William Andrés Gómez
Juan Manuel Bermúdez

PRUEBAS DE HIPOTESIS

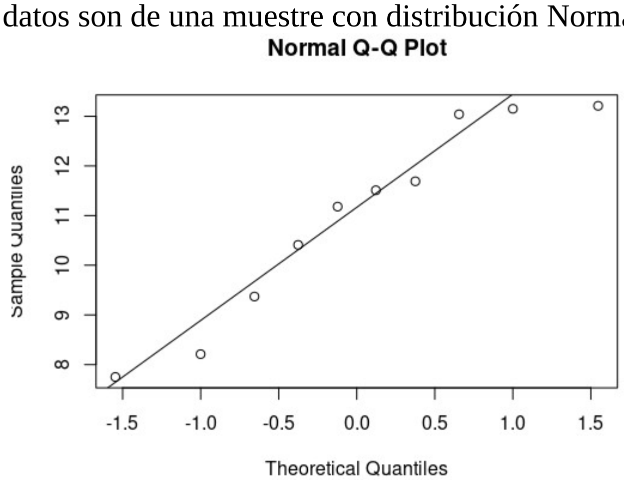
Quiz#3

Problems: 2.5 2.8 2.9 2.20 2.22 2.26 2.30 2.31 2.32 2.33 2.35

#2.5 Considere el resultado siguiente:
Test de mu=30 vs no=30
La desviación estándar asumida=1.2
N=16
Mean=31.2000
SE_Mean=0.3000
CI=0.95
Z=?
P=?

```
> qnorm(0.975)
Z=  1.959964
> pvalue <- 2*(1-pnorm(qnorm(0.975)))
> pvalue
P= 0.05
```

#2.8 Considere la siguiente muestra de datos:
[9.37, 13.04, 11.69, 8.21, 11.18, 10.41, 13.15, 11.51, 13.21, 7.75]
Se puede asumir que estos datos son de una muestre con distribución Normal.



Se puede asumir que los datos tienen distribucion normal, a pesar de que no se encuentran completamente sobre la recta. Pero apesar de ser pocos datos tienen un distribucion aproximadamente normal.

¿Existe evidencia para soportar que la Media de la población es 10?

```
datos <- c(9.37, 13.04, 11.69, 8.21, 11.18, 10.41, 13.15, 11.51, 13.21, 7.75)
qqnorm((datos))
qqline(datos)
```

n=length((datos))
qt(0.95,n-1)
t.test(datos, mu=10)

> qt(0.95,n-1)
[1] 1.833113
> t.test(datos, mu=10)
One Sample t-test

data: datos
t = 1.5102, df = 9, p-value = 0.1653
alternative hypothesis: true mean is not equal to 10
95 percent confidence interval:
9.525956 12.378044
sample estimates:
mean of x
10.952
R//: Ya que la estadística t=1,51 es menor que el punto crítico tc=1,833, entonces no podemos rechazar la hipótesis nula. De la misma forma que el pvalue de 0,165 es mucho mayor que alfa=0,05. En conclusión, no hay evidencia suficiente para afirmar que la media de los datos sea igual a 10

#2.9 Un programa de computador ha producido el siguiente resultado para una prueba de hipotesis:
Diferencia en la media de las muestras: 2.35
Grados de libertad= 18
SE de la diferencia de las medias muestrales=?
Test estadistico: t_0= 2.01
P-value=0.0298

a)
b) Prueba Bilateral
c) Como el Pvalue es menor que alfa, se rechaza la hipotesis nula.

#2.20 La duración de una bebida carbonatada es de interes. 10 botellas son seleccionadas aleatoriamente y testeadas, obteniendo los siguientes resultados.
Days=[108, 124, 124, 106, 115, 138, 163, 159, 134, 139]

a) Nos gustaría demostrar que la vida media excede los 120 días. Establezca una hipotesis para invesitigar esta afirmación.
Ho: mu<= 120
H1: mu>120

b) Pruebe la hipoteisis utilizando alfa=0.01. Cuales son sus conclusiones.

