Tests de hipótesis

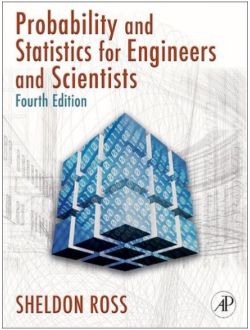
Estadística, Grado en Sistemas de Información

Constantino Antonio García Martínez

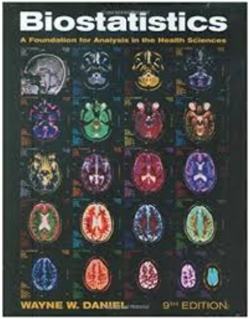
Universidad San Pablo Ceu

Contenidos

References



S. Ross. Introduction to Probability and Statistics for Engineers and Scientists. Chapter 8.



Wayne W. Daniel, Biostatistics: A Foundation for Analysis in the Health Sciences. Chapter 7.

Introducción

Ejemplo: Racismo en la selección de jurados

Durante los 60s-70s, se dieron casos de racismo en la elección de jurados populares. Supuestamente, la elección es al azar entre un listado de todos los ciudadanos. Sin embargo, se daban situaciones como que en una preselección de 80 posibles jurados solo 4 fuesen afroamericanos (de una población con $p = 0.5$ de afroamericanos). Las autoridades se defendían diciendo que era pura casualidad. ¿Es esto creíble?

1. Como en un juicio, asumimos la inocencia de las autoridades y aceptamos que realmente la selección es al azar. Entonces $X \sim \mathcal{B}(80, 0.5)$.
2. Bajo esta hipótesis, la probabilidad de obtener un jurado con solo 4 afroamericanos es
$$P(X \leq 4) = 1.4 \cdot 10^{-18}.$$
3. La probabilidad es tan baja que tenemos una **evidencia muy fuerte** contra la **hipótesis inicial**.

Podemos resumir el procedimiento como sigue

- Formulamos una hipótesis. Generalmente aceptamos como hipótesis de partida (**hipótesis nula**, H_0) lo contrario de lo que queremos probar (**hipótesis alternativa**, H_a) [Pensad en la presunción de inocencia].
- Buscamos un **estadístico de contraste** T que nos permita testear la veracidad de H_0.
- Calculamos el **p-valor**: bajo la H_0, la probabilidad de observar un estadístico tan extremo como el realmente observado. *Cuanto más pequeño sea el p-valor, más pequeña es la evidencia a favor de* H_0.
- Comparamos el p-valor con un umbral llamado **nivel de significancia** α. Si p-valor $< \alpha$, descartamos la hipótesis nula y diremos que la hipótesis ha sido **rechazada al** α **nivel de significancia**. Los valores más típicos de α son 0.05 y 0.01.

Errores tipos I y II

	H_0 es cierta	H_a es cierta
Se escogió H_0	No hay error	Error de tipo II/Falso negativo (β)
Se escogió H_a	Error de tipo I/Falso positivo (α)	No hay error

Tenemos que

$$P(\text{escoger } H_a | H_0 \text{ es cierta}) = \alpha$$

$$P(\text{escoger } H_0 | H_a \text{ es cierta}) = \beta$$

Se denomina **potencia del contraste** al valor $1 - \beta$, esto es, a la probabilidad de escoger H_a cuando esta es cierta.

Tests para medias

Tests para medias

Varianza conocida

Con varianza conocida
Asumimos que $X_1, X_2, \cdots X_n$ provienen de un muestreo con reemplazamiento de una población Normal o bien, si la población no es normal que el muestreo es grande $n \geq 30$.

