UNIVERSIDAD SIMÓN BOLÍVAR
DEPARTAMENTO DE CÓMPUTO CIENTÍFICO Y ESTADÍSTICA
CÁTEDRA: ESTADISTICA PARA INGENIEROS (CO3321)

Laboratorio n° 4: Pruebas de Hipótesis.

1.- Para la media con varianza poblacional conocida:

Un grupo de vendedores de artefactos electrónicos está interesado en conocer el tiempo de duración promedio de un tipo de pilas. Para eso se seleccionaron 10 pilas y se calculó su media muestral, la cual dio como resultado 27 días. Se sabe que la población es normal y su varianza poblacional es conocida e igual a 20 días^2. Se quiere probar si el tiempo promedio de duración de las pilas es menor de 30 días, a un nivel de significación del 5%.

<u>Solución:</u>

Tenemos una prueba de hipótesis unilateral izquierda, donde
Ho:mu>=30 contra Ha: mu<30

```
>n<-10
>xbarra<-27
>sigma.cuadrado<-20
>sigma<-sqrt(sigma.cuadrado)
>alpha<-0.05
>mu<-30
>z.obs<-sqrt(n)*(xbarra-mu)/sigma
>z.alpha<-qnorm(1-alpha,0,1)
>z.alpha
```

Nos da como resultado 1.644854

Por lo tanto la región de rechazo (unilateral izquierda) es RR=(-infinito, -1.64)

Como z.obs pertenece a RR entonces los datos presentan suficientes evidencias para rechazar Ho al nivel de significación 0,05. Se concluye entonces que el tiempo de duración de las pilas es menos a 30 días.

Por otra parte, el p-valor se calcula:

```
>p.valor<-pnorm(z.obs,0,1)
>p.valor
```

Nos da como resultado 0.0169.

Como p.valor <0.05 se rechaza Ho al nivel de significación 0,05.

Nota:

- **Cuando la prueba es bilateral:**

  La región de rechazo es RR=(-infinito, -z.alphamedio) U (z.alphamedio, infinito), donde z.alphamedio se calcula como sigue:

  ```
  >z.alphamedio<-qnorm(1-alpha/2,0,1)
  ```

  El p-valor es 2-2P(Z<=|z.obs|) y se calcula como:

  ```
  >p.valor<-2-2*pnorm(abs(z.obs),0,1)
  ```

- **Cuando la prueba es unilateral derecha:**

  La región de rechazo es RR= (z.alpha, infinito), donde z.alpha se calcula como sigue:

  ```
  >z.alpha<-qnorm(1-alpha,0,1)
  ```

  El p-valor es 1-P(Z<=z.obs) y se calcula como:

  ```
  >p.valor<-1-pnorm(z.obs,0,1)
  ```

- **Cuando la prueba es unilateral izquierda:**

  La región de rechazo es RR= (- infinito, z.alpha), donde z.alpha se calcula como sigue:

  ```
  >z.alpha<-qnorm(1-alpha,0,1)
  ```

  El p-valor es P(Z<=z.obs) y se calcula como:

  ```
  >p.valor<-pnorm(z.obs,0,1)
  ```

2.- Para la media con varianza poblacional desconocida y n<30:

Los datos siguientes datos corresponden al ingreso anual (en miles de dólares) de 20 empleados de una empresa:

40 32 54 12 18 84 31 41 52 80 14 16 17 18 15 19 20 19 14 18

Se quiere ver si el ingreso anual promedio de los empleados de dicha empresa es mayor de 40.000 dólares, suponiendo que la varianza poblacional es desconocida. Use un nivel de significación del 10%.

Solución:

Ho: mu<=40 contra Ha>40

```
>datos<-c(40,32,54,12,18,84,31,41,52,80,14,16,17,18,15,19,20,19,14,18)
>t.test(datos,alternative=”greater”,mu=40,conf.level=0.90)
```

Esto genera:

tobs=-1.9225
p-valor=0.9652

Como p-valor>alpha, entonces los datos presentan suficientes evidencias para no rechazar Ho al nivel de significación del 10%. Se concluye entonces que el ingreso promedio de los empleados de la empresa es mayor de 40.000 dólares.

Otra manera de hacerlo:

```
>xbarra<-mean(datos)
>s<-sd(datos)
>alpha<-0.1
>n<-20
>mu<-40
>t.obs<-(xbarra-mu)*sqrt(n)/s
>t.obs
>t.alpha<-qt(1-alpha,n-1)
>t.alpha
```

Nos da como resultado

t.obs= -1.9225
t.alpha=1.327728

Entonces, la región de rechazo es RR=(1.327728, infinito)

Como t.obs no pertenece a RR, entonces los datos presentan suficientes evidencias para no rechazar Ho al nivel de significación del 10%.

Nota:

- Si la prueba es unilateral derecha (cola superior) se coloca dentro de t.test: alternative=”greater”.