>days=c(108, 124, 124, 106, 115, 138, 163, 159, 134, 139)
>t.test(days,mu=120,alternative=c("greater"), conf.level=0.99)
estadistica: t = 1.7798
p-value = 0.05441
>qt(0.99,9)

2.821438

Debido a que el valor critico 2,82 es mayor que la estadistica 1,77 no se puede rechazar la hipotesis nula. Por tanto no hay evidencia suficiente para afirmar que la media excede los 120 dias.

c) Encuentre el P-value para el test

El p-value fue de 0,05441 > alfa=0,01

Como el p-value no es menor que alfa no hay evidencia suficiente para rechazar la hipótesis nula.

#2.22 El tiempo para reparar un instrumento electrónico es una variable aleatoria distribuida de forma normal. EL tiempo de reparación para 16 instrumentos seleccionados aleatoriamente es:
Hours=[159, 224, 222, 149, 280, 379, 362, 260, 101, 179, 168, 485, 212, 264, 250, 170]
a) Te gustaría saber si el tiempo medio de reparación excede 225 horas. Establezca una hipotesis apropiada para invesitgar esta afirmación.
Ho: mu<=225
H1: mu>225

b) Pruebe la hipotesis con alfa=0.05, Cuales sons sus conclusiones.

```
Hours=c(159, 224, 222, 149, 280, 379, 362, 260, 101, 179, 168, 485, 212, 264, 250, 170)
n=length((Hours))
t.test(Hours, mu=225,alternative=c("greater") )
qt(0.95, n-1)
```

```
> t.test(Hours, mu=225,alternative=c("greater") )
	One Sample t-test
data:  Hours
t = 0.66852, df = 15, p-value = 0.257
alternative hypothesis: true mean is greater than 225
95 percent confidence interval:
 198.2321      Inf
sample estimates:
mean of x
   241.5
> qt(0.95, n-1)
[1] 1.75305
```

Ya que la estadística t=0,668 es menor que el punto critico tc=1,75, no cae en la región de rechazo y no se puede rechazar la hipótesis nula. Entonces se puede decir que no hay suficiente evidencia para afirmar que la media de horas sea mayor a 225

c) Encuentre un P-value para la prueba
El p-value es = 0,257 el cual no es menor a alfa=0.05 y por tanto no es posible rechazar la hipotesis nula. Es decir, la media no es mayor a 225, o al menos con la muestra que tenemos no hay suficiente evidencia para demostrarlo

#2.26 Los siguientes son los tiempos de combustión de dos tipos de bengalas químicas. Los ingenieros están interesados tanto en la media como en la varianza de los tiempos de combustión.

a) Pruebe la hipótesis de que las dos varianzas son iguales. Utilice alfa=0.05

Ho: Cociente de varianzas igual a 1
H1: Cociente de varianzas diferente a 1

```
TypeI=c(65, 81, 57, 66, 82, 82, 67, 59, 75, 70)
TypeII=c(64, 71, 83, 59, 65, 56, 69, 74, 82, 79)
var.test(TypeI, TypeII)

> var.test(TypeI, TypeII)

        F test to compare two variances
data:  TypeI and TypeII
F = 0.97822, num df = 9, denom df = 9, p-value = 0.9744
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.2429752 3.9382952
sample estimates:
ratio of variances
         0.9782168
> qf(0.975,n-1,n-1)
[1] 4.025994
> qf(0.025,n-1,n-1)
[1] 0.2483859
```

De esta prueba obtenemos una estadistica F=0.978. Si obervamos los puntos criticos calulados para la distribución F, podemos ver que son (0.248 , 4.0259), ya que la estadística F esta dentro del intervalo de los puntos críticos, NO podemos rechazar la hipotesis nula, y por ende las varianzas SI SON IGUALES.