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

Rechazaremos la hipótesis nula si $\hat{X} - \mu_0$ es demasiado grande o pequeño. Dado que que bajo la hipótesis nula $Z = \frac{\hat{X}-\mu_0}{\sigma/\sqrt{n}}$ tiene distribución Normal, rechazaremos la hipótesis nula si el p-valor es menor que α:

$$P\left(|Z| > \frac{\hat{x}-\mu_0}{\sigma/\sqrt{n}}\right) < \alpha.$$

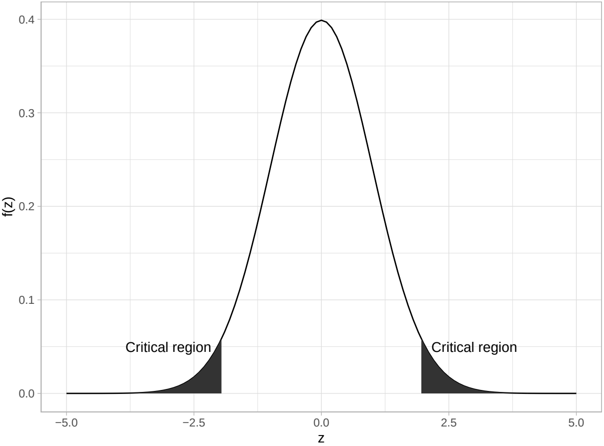
De aquí se observa que rechazamos la hipótesis nula si

$$\frac{\hat{x}-\mu_0}{\sigma/\sqrt{n}} > z_{\alpha/2} \qquad \text{o bien} \qquad \frac{\hat{x}-\mu_0}{\sigma/\sqrt{n}} < -z_{\alpha/2},$$

que definen una región de rechazo llamada **zona crítica**.

Zona crítica:

$$\left| \frac{\hat{x} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\alpha/2}.$$



Ejercicio:
Una señal de valor μ se envía desde A hasta B. Cuando se recibe en B, la señal ha sido distorsionada por un error de media 0 y desviación estándar 2. B sospecha que la señal de hoy es $\mu \neq 8$. Testea esta hipótesis al 95 % de confianza (5 % de significancia) si la misma señal se ha enviada 5 veces y el valor medio recibido en X es 9.5.

```
mu_sample = 9.5;  mu_real = 8; sigma = 2; n = 5

Z = (mu_sample - mu_real) / (sigma / sqrt(n))
if (Z > qnorm(1 - 0.025) | Z < qnorm(0.025)) {
  cat("Reject H0\n")
} else {
  cat("Accept H0\n")
}

## Accept H0

p_value = 2 * min(pnorm(Z), 1 - pnorm(Z))
print(p_value)

## [1] 0.09353251
```

Prueba de una sola cola

Si testeamos

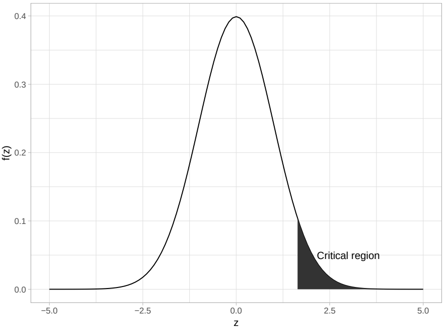
$$H_0 : \mu < \mu_0 \text{ versus } H_a : \mu > \mu_0.$$

tenemos un prueba de una sola cola, ya que sólo rechazaremos H_0 si

$$P(X - \mu_0 > c) = \alpha.$$

Procediendo como antes llegamos a que rechazaremos H_0 si

$$\frac{\sqrt{n}}{\sigma}(\hat{X} - \mu_0) > z_\alpha$$



Test para medias con varianza conocida

TABLE 8.1 $X_1, \ldots, X_n$ *Is a Sample from a* $\mathcal{N}(\mu, \sigma^2)$ *Population*
σ^2 *Is Known,* $\overline{X} = \sum_{i=1}^{n} X_i/n$

H_0	H_1	Test Statistic TS	Significance Level α Test	p-Value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $\lvert TS \rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t \rvert\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $TS > z_\alpha$	$P\{Z \geq t\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $TS < -z_\alpha$	$P\{Z \leq t\}$

Z is a standard normal random variable.

Tests para medias

Varianza desconocida

Test para medias con varianza desconocida

Para poblaciones Normales con varianza desconocida consideremos el test

$$H_0 : \mu = \mu_0 \text{ versus } H_a : \mu \neq \mu_0.$$

De forma similar a los ICs, explotamos que $T = \sqrt{n}\frac{\hat{X}-\mu_0}{\hat{S}}$ tiene distribución T-Student con $n-1$ grados de libertad. Se obtiene que debemos rechazar H_0 si

$$\left|\sqrt{n}\frac{\hat{X}-\mu_0}{\hat{S}}\right| > t_{\alpha/2;n-1}$$

y aceptarla en otro caso.

