Proyecto 5

Rudik Rompich - Carlos Martínez

Problema 3

Una comparación de la durabilidad de dos tipos de llantas para automóvil se obtuvo de muestras de pruebas en carretera de $n_= n_2 = 100$ llantas de cada tipo. Se registró el número de millas hasta quedar inútiles, el desgaste se definió como el número de millas hasta que la cantidad restante de superficie de rodamiento llegó a un valor pequeño especificado previamente. Las mediciones para los dos tipos de llantas se obtuvieron de manera independiente y se calcularon las siguientes medias y varianzas:

$$\overline{y_1} = 26,400 \text{ millas} \qquad \overline{y_2} = 25,100 \text{ millas}$$
$$s_1^2 = 1,440,000 \qquad s_2 = 1,960,000.$$

Estime la diferencia en la media de millas hasta quedar inútiles y precise un límite de error estándar de 2 en el error de estimación.

Solución

Comenzamos colocando los datos de las llantas.

```
#Tipo 1
n1= 100
m1= 26400
s1= 1440000

#Tipo 2
n2= 100
m2= 25100
s2= 1960000
```

Ahora bien, considerando que existe una cantidad de datos significativa (n $\geq$ 30); es posible hacer una aproximación del error estándar del estimador (diferencia de medias), se calcula por medio de:

$$\sigma_{\overline{Y_1}-\overline{Y_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

```
#Función basada en el error estándar de la tabla 8.1
error_estandar <- function(n1,s1,n2,s2){
  error <- sqrt((s1)/n1 + (s2)/n2)
  return(error)
}

#Resultado
errorcito <- error_estandar(n1,s1,n2,s2)
```

Por lo tanto, la diferencia de desgaste medio es:

```
desgaste= m1-m2
sprintf("%d millas", desgaste)
```

```
## [1] "1300 millas"
```

Por otra parte, el error de estimación de 2:

```
sprintf("%f millas", 2*errorcito)
```

```
## [1] "368.781778 millas"
```

con una probabilidad de aproximadamente 0.95.

Problema 10

Un experimentador desea comparar la efectividad de dos métodos de capacitación para obreros que van a realizar una operación de ensamble. Los obreros seleccionados han de dividirse en dos grupos de igual tamaño, el primero para recibir el método 1 de capacitación y el segundo el método 2 de capacitación. Después de la capacitación cada obrero realizará la operación de ensamble y se registrará el tiempo que le tome hacerlo. El experimentador espera que las mediciones para ambos grupos tengan una amplitud de aproximadamente 8 minutos. Si la estimación de la diferencia en los tiempos promedio de ensamble debe ser correcta con una variación de no más de 1 minuto con probabilidad 0.95, ¿cuántos trabajadores deben incluirse en cada grupo de capacitación?

Solución

Ya que la probabilidad es 0.95, entonces $1 - \alpha = 0.95$, lo quiere decir que $\alpha = 0.05$. Ahora

```
zeta <-qnorm(0.025, mean = 0, sd=1, lower.tail = FALSE, log= FALSE)
```

$z_{\alpha/2} = z_{0.25}$ es igual a

```
zeta
```

```
## [1] 1.959964
```

Ahora bien como se espera que la diferencia en los tiempos promedio no varíe más de 1 minuto, entonces

$$1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 1$$

Como el tamaño de los dos grupos debe tener igual tamaño, quiere decir que $n_1 = n_2 = n$. Tenemos que:

$$1.96\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} = 1$$

Además, en ambos métodos de esamble, no se desea más de 1 minuto de variación y que los grupos tengan una amplitud de 8 minutos, entonces sabemos que $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Tomando en cuenta que los datos, con 0.95 de probabilidad, caen a en un rango de $4\sigma = 8$ minutos $\Rightarrow 1.96\sqrt{\frac{2^2}{n} + \frac{2^2}{n}} = 1 \Rightarrow n = 30.73$

Concluimos que bastan aproximadamente 31 obreros en cada uno de los dos grupos.

Problema 12

Para alcanzar la máxima eficiencia al realizar una operación de ensamble en una planta manufacturera, obreros nuevos requieren aproximadamente un periodo de capacitación de 1 mes.Se sugirió un nuevo método de capacitación y se realizó un examen para comparar el nuevo método contra el procedimiento estándar. Dos grupos de nueve obreros nuevos cada uno fueron capacitados durante 3 semanas, un grupo usando el nuevo método y el otro siguiendo el procedimiento estándar de capacitación. El tiempo (en minutos) requerido por cada obrero para ensamblar el dispositivo se registró al final del período de 3 semanas. Las mediciones resultantes son las que se muestran en la Tabla 8.3. Calcule la diferencia real de las medias $(m_1 - m_2)$ con coeficiente de confianza 0.95. Suponga que los tiempos de ensamble están distribuidos normalmente en forma aproximada, que las varianzas de los tiempos de ensamble son aproximadamente iguales para los dos métodos y que las muestras son independientes.