b) En base al resultado anterior, pruebe la hipótesis de que los tiempos de combustión de los 2 tipos de bengalas son iguales. Utilice alfa=0.05. ¿Cuál es el P-value para este test?
Ho: mu1=mu2
H1: mu1 ≃mu2

```
> t.test(TypeI, TypeII, var.equal=TRUE)
        Two Sample t-test
data:  TypeI and TypeII
t = 0.048008, df = 18, p-value = 0.9622
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -8.552441  8.952441
sample estimates:
mean of x mean of y
     70.4      70.2
> qt(0.975, n-1)
[1] 2.262157
```

Ya que la estadística t=0.048 esta dentro de la región de No rechazo, No se puede rechazar la hipótesis nula de que las medias son iguales. Y por tanto, existe suficiente evidencia para decir que las medias de ambas muestras SON IGUALES.

#2.30 Cubiertas plásticas para celular son fabricadas por un proceso de inyección en un molde. Se cree que el tiempo que la parte es dejada enfriar antes de ser retirada del molde puede influir en la calidad final del producto. Después de fabricadas las piezas son inspeccionadas y asignadas un valor entre 1 y 10 basado en su apariencia, donde 1 es completamente defectuoso y 10 es perfecta. Un experimento fue llevado a cabo utilizando 2 tiempos de enfriamiento 10 segundos y 20 segundos.

10_seconds=[1,2,1,3,5,1,5,2,3,5,3,6,5,3,2,1,6,8,2,3]
20_seconds=[7,8,5,9,5,8,6,4,6,7,6,9,5,7,4,6,8,5,8,7]

a) Existe evidencia que demuestre que entre más tiempo se deja enfriar la pieza , se resulta en menos defectos?. Utilice alfa=0.05

*Primero hace test de varianzas:
Ho: Cociente de varianzas igual a 1
H1: Cociente de varianzas diferente a 1

```
> var.test(ten_seconds, twenty_seconds)
        F test to compare two variances
data:  ten_seconds and twenty_seconds
F = 1.7011, num df = 19, denom df = 19, p-value = 0.2559
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6733205 4.2977738
sample estimates:
ratio of variances
        1.701111
> qf(0.975,n-1,n-1)
[1] 2.526451
> qf(0.025,n-1,n-1)
[1] 0.3958122
```

Como la estadística F cae dentro de la región de No rechazo, no rechazamos la Hipótesis nula y decimos que las varianzas SON IGUALES.

*Prueba de diferencia de medias
Ho: mu1=mu2
H1: mu1 ≃mu2

```
> t.test(ten_seconds, twenty_seconds, var.equal=TRUE)
        Two Sample t-test
data:  ten_seconds and twenty_seconds
```

t = -5.5696, df = 38, p-value = 2.217e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.294935 -2.005065
sample estimates:
mean of x mean of y
    3.35   6.50
> qt(0.025, n-1)
[1] - 2.093024

Podemos ver que la estadística t es mucho menor que el punto critico inferior. Por tanto rechazamos la hipótesis nula ya que existe suficiente evidencia para decir que las medias No son iguales. En otras palabras, las medias son diferentes. Esto quiere decir que en estas muestras si existe una diferencia estadísticamente significativa para decir que entre mayor tiempo se dejen las piezas enfriar, mejor sera su calidad.

b) Cual es el P-value para el test anterior?
El p-value = 2.217e-06 es mucho menor que alfa=0.05. Por lo que rechazamos la hipótesis nula.
d) Revise la presuncion de Normalidad para este experimento.

#2.31 Observaciones de la uniformidad del gravado sobre silicona se evaluó en un experimento. Los datos son los siguientes.