Procedimiento de forma similar se podrían hallar las zonas de rechazo para los tests unilaterales.

Ejercicio:
Entre pacientes con un alto nivel de colesterol (> 220 mg/dL), se ha probado una nueva droga diseñada para reducir el nivel de colesterol. Un grupo de 50 voluntarios toma la droga durante 1 mes, produciéndose una reducción media de 14.8 con una desviación típica muestral de 6.4. ¿Qué conclusiones podemos obtener (al 95 %)? Datos en cholesterol.csv.

Test para medias con varianza desconocida

$$H_0 : \mu < 0 \text{ versus } H_a : \mu > 0.$$

```
cholesterol = read.csv("data/cholesterol.csv")$cholesterol
t.test(cholesterol, alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  cholesterol
## t = 16.352, df = 49, p-value < 2.2e-16
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  13.28256      Inf
## sample estimates:
## mean of x
##      14.8
```

Ejercicio:

Nakamura et al.[1] estudiaron pacientes con roturas del ligamento colateral tibial (medial collateral ligament, MCL) y el ligamento cruzado anterior (anterior cruciate ligament, ACL) y pacientes con lesiones combinadas de ACL y MCL. Todos los pacientes fueron tratados por un fisio en el centro del estudio. Una de las variables del estudio era el tiempo en días entre que se produjo la lesión y la primera MRI. Los datos se encuentran disponibles en *injuries.csv* ([2]). Queremos testear si el número medio de días para los pacientes de "MCL + ACL" es inferior a 15 días.

[1] N. Nakamura, S. Horibe, Y. Toritsuka, T. Mitsuoka, H. Yoshikawa, and K. Shino, "Acute Grade III Medial Collateral Ligament Injury of the Knee Associated with AnteriorCruciate Ligament Tear", American Journal of Sports Medicine, 31 (2003), 261–267.

[2] los datos para esta clase son reales (del paper citado), los demás datos son sintéticos para hacer el ejercicio más interesante

Tenemos que asumir normalidad

```
data = read.csv("data/injuries.csv")
days = data[data$injury == "MCL+ACL", "days"]
days = na.omit(days)
cat("mean = ", mean(days), ", sd = ", sd(days), ", n = ",  length(days), "\n")

## mean =  13.29412 , sd =  8.88654 , n =  17

t.test(days, mu = 15, alternative="less")

##
##  One Sample t-test
##
## data:  days
## t = -0.79148, df = 16, p-value = 0.2201
## alternative hypothesis: true mean is less than 15
## 95 percent confidence interval:
##      -Inf 17.05703
## sample estimates:
## mean of x
##  13.29412
```

Test para medias con varianza desconocida

Tenemos que asumir normalidad

```r
data = read.csv("data/injuries.csv")
days = data[data$injury == "MCL+ACL", "days"]
days = na.omit(days)
cat("mean = ", mean(days), ", sd = ", sd(days), ", n = ",  length(days), "\n")
```

```
## mean =  13.29412 , sd =  8.88654 , n =  17
```

```r
t.test(days, mu = 15, alternative="less")
```

```
##
##  One Sample t-test
##
## data:  days
## t = -0.79148, df = 16, p-value = 0.2201
## alternative hypothesis: true mean is less than 15
## 95 percent confidence interval:
##      -Inf 17.05703
## sample estimates:
## mean of x
##  13.29412
```

¿Es realista asumir normalidad?

