == "Masculino"], na.rm = TRUE)
n2 <- sum(dat$gender == "Masculino", na.rm = TRUE)

prop_test_severidad <- prop.test(c(x1, x2), c(n1, n2), alternative = "greater", correct = FALSE)

prop_test_severidad

##
## 2-sample test for equality of proportions without continuity correction
##
## data:  c(x1, x2) out of c(n1, n2)
## X-squared = 0.38732, df = 1, p-value = 0.7331
## alternative hypothesis: greater
## 95 percent confidence interval:
## -0.1858486  1.0000000
## sample estimates:
##   prop 1   prop 2
## 0.2978723 0.3495146
```

No se rechaza la hipótesis nula, $p\text{-value} > 0.05$, no hay evidencia suficiente, el intervalo de confianza del 95% incluye el valor cero, lo que indica que no hay una diferencia significativa. Aunque la proporción de mujeres con severidad es inferior a la de hombres, 29,79 vs 34,95 la diferencia no es significativa estadísticamente o si lo fuera se tendría que interpretar en el contexto de la enfermedad y con otros estudios. En el contexto de la pregunta, no podemos afirmar la hipótesis alternativa de que la proporción de pacientes con severidad es mayor en mujeres que en hombres.

- e) Se cree que el alto impacto sintomatológico de la EPOC (puntuación CAT ≥ 20) afectará negativamente a la capacidad de ejercicio del paciente. La prueba del 6MWT (variable MWT1Best) mide la distancia

(metros) que el paciente puede caminar en 6 minutos. Realiza un test de homogeneidad de varianzas de los metros recorridos entre pacientes entre grupos de carga sintomatológica. Plantea el contraste (H_0 y H_1), calcula el contraste con un nivel de significación de 0.05 y comenta los resultados.

```
library(carData)
library(car)
grupo_cat <- as.factor(ifelse(dat$CAT >= 20, "Alta", "Baja-Media"))
dat$Grupo_CAT <- grupo_cat
str(dat$Grupo_CAT)
```

```
## Factor w/ 2 levels "Alta","Baja-Media": 2 2 2 1 2 1 1 2 1 1 ...
```

```
levene_test <- leveneTest(MWT1Best ~ Grupo_CAT, data = dat)
```

```
levene_test
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value   Pr(>F)
## group  1      10 0.001902 **
##      147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tenemos un F value de 10 y p-value es 0,00190, hay una diferencia significativa de varianzas entre los dos grupos.

f) ¿Cuál es el valor crítico superior del contraste previo. Y el valor de F si hubiesemos planteado un nivel de significación del 0.025?

```
qf(1 - 0.025, sum(grupo_cat == "Alta", na.rm = TRUE)-1, sum(grupo_cat == "Baja-Media", na.rm = TRUE)-1)
```

```
## [1] 1.577999
```

En el test de levene el valor de F fue 10, y el valor critico superior es 1,577, que es inferior, por lo que no hay igualdad de varianzas y lleva a rechazar H_0 ya que $F \text{ value } (1.577999) < F \text{ value calculado } (10)$.

g) Contrasta la hipótesis de si hay diferencias en los metros medios recorridos entre grupos de severidad del enunciado e. Plantea el contraste (H_0 y H_1). ¿Cuál es la estimación puntual y intervalo de confianza al 95% de la diferencia de metros entre grupos? y el valor p del contraste? (nota: asume varianzas iguales o no, según el resultado del apartado e y un nivel de significación de 0.05).

Ambos grupos presentan valores de p menores que 0.05 entonces no se puede asumir la normalidad de las distribuciones. Procedo a hacer el test U.

```
grupo_severo <- subset(dat, Severo == 1)$MWT1Best
grupo_no_severo <- subset(dat, Severo == 0)$MWT1Best
shapiro_severo <- shapiro.test(grupo_severo)
shapiro_no_severo <- shapiro.test(grupo_no_severo)

shapiro_severo
```

```
##
## Shapiro-Wilk normality test
##
## data: grupo_severo
## W = 0.90914, p-value = 0.0009761
```

```
shapiro_no_severo
```

```
##
## Shapiro-Wilk normality test
##
## data: grupo_no_severo
## W = 0.95643, p-value = 0.002403
```

```
dat$Severo <- factor(dat$Severo, levels = c(0, 1), labels = c("No Severo", "Severo"))
levene_test <- leveneTest(MWT1Best ~ Severo, data = dat)
levene_test
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 1  0.8899  0.347
##      147
```

```
wilcox_test_severidad <- wilcox.test(grupo_severo, grupo_no_severo,
                                     alternative = "two.sided",
                                     conf.int = TRUE,
                                     conf.level = 0.95)
wilcox_test_severidad
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: grupo_severo and grupo_no_severo
## W = 1454, p-value = 4.056e-05
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
##  -110.00000  -37.00006
## sample estimates:
## difference in location
##                -74.99996
```

```
mediana_severo <- median(grupo_severo, na.rm = TRUE)
mediana_no_severo <- median(grupo_no_severo, na.rm = TRUE)
dif_mediana <- mediana_severo - mediana_no_severo

cat("Mediana MWT1Best - Severo:", mediana_severo, "metros\n")
```

```
## Mediana MWT1Best - Severo: 369 metros
```

```
cat("Mediana MWT1Best - No Severo:", mediana_no_severo, "metros\n")
```

```
## Mediana MWT1Best - No Severo: 453 metros
```

```
cat("Diferencia de Medianas (Severo - No Severo):", dif_mediana, "metros\n")
```

```
## Diferencia de Medianas (Severo - No Severo): -84 metros
```

```
media_severo <- mean(grupo_severo, na.rm = TRUE)
media_no_severo <- mean(grupo_no_severo, na.rm = TRUE)
dif_media <- media_severo - media_no_severo

cat("Media - Severo:", media_severo, "metros\n")
```

```
## Media - Severo: 366.62 metros
```

```
cat("Media - No Severo:", media_no_severo, "metros\n")
```

```
## Media - No Severo: 432.9293 metros
```

```
cat("Diferencia de medias (Severo - No Severo):", dif_media, "metros\n")
```

```
## Diferencia de medias (Severo - No Severo): -66.30929 metros
```