Etch=c(5,34, 6,00, 5,97, 5,25, 6,65, 7,55, 7,35, 6,35, 4,76, 5,54, 5,44, 4,61, 5,98, 5,62, 4,39, 6,00, 7,25, 6,21, 4,98, 5,32)

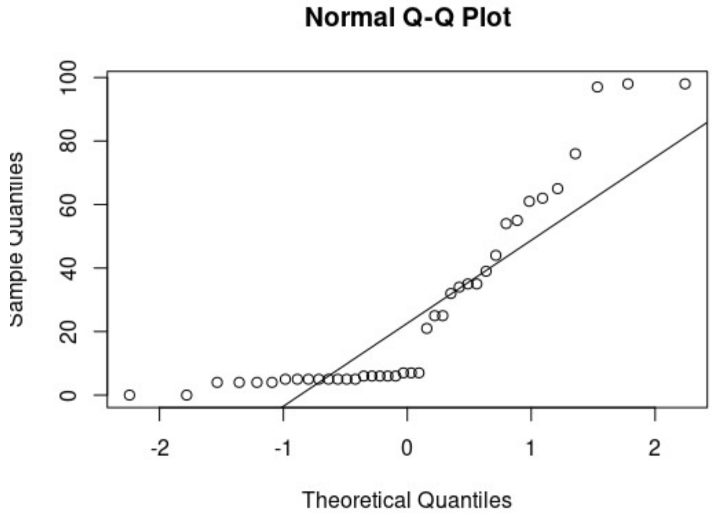
b) Pruebe la hipotesis de que la varianza=1. Utilice alfa=0.05. Cuales son sus conclusiones?

Ho: var=1
H1: var≃1
library(EnvStats)
gl<-length(Etch)-1
PHV<-varTest(Etch,alternative = "two.sided",sigma.squared=1)
# Valores cr?ticos
qchisq(0.025,gl)
qchisq(0.975,gl)
PHV$statistic
> qchisq(0.025,gl)
[1] 23.65432
> qchisq(0.975,gl)
[1] 58.12006
> PHV$statistic
Chi-Squared
  34243.78

Dado que la estadistica Chi=34243,78 cae en la zona de recahzo, Rechazamos la hipótesis nula de que la varianza de la muestra es igual a 1.

c) Discuta la suposición de Normalidad y el papel ene este problema.

Ay un problema serio al asumir la normalidad en este caso, lo que hace que la prueba T no sea correcta. Como se observa a continuación, los datos no tienen distribución normal.
d) Revise la normalidad construyendo una grafica de probabilidad de normalidad. Cuales son sus conclusiones.



#2.32 El diametro de una bola de un rodamieno fue medido por 12 inspectores, cada uno usando 2 diferentes calibradores.

a) Existe una diferencia significativa entre las medias ? Utilice alfa=0.05
* Prueba para saber si las varianzas son iguales o diferentes:
Ho: Cociente de varianzas igual a 1
H1: Cociente de varianzas diferente a 1

```
caliper1=c(0.265, 0.265, 0.266, 0.267, 0.267, 0.265, 0.267, 0.267, 0.265, 0.268, 0.268, 0.265)
caliper2=c(0.264, 0.265, 0.264, 0.266, 0.267, 0.268, 0.264, 0.265, 0.265, 0.267, 0.268, 0.269)
var.test(caliper1, caliper2)
n=length((caliper2))
qf(0.975,n-1,n-1)
qf(0.025,n-1,n-1)

> var.test(caliper1, caliper2)
        F test to compare two variances
data:  caliper1 and caliper2
F = 0.47794, num df = 11, denom df = 11, p-value = 0.2364
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1375885 1.6602238
sample estimates:
ratio of variances
      0.4779412

> n=length((caliper2))
> qf(0.975,n-1,n-1)
[1] 3.473699
```

```
> qf(0.025,n-1,n-1)
[1] 0.2878776
```

Como la estadística F cae dentro de la región de no rechazo, entonces no rechazamos la hipótesis nula y decimos que las varianzas SON IGUALES.