Tests para medias con varianza desconocida y no-normalidad

```r
bp = read.csv("data/bp.csv")$blood_pressure
cat("mean = ", mean(bp), ", si-sd = ", sd(bp), ", n = ", length(bp))

## mean =  146 , si-sd =  27 , n =  157

t.test(bp, mu = 140, alternative = "greater")$p.value

## [1] 0.003013078

# Another method, using the Z distribution
Z = sqrt(length(bp)) * (mean(bp) - 140) / sd(bp)
p_value =  1 - pnorm(Z)
cat("Stat: ", Z, " p-value: ", p_value)

## Stat:  2.784436  p-value:  0.002681041
```

Test para medias con varianza desconocida

TABLE 8.2 $X_1, \ldots, X_n$ *Is a Sample from a* $\mathcal{N}(\mu, \sigma^2)$ *Population*
σ^2 *Is Unknown,* $\overline{X} = \sum_{i=1}^{n} X_i/n \; S^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2/(n-1)$

H_0	H_1	Test Statistic TS	Significance Level α Test	p-Value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $\lvert TS\rvert > t_{\alpha/2,\, n-1}$	$2P\{T_{n-1} \geq \lvert t\rvert\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $TS > t_{\alpha,\, n-1}$	$P\{T_{n-1} \geq t\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $TS < -t_{\alpha,\, n-1}$	$P\{T_{n-1} \leq t\}$

T_{n-1} is a t-random variable with $n-1$ degrees of freedom: $P\{T_{n-1} > t_{\alpha,\, n-1}\} = \alpha$.

Tests para comparar medias

Tests para comparar medias

Varianzas desconocidas

Comparación de medias de poblaciones normales con varianzas desconocidas pero iguales

Para poblaciones Normales con varianzas desconocidas pero iguales consideramos el test

$$H_0 : \mu_x = \mu_y \text{ versus } H_a : \mu_x \neq \mu_y$$

En este caso, explotamos que

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{(1/n + 1/m)}} \sim \mathcal{T}_{n+m-2},$$

donde S_p^2 es la **varianza combinada**

$$\hat{S}_p^2 = \frac{(n-1)\hat{S}_1^2 + (m-1)\hat{S}_2^2}{n+m-2}.$$

Se obtiene que debemos rechazar H_0 si

$$|T| > t_{\alpha/2;n+m-2}$$

y aceptarla en otro caso.

Comparación de medias de poblaciones normales con varianzas desconocidas pero iguales

Ejemplo:
El propósito del estudio de Tam et al. [3] era investigar maniobras en sillas de ruedas en individuos con lesiones en la parte inferior de la médula espinal (SCI) e individuos de control (C). Durante las maniobras se midió la presión bajo las tuberosidades isquiáticas. Asumiendo que las poblaciones son normales y la varianza es igual en las poblaciones, ¿podemos concluir que los individuos de control muestran menor presión que los pacientes de SCI? Datos en *sci.csv*.

[3] ERIC W. TAM, ARTHUR F. MAK, W AI NGA LAM, JOHN H. EVANS, AND YORK Y. CHOW, "Pelvic Movement and Interface Pressure Distribution During Manual Wheelchair Propulsion", Archives of Physical Medicine and Rehabilitation, 84 (2003), 1466–1472.

Comparación de medias de poblaciones normales con varianzas desconocidas pero iguales

Ejemplo:
El propósito del estudio de Tam et al. [3] era investigar maniobras en sillas de ruedas en individuos con lesiones en la parte inferior de la médula espinal (SCI) e individuos de control (C). Durante las maniobras se midió la presión bajo las tuberosidades isquiáticas. Asumiendo que las poblaciones son normales y la varianza es igual en las poblaciones, ¿podemos concluir que los individuos de control muestran menor presión que los pacientes de SCI? Datos en *sci.csv*.

¡Ojo! Las hipótesis son

$$H_0 : \mu_c > \mu_{sci} \text{ versus } \mu_c < \mu_{sci}$$

[3] ERIC W. TAM, ARTHUR F. MAK, W AI NGA LAM, JOHN H. EVANS, AND YORK Y. CHOW, "Pelvic Movement and Interface Pressure Distribution During Manual Wheelchair Propulsion", Archives of Physical Medicine and Rehabilitation, 84 (2003), 1466–1472.