La capacidad de ejercicio medida por la prueba MWT1Best es significativamente menor en los pacientes clasificados como Severo en comparación con los No Severo. La diferencia entre los valores se debe a que he calculado las medianas y medias sobre el test de Wilcoxon es sobre los rangos. Los p-value de ambos grupos es inferior a 0.05, entonces no podemos hablar de normalidad de las distribuciones y tenemos que usar un test no paramétrico como Wilcoxon. Existe una diferencia significativa entre los metros recorridos de los pacientes severos y no severos.

- h) En clave contraste de hipótesis, ¿podemos rechazar la hipótesis nula de igualdad de medias entre grupos? razona la respuesta.

Sí, podemos rechazar la hipótesis nula de igualdad de medias entre los grupos Severo y No Severo. Las varianzas entre grupos son homogéneas según el test Levene. Hay evidencia estadísticamente significativa de que los pacientes clasificados como Severo tienen una capacidad de ejercicio significativamente menor que los pacientes no severo según el test de Wilcoxon. Además existen diferencias importantes en los metros recorridos en los dos grupos.

- i) Contrasta si la proporción de pacientes con buena calidad de vida (SGRQ>50) es diferente entre hombres y mujeres en los pacientes sin hipertensión arterial. ¿Qué proporción de pacientes con buena calidad hay en cada grupo? Comenta los resultados del contraste.

```
sin_hipertension <- subset(dat, hypertension == "No")
```

```
hombres_buena <- sum(sin_hipertension$SGRQ > 50 & sin_hipertension$gender == "Masculino", na.rm = TRUE)
mujeres_buena <- sum(sin_hipertension$SGRQ > 50 & sin_hipertension$gender == "Femenino", na.rm = TRUE)
total_hombres <- sum(sin_hipertension$gender == "Masculino", na.rm = TRUE)
total_mujeres <- sum(sin_hipertension$gender == "Femenino", na.rm = TRUE)
```

```
prop_test_calidad <- prop.test(c(hombres_buena, mujeres_buena), c(total_hombres, total_mujeres), alternative = "two.sided")
prop_test_calidad
```



```
##
## 2-sample test for equality of proportions without continuity correction
##
## data: c(hombres_buena, mujeres_buena) out of c(total_hombres, total_mujeres)
## X-squared = 0.047388, df = 1, p-value = 0.8277
## alternative hypothesis: two.sided
## 95 percent confidence interval:
## -0.1476467 0.1847652
## sample estimates:
## prop 1 prop 2
## 0.3296703 0.3111111
```

Tenemos por una parte un 32,97% de hombres y un 31,11% de mujeres con buena calidad de vida (SGRQ>50), del total de hombres y mujeres sin hipertension, un p-value de 0.8277, lo que lleva a decir que no hay evidencia de que la proporción de pacientes es diferente entre hombres y mujeres.

- j) Se quiere testear una nueva intervención fisioterapéutica en estos pacientes para mejorar su función pulmonar. Se decide realizar una medición de la función pulmonar (medida por FEV1) de los pacientes después de seis meses de terapia (FEV1PRED_post). Plantea el contraste de hipótesis que realizarías para contrastar si los pacientes mejoraron significativamente, calcula y discute los resultados.

```
dat$diferencia_FEV1 <- dat$FEV1PRED_post - dat$FEV1PRED
diferencia_FEV1 <- na.omit(dat$diferencia_FEV1)

shapiro_diferencia <- shapiro.test(diferencia_FEV1)

t_test_pareado <- t.test(dat$FEV1PRED_post, dat$FEV1PRED, paired = TRUE, alternative = "greater")
list(shapiro_test = shapiro_diferencia, t_test = t_test_pareado)
```

```
## $shapiro_test
##
## Shapiro-Wilk normality test
##
## data: diferencia_FEV1
## W = 0.93641, p-value = 2.849e-06
##
##
## $t_test
##
## Paired t-test
##
## data: dat$FEV1PRED_post and dat$FEV1PRED
## t = 21.237, df = 149, p-value < 2.2e-16
## alternative hypothesis: true mean difference is greater than 0
## 95 percent confidence interval:
## 12.85213 Inf
## sample estimates:
## mean difference
## 13.93846
```

El análisis muestra una mejora significativa en la función pulmonar de los pacientes con EPOC tras la terapia fisioterapéutica, medida mediante FEV1PRED y FEV1PRED_post respectivamente. El Test t pareado evidencia que la diferencia promedio entre los valores antes y después de la terapia es 13.94 unidades de

volumen Espiratorio Forzado de la espirometría (%). Con un intervalo de confianza al 95% que asegura que esta mejora es al menos de 12.85 en la media. Además, el p-value obtenido indica una alta significancia estadística (p value muy, muy bajo). Por último si bien la prueba de distribución de diferencias no es normal, el test t. Tendríamos que decir formalmente que no se acepta la hipótesis nula y se acepta la alternativa de mejora importante de la función pulmonar de los pacientes.