** Probamos si existe diferencia significativa entre las medias
Ho: mu1=mu2
H1: mu1 ≃mu2

```
t.test(caliper1, caliper2, var.equal=TRUE)
qt(0.975, n-1)

> t.test(caliper1, caliper2, var.equal=TRUE)
        Two Sample t-test
data:  caliper1 and caliper2
t = 0.40519, df = 22, p-value = 0.6893
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.001029568  0.001529568
sample estimates:
mean of x mean of y
  0.26625   0.26600
> qt(0.975, n-1)
[1] 2.200985
```

La estadística t=0,4051 cae dentro de la región de No rechazo, por tanto no rechazamos Ho y decimos que con un 95% de confidencia que las medias son iguales

b) Encuentre el P-value para el test anterior

p-value = 0.6893

Como el p-value no es menor que 0.05, no se rechaza la hipótesis nula y decimos que hay suficiente evidencia estadística para pensar que las medias SI son iguales.

#2.33 Un articulo de Neurología observó que los gemelos monocigóticos comparten numerosos rasgos físicos, fisiológicos y patológicos. Los investigadores midieron el puntaje de inteligencia de 10 parejas de gemelos. Los datos obtenidos son los siguientes.
a) La suposición de normalidad en la diferencia de puntajes es valida?

c) Encuentre un set de hipotiposis apropiados indicando que la media de puntajes no depende del orden de nacimiento.
Ho: mu1=mu2
H1: mu1 ≃mu2

*Primero probamos la suposición de las Varianzas:

```
b1<-c(6.08, 6.22, 7.99, 7.44, 6.48, 7.99, 6.32, 7.60, 6.03, 7.52)
b2<-c(5.73, 5.80, 8.42, 6.84, 6.43, 8.76, 6.32, 7.62, 6.59, 7.67)
```

```
var.test(b1, b2)
n=length((b1))
qf(0.975,n-1,n-1)
qf(0.025,n-1,n-1)
```

Ho: Cociente de varianzas igual a 1
H1: Cociente de varianzas diferente a 1

```
> var.test(b1, b2)
        F test to compare two variances
data:  b1 and b2
F = 0.59148, num df = 9, denom df = 9, p-value = 0.4461
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1469164 2.3813137
sample estimates:
ratio of variances
         0.5914846
> n=length((b1))
> qf(0.975,n-1,n-1)
[1] 4.025994
> qf(0.025,n-1,n-1)
[1] 0.2483859
```

Como la estadística F=0,591 esta dentro de la región de No rechazo dado por los dos cuantiles, No podemos rechazar la hipótesis nula, y por tanto decimos que las varianzas son IGUALES.

* Prueba de Hipotesis para diferencia de medias:
Ho: mu1=mu2
H1: mu1 ≃mu2

```
t.test(b1, b2, var.equal=TRUE)
qt(0.975, n-1)
```

```
> t.test(b1, b2, var.equal=TRUE)
        Two Sample t-test
data:  b1 and b2
t = -0.12141, df = 18, p-value = 0.9047
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.933492  0.831492
sample estimates:
mean of x mean of y
    6.967     7.018
> qt(0.975, n-1)
[1] 2.262157
```

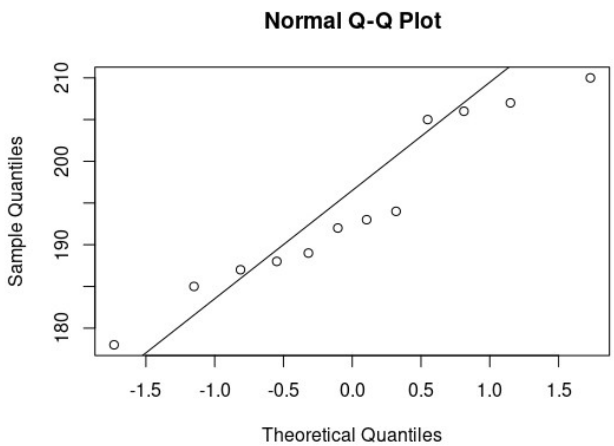
Ya que la estadística t=-0,124, cae dentro de los puntos críticos, o en la región de no rechazo, entonces no hay suficiente evidencia para rechazar la hipótesis nula. Por tanto las medias son iguales.
Esto significa que el orden de nacimiento no importa, ya que al tener la misma media, se demuestra que ambos gemelos tienen estadísticamente el mismo puntaje de inteligencia, con un 95% de confidencia.