Comparación de medias de poblaciones normales con varianzas desconocidas pero iguales

```
df = read.table("data/sci.tsv")
# Arrange data.frame as list to compute summary stats. Note the -1!
dat_list = list('control' = as.numeric(df[1, -1]), 'sci' = as.numeric(df[2, -1]))
summ_list = lapply(dat_list, function(x) {
  list('mean' = mean(x), 'sd' = sd(x), 'n' = length(x))
})

# Let's use with to illustrate its use
s2_pooled = with(summ_list,
     ((control$n - 1) * control$sd ^ 2 + (sci$n - 1) * sci$sd ^ 2) / (
         control$n + sci$n - 2
     ))
stat = with(summ_list,
 (control$mean - sci$mean) / (sqrt(s2_pooled) * sqrt(1 / control$n + 1 / sci$n))
)
p_value = pt(stat, df = summ_list$control$n + summ_list$sci$n - 2)
cat("Stat: ", stat, " p-value: ", p_value, "\n")

## Stat:  -0.5693629  p-value:  0.2880734
```

Comparación de medias de poblaciones normales con varianzas desconocidas pero iguales

```r
with(dat_list,
     t.test(control, sci, alternative = "less", var.equal = TRUE)
)
```

```
##
##  Two Sample t-test
##
## data:  control and sci
## t = -0.56936, df = 18, p-value = 0.2881
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 14.31935
## sample estimates:
## mean of x mean of y
##     126.1     133.1
```

Comparación de medias de poblaciones normales con varianzas desconocidas y distintas

Para poblaciones Normales con varianzas desconocidas. consideramos el test

$$H_0 : \mu_x = \mu_y \text{ versus } H_a : \mu_x \neq \mu_y.$$

Se considera el estadístico

$$T' = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\hat{S}_x^2}{n} + \frac{\hat{S}_y^2}{m}}}$$

junto con el valor crítico t':

$$t'_{1-\alpha/2} = \frac{w_1 t_1 + w_2 t_2}{w1 + w2}$$

con $w_1 = \hat{s}_1^2/n_1$, $w_2 = \hat{s}_2^2/n_2$, $t_1 = t_{n_1-1;1-\alpha/2}$, $t_2 = t_{n_2-1;1-\alpha/2}$ que es la llamada **aproximación de Welch**.

Ejercicio:

Dernellis and Panaretou [4] compararon sujetos con hipertensión y de control. Una de las variables era el índice de rigidez de la arteria aorta. En 15 pacientes con hipertensión, el índice medio fue de 19.16 ± 5.29. En los 30 pacientes de control la rigidez media fue de 9.53 ± 2.69. ¿Existen diferencias significativas? Datos en hypertension.csv.

[4] J. DERNELLIS and M. PANARETOU, "Effects of Thyroid Replacement Therapy on Arterial Blood Pressure in Patients with Hypertension and Hypothyroidism", American Heart Journal, 143 (2002), 718–724.

Comparación de medias de poblaciones normales con varianzas desconocidas y distintas

Asumiendo normalidad...

```
# Agggh, spanish-like data!
dat = read.csv2("data/hypertension.csv")
# Let's use split to illustrate it
dat = split(dat, dat$group) %>% sapply(function(x) x$rigidity)
t.test(dat$hypertension, dat$control)
```

```
##
##  Welch Two Sample t-test
##
## data:  dat$hypertension and dat$control
## t = 6.6346, df = 17.711, p-value = 3.426e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   6.576974 12.683026
## sample estimates:
## mean of x mean of y
##     19.16      9.53
```

Tests para comparar medias

TABLE 8.4 $X_1, \ldots, X_n$ *Is a Sample from a* $\mathcal{N}(\mu_1, \sigma_1^2)$ *Population;* $Y_1, \ldots, Y_m$ *Is a Sample from a* $\mathcal{N}(\mu_2, \sigma_2^2)$ *Population*

The Two Population Samples Are Independent

to Test

$H_0 : \mu_1 = \mu_2$ versus $H_0 : \mu_1 \neq \mu_2$

Assumption	Test Statistic TS	Significance Level α Test	p-Value if $TS = t$
σ_1, σ_2 known	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}$	Reject if $\lvert TS\rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t\rvert\}$
$\sigma_1 = \sigma_2$	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}\sqrt{1/n+1/m}}$	Reject if $\lvert TS\rvert > t_{\alpha/2,\, n+m-2}$	$2P\{T_{n+m-2} \geq \lvert t\rvert\}$
n, m large	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2/n+S_2^2/m}}$	Reject if $\lvert TS\rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t\rvert\}$

Test T-Student para datos apareados

Ejemplo:
Queremos comprobar si la instalación de un nuevo filtro afecta al kilometraje de los coches. Para testearlo medimos los kilómetros/litro antes y después de su instalación. Por tanto, los datos son n pares (X_i, Y_i) correspondientes a n coches.