- Si la prueba es unilateral izquierda (cola inferior) se coloca dentro de t.test alternative=”less”

- Si la prueba es bilateral (dos colas) se coloca dentro de t.test: alternative=”two.sided”

3.-Para la varianza:

Los datos siguientes datos corresponden al ingreso anual (en miles de dólares) de 20 empleados de una empresa:

40 32 54 12 18 84 31 41 52 80 14 16 17 18 15 19 20 19 14 18

Se quiere contrastar la hipótesis de que la varianza es de 10 contra la varianza es distinta de 10. Use un nivel de significación del 5%.

<u>Solución:</u>

Ho: sigma^2=10 contra Ha: sigma^2 distinta de 10

```
>datos<-c(40,32,54,12,18,84,31,41,52,80,14,16,17,18,15,19,20,19,14,18)
>n<-length(datos)
>alpha<-0.05
>sigma.cuadrado<-10
>s.cuadrado<-var(datos)
>a<-qchisq(alpha/2,n-1)
>a
>b<-qchisq(1-alpha/2,n-1)
>b
>chi.obs<-(n-1)*s.cuadrado/sigma.cuadrado
>chi.obs
```

Esto nos da como resultado:

a=8.906516
b=32.85233
chi.obs=889.22

Por lo tanto, la región de rechazo es RR=(0, 8.906516)U(32.85233, infinito)
Como chi.obs pertenece a RR, los datos presentan suficientes evidencias para rechazar Ho al nivel de significación del 5%.

4.- Para proporciones:

Se considera que el 20% de las computadoras fabricadas por una determinada empresa presentan fallas. Sin embargo, en una muestra de 100 computadoras fabricadas por dicha empresa, se encontraron a 25 defectuosas. Para un nivel de significación del 1%. ¿Aceptaríamos la hipótesis de que la proporción de computadoras defectuosas fabricadas por dicha empresa es superior al 20%?

<u>Solución:</u>

Tenemos Ho: p<=0.2 contra Ha: p>0.2

```
>n<-100
>x<-25
>p0<-0.2
>ptecho<-x/n
>alpha<-0.01
>z.obs<-(ptecho-p0)/sqrt(p0*(1-p0)/n)
>z.obs
>z.alpha<-qnorm(1-alpha,0,1)
>z.alpha
```

Esto nos da como resultado:

z.obs=1.25
z.alpha=2.326348

Entonces la región de rechazo es RR = (2.32638, infinito)

Como z.obs no pertenece a la región de rechazo, los datos presentan suficientes evidencias para no rechazar Ho al nivel de significación del 1%. Por lo tanto, la proporción de computadoras defectuosas es igual o menor al 20%.

El p-valor es

```
>p.valor<-1-pnorm(z.obs,0,1)
>p.valor
```

Este da como resultado 0.1056498 y se concluye lo mismo que en la parte anterior.

6.-Para el cociente de varianzas:

Los siguientes datos corresponden a los ingresos mensuales (en miles de dólares) de 10 empleados de la empresa A y de 12 empleados de la empresa B, respectivamente.

Empresa A: 12   15   19   3   8   10   19   21   10   12
Empresa B: 15   12   10   8   8   7   19   4   19   17   26   18

Se quiere probar si las varianzas de los ingresos mensuales de los empleados en ambas empresas son iguales a un nivel de significación del 1%.

<u>Solución:</u>

Tenemos Ho: $\sigma 1=\sigma 2$ contra Ha: $\sigma 1 \neq \sigma 2$

```
>data1<-c(12,15,193,8,10,19,21,10,12)
>data2<-c(15,12,10,8,8,7,19,4,19,17,26,18)
>var.test(data1,data2,alternative=”two.sided”,conf.level=0.99)
```

Esto nos da los resultados:

f.obs=0.7588
p-valor=0.6895

Como p-valor>alpha, entonces los datos presentan suficientes evidencias para no rechazar Ho al nivel de significación del 1%. Se supone entonces que las varianzas son iguales.

7.- Para la diferencia de medias:

Del problema anterior queremos probar ahora si ingresos mensuales promedios de los empleados de ambas empresas son iguales al mismo nivel de significación.

<u>Solución:</u>

Tenemos Ho: mu1=mu2 contra Ha: mu1≠ mu2

Con los mismos comandos del ejemplo anterior, le agregamos:

```
>t.test(data1,data2,alternative=”two.sided”,var.equal=TRUE,conf.level=0.99)
```

Se tiene entonces:

tobs=-0.2617
p-valor=0.7962

Como p-valor>alpha, entonces los datos presentan suficientes evidencias para no rechazar Ho al nivel de significación del 1%, por lo tanto se puede suponer que las medias son iguales, es decir, el ingreso mensual promedio de los empleados en ambas empresas son iguales.