Procedimiento	Mediciones
Estándar	32, 37, 35, 28, 41, 44, 35, 31, 34
Nuevo	35, 31, 29, 25, 34, 40, 27, 32, 31

Solución

Sabemos que los tiempos de ensamblaje están distribuidos uniformemente, las muestras son independientes y las varianzas son aproximadamente iguales. Se procede a calcular las medias y varianzas:

```
#DB método estándar
estandar<- c(32, 37, 35, 28, 41, 44, 35, 31, 34)
#media
m1 <-mean(estandar)
#varianza
s1 <-sd(estandar)^2
#n
n1 <- length(estandar)

#DB método nuevo
nuevo <- c(35, 31, 29, 25, 34, 40, 27, 32, 31)
#media
m2 <-mean(nuevo)
#varianza
s2 <-sd(nuevo)^2
#n
n2 <- length(nuevo)
```

Sabemos que el intervalo de confianza para $(\mu_1 - \mu_2)$ se calcula por medio de:

$$(\overline{y}_1 - \overline{y}_2) \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

donde $t_{\alpha/2}$ es determinado de la distribución t con $(n_1 + n_2 - 2)$ gl. Además, s_p se calcula con:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Por lo tanto, procedemos a calcular el intervalo de confianza declarando las siguientes funciones:

```
#Función para calcular sp
sp_cuadrado <- function(n1,n2,s1,s2){
  sp_square <- ((n1-1)*s1+(n2-1)*s2)/(n1+n2-2)
  return(sp_square)
}

# Función para calcular t_{\alpha/2}

t_alpha <- function(gl,alpha){
  significancia = alpha/2
```

```r
  t<- qt(p=significancia, df = gl,lower.tail=FALSE)
  return(t)
}

#Función para calcular intervalo de confianza (segunda parte)

intervalo_confianza <- function(alpha,n1,n2,s1,s2){
  gl <- n1+n2-2
  t <- t_alpha(gl,alpha)
  sp <- sqrt(sp_cuadrado(n1,n2,s1,s2))
  resultado<- t*sp*sqrt(1/n1+1/n2)
  return(resultado)
}
```

Como resultado obtenemos el intervalo de confianza:

```r
medias <- m1-m2
dev<-intervalo_confianza(0.05,n1,n2,s1,s2)
sprintf("Resultado: %f +/- %f ",medias,dev)
```

```
## [1] "Resultado: 3.666667 +/- 4.712373 "
```

```r
sprintf("Resultado equivalente: [%f,%f] ",medias-dev,medias+dev)
```

```
## [1] "Resultado equivalente: [-1.045706,8.379039] "
```

Citando al libro de texto: «El intervalo es bastante ancho e incluye valores positivos y negativos. Si $m_1 - m_2$ es positivo, $m_1 > m_2$ y el procedimiento estándar tiene un tiempo de ensamble esperado mayor que el nuevo procedimiento. Si $m_1 - m_2$ es realmente negativo, lo inverso es verdadero. Como el intervalo contiene valores positivos y negativos, se puede decir que ninguno de los métodos de capacitación produce un tiempo medio de ensamble que difiera del otro.»

Problema 13

Un experimentador desea comprobar la variabilidad de mediciones obtenidas al usar equipo diseñado para medir el volumen de una fuente de audio. Tres mediciones independientes registradas por este equipo para la misma fuente de sonido fueron 4.1, 5.2 y 10.2. Estime s2 con coeficiente de confi anza .90

Solución

Suponiendo que los datos tienen una distribución normal, se puede aplicar un intervalo de confianza $100(1-\alpha)\%$ para σ^2

$$\left(\frac{(n-1)S^2}{\mathcal{X}^2_{\alpha/2}}, \frac{(n-1)S^2}{\mathcal{X}^2_{1-(\alpha/2)}}\right)$$

```r
var <- (sd(c(4.1, 5.2, 10.2)))^2
```

Entonces S^2 es

```r
var
```

```
## [1] 10.57
```

En este caso $\alpha = 1 - 0.90 = 0.10 \Rightarrow \alpha/2 = 0.05$ y $1 - (\alpha/2) = 0.95$ y $(n-1) = 3 - 1 = 2$ grados de libertad.

```r
gl1 <- 2
probabilidad1 <- 0.95
```

```
x1 <- qchisq(p=probabilidad1, df=gl1, ncp=0, lower.tail=FALSE, log.p=FALSE)
```

Entonces $\mathcal{X}^2_{1-(\alpha/2)} = \mathcal{X}^2_{0.95}$ es

```
x1
```

```
## [1] 0.1025866
```

Por otra parte

```
probabilidad2 = 0.05
x2 <- qchisq(p=probabilidad2, df=gl1, ncp=0, lower.tail=FALSE, log.p=FALSE)
```

Entonces $\mathcal{X}^2_{\alpha/2} = \mathcal{X}^2_{0.05}$ es

```
x2
```

```
## [1] 5.991465
```