La dificultad es que los datos $X_1, X_2, \cdots, X_n$ e $Y_1, Y_2, \cdots, Y_n$ **no son independientes**. Sin embargo, podemos crear la variable

$$W_i = X_i - Y_i$$

y testear

$$H_0 : \mu_w = 0 \text{ versus } H_a : \mu_w \neq 0.$$

Usando el test-T presentado anterior para media y varianza desconocidas, rechazaremos H_0 si

$$\left| \frac{\bar{W}}{\hat{S}_w/\sqrt{n}} \right| > t_{\alpha/2;n-1}.$$

Ejemplo:

John M. Morton et al. [5] examinaron la función de la vesícula biliar antes y después de una cirugía para detener el reflujo. Los autores midieron la funcionalidad de la vesícula biliar calculando la fracción de eyección de la vesícula biliar (GBEF) antes y después de la operación, cuyo objetivo de la fundoplicatura es aumentar la GBEF. ¿Hay evidencia para concluir que la operación aumenta el GBEF? Datos en *gbef.csv*.

[5] J. MORTON, S. BOWERS, T. LUCKTONG, S. MATTAR, W. BRADSHAW , K. BEHRNS, M. ORUDA, C. HERBST , W. MCCARTNEY , R. HALKAR, C. SMITH, and T.FARRELL, "Gallbladder Function Before and After Fundoplication", Journal of Gastrointestinal Surgery, 6 (2002), 806–811.

Ejemplo:
John M. Morton et al. [5] examinaron la función de la vesícula biliar antes y después de una cirugía para detener el reflujo. Los autores midieron la funcionalidad de la vesícula biliar calculando la fracción de eyección de la vesícula biliar (GBEF) antes y después de la operación, cuyo objetivo de la fundoplicatura es aumentar la GBEF. ¿Hay evidencia para concluir que la operación aumenta el GBEF? Datos en *gbef.csv*.

Asumiendo normalidad, y teniendo en cuenta que las hipótesis son

$$H_0 : \mu_d < 0 \qquad H_a : \mu_d > 0 \qquad (d_i = \text{post} - \text{preop})$$

[5] J. MORTON, S. BOWERS, T. LUCKTONG, S. MATTAR, W. BRADSHAW , K. BEHRNS, M. ORUDA, C. HERBST , W. MCCARTNEY , R. HALKAR, C. SMITH, and T.FARRELL, "Gallbladder Function Before and After Fundoplication", Journal of Gastrointestinal Surgery, 6 (2002), 806–811.

Test T-Student para datos apareados

```
df = read.csv("data/gbef.csv", header = FALSE)
# compute differences
postop = df[2, -1] %>% as.numeric
preop = df[1, -1] %>% as.numeric
d = postop - preop
cat("mean = ", mean(d), "\t sd = ", sd(d), "n = ", length(d), "\n")

## mean =  18.075   sd =  32.68169 n =  12

t.test(postop, preop, paired = TRUE, alternative = "greater")

##
##  Paired t-test
##
## data:  postop and preop
## t = 1.9159, df = 11, p-value = 0.04086
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.131919      Inf
## sample estimates:
## mean of the differences
##                  18.075

# Accept Ha
```

Tests para la varianza

Tests para la varianza

Tests para la varianza en una población normal

Tests para la varianza en una población normal

Sea la hipótesis

$$H_0 : \sigma^2 = \sigma_0^2 \text{ versus } H_a : \sigma^2 \neq \sigma_0^2.$$

Empleando el hecho que bajo H_0 $(n-1)\frac{\hat{S}^2}{\sigma_0^2} \sim \chi^2_{n-1}$ concluimos que tenemos que

$$\text{Aceptar si } \chi^2_{1-\alpha/2;n-1} \leq (n-1)\frac{\hat{S}^2}{\sigma_0^2} \leq \chi^2_{\alpha/2;n-1}$$

$$\text{Rechazar en otro caso}$$

De forma análoga, podríamos derivar los test de una sola cola.