#2.35 Las deflexiones de temperatura para unos tubos de plástico ABS se esta estudiando. 2 muestras de 12 observaciones son preparadas utilizando 2 formulas diferentes y las deflexiones de temperatura son reportadas a continuación:

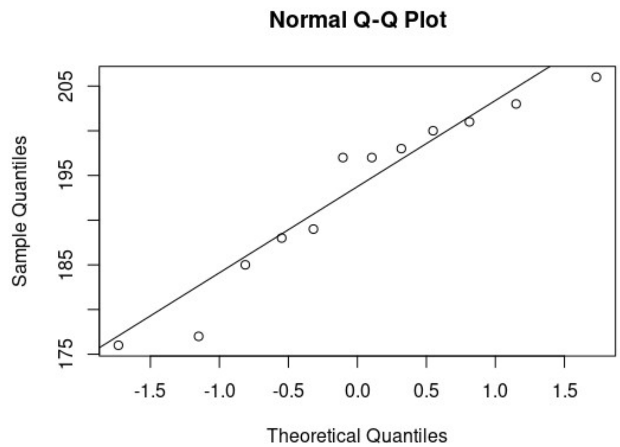
```
Formulation_1=c(206, 188, 205, 187, 193, 207, 185, 189, 192, 210, 194, 178)
Formulation_2=c(177, 197, 206, 201, 176, 185, 200, 197, 198, 188, 189, 203)
qqnorm(Formulation_1)
qqline(Formulation_1)
qqnorm(Formulation_2)
qqline(Formulation_2)
var.test(Formulation_1, Formulation_2)
n=length((Formulation_1))
qf(0.975,n-1,n-1)
qf(0.025,n-1,n-1)
```

a) Construya graficas de normalidad para ambas muestras. Estas graficas soportan la suposicion de nomalidad y varianzas iguales?

Formulation 1:



Formulation 2:



Se puede ver que para los pocos datos que se tienen, la presunción de normalidad no es completamente correcta, sin embargo los datos si tienden a tener una relación lineal.

```
> var.test(Formulation_1, Formulation_2)
```

Ho: Cociente de varianzas igual a 1

H1: Cociente de varianzas diferente a 1

```
        F test to compare two variances
data:  Formulation_1 and Formulation_2
F = 1.046, num df = 11, denom df = 11, p-value = 0.9419
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3011181 3.6334674
sample estimates:
ratio of variances
        1.045994

> n=length((Formulation_1))
```

```
> qf(0.975,n-1,n-1)
[1] 3.473699
> qf(0.025,n-1,n-1)
[1] 0.2878776
```

Como la estadistica F=1,046 cae dentro de la región de no rechazo, no rechazamos la hipotesis nula y decimos que las varianzas SON IGUALES

b)Los datos suportan la afirmación de que la media de deflexión de temperatura para la formula 1 excede a la formula 2?. Utilice alfa=0.05

Ho: mu1<=mu2
H1: mu1 >mu2

```
t.test(Formulation_1,Formulation_2, var.equal=TRUE, c("greater"))
qt(0.975, n-1)
> t.test(Formulation_1,Formulation_2, var.equal=TRUE, c("greater"))

	Two Sample t-test
data:  Formulation_1 and Formulation_2
t = 0.34483, df = 22, p-value = 0.3667
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -5.637883       Inf
sample estimates:
mean of x mean of y
 194.5000  193.0833
> qt(0.975, n-1)
[1] 2.200985
```

Podemos ver la estadística t=0,3448 se encuentra dentro de la región de no rechazo, por lo tanto no hay evidencia suficiente para afirmar que la formula 1 exceda en la media a la formula 2.

c) Cual es el P-value?
P-value = 0.3667 . Lo cual nos dice igualmente que no podemos rechazar la hipótesis nula, ya que no es menor a 0.05