Ejemplo:

El objetivo de Wilkins et al. [6] fue medir la efectividad de la hormona de crecimiento humano recombinante (rhGH) en niños con quemaduras. 16 sujetos recibieron inyecciones diarias de rhGH. Al inicio, los investigadores querían conocer los niveles del factor de crecimiento de la insulina (IGF-I) antes de la administración de rhGH. La varianza muestral S^2 de los niveles de IGF-I (medida en ng/ml) fue de 628.8844. ¿Podemos concluir que la varianza de la población no es 600?

[6] J. WILKINS, O. SUMAN, D. BENJAMIN, and D. HEMDON, "Comparison of Self-Reported and Monitored Compliance of Daily Injection of Human Growth Hormone in Burned Children", Burns, 29 (2003), 697–701.

Asumiendo normalidad...

```
sample_var = 628.8844
stat = (16 - 1) * sample_var / 600
critical_values = c(qchisq(0.025, 16 - 1), qchisq(0.975, 16 - 1))
if (stat <= critical_values[1] || stat >= critical_values[2]) {
  cat("Stat", stat, "Reject H0\n")
} else {
  cat("Stat", stat, "Accept H0\n")
}

## Stat 15.72211 Accept H0

#  we are unable to conclude that the population variance is not 600.
```

Tests para la igualdad de varianzas entre poblaciones normales

Tests para la igualdad de varianzas entre poblaciones normales

Sea la hipótesis

$$H_0 : \sigma_x^2 = \sigma_y^2 \text{ versus } H_a : \sigma_x^2 \neq \sigma_y^2.$$

Empleando el hecho que bajo H_0 $\hat{S}_x^2/\hat{S}_y^2 \sim F_{n-1,m-1}$ concluimos que tenemos que

$$\text{Aceptar si } F_{1-\alpha/2;n-1,m-1} \leq \frac{\hat{S}_x^2}{\hat{S}_y^2} \leq F_{\alpha/2;n-1,m-1}$$

$$\text{Rechazar en otro caso}$$

Ejemplo:
Borden et al. [7] compararon técnicas de reparación de meniscos utilizando muestras de rodilla de cadáver. Una de las variables de interés fue la carga de fallo (en Newtons) para rodillas fijadas con la técnica FasT-FIX (grupo 1) y el método de sutura vertical (grupo 2). Cada técnica se aplicó a seis muestras. La desviación estándar muestral para el método FasT-FIX fue de 30.62, y la desviación estándar muestral para el método de sutura vertical fue de 11.37. ¿Podemos concluir que, en general, la varianza de la carga de fallo es mayor para la técnica FasT-FIX que para el método de sutura vertical?

[7] P. BORDEN, J. NYLAND, D. CABORN, and D. PIENOWSKI, "Biomechanical Comparison of the FasT-FIX Meniscal Repair Suture System with Vertical Mattress and Meniscus Arrows", American Journal of Sports Medicine, 31 (2003), 374–378.

Tests para la igualdad de varianzas entre poblaciones normales

Asumiendo normalidad y dado que las hipótesis son

$$H_0 : \sigma_1^2 < \sigma_2^2 \qquad H_a : \sigma_1^2 > \sigma_2^2$$

```
dat = read.csv('data/knees.csv')
dat = split(dat, dat$class)
var.test(dat$group_1$data, dat$group_2$data, alternative = "greater")

##
##  F test to compare two variances
##
## data:  dat$group_1$data and dat$group_2$data
## F = 7.2525, num df = 5, denom df = 5, p-value =
## 0.02425
## alternative hypothesis: true ratio of variances is greater than 1
## 95 percent confidence interval:
##  1.436051      Inf
## sample estimates:
## ratio of variances
##           7.252528

# The failure load variability is higher when using the FasT-
# FIX method than the vertical suture method.
```