



Universidad Simón Bolívar

Departamento de Cómputo Científico y Estadística

CO3321 Estadística para Ingenieros

Laboratorio 4: Pruebas de hipótesis

Hecho por:

Jorge M., María Victoria

11-10495

Sartenejas, agosto de 2015

Tomaremos X como la variable asociada al número de hijos nacidos vivos de una mujer. Las muestras tomadas que se utilizarán para las pruebas de hipótesis son las de los estados Miranda y Zulia, ubicadas en los archivos 'mirmus.txt' y 'zulmus.txt', respectivamente.

En cada estado se tiene una muestra de tamaño 250, por lo tanto se realizarán las pruebas de hipótesis correspondientes a muestras grandes. Además, como es una muestra grande podemos decir que la varianza poblacional (desconocida en el enunciado) es igual a la varianza muestral, que es posible calcularla con los datos provistos en cada archivo.

Prueba de hipótesis para la diferencia de medias

Para determinar si las medias poblacionales del número de hijos son iguales o no entre los estados Miranda y Zulia, se realizó una prueba de hipótesis, usando una significancia del 3%. Las hipótesis planteadas fueron las siguientes: $H_0: \mu_Z - \mu_M = 0$, $H_a: \mu_M \neq \mu_Z$. Siendo μ_M y μ_Z , las medias poblacionales de los estados Miranda y Zulia, respectivamente. La región de rechazo será $Z > Z_{\alpha/2} = Z_{0.015} = 2.17009$, es decir, $RR: \{|Z| > 2.17009\}$. El valor obtenido para el estadístico Z fue 1.684346. Como el estadístico Z no se encuentra dentro de la región de rechazo, aceptamos la hipótesis nula. Esto quiere decir que con un nivel de significancia del 3% no se puede concluir que las medias poblacionales en Miranda y Zulia son distintas.

Por otro lado, para determinar el p-valor debemos calcular $P(Z > 1.684346) + P(Z < -1.684346)$, que es lo mismo que calcular $2 * P(Z > 1.684346)$. Por lo tanto, el p-valor obtenido fue 0.09211487. Finalmente, como el p-valor es mayor que el alfa utilizado (0.03) no se puede rechazar la hipótesis nula, llegando a la misma conclusión que en el estudio realizado antes.

Pruebas de hipótesis para las medias

Para determinar, por medio de una prueba de hipótesis, si el número medio de hijos para una mujer en los estados Miranda y Zulia es mayor a dos, usando una significancia de 7%, se estudiaron los dos estados por separado con las siguientes hipótesis: $H_0: \mu = 2$, $H_a: \mu > 2$. Para determinar la región de rechazo, se calcularon los estadísticos Z . Luego, para determinar las regiones de rechazo se utilizó la región $Z > Z_{\alpha}$. La región de rechazo para ambos casos es $\{Z > 1.475791\}$. Los resultados obtenidos para cada estado fueron los siguientes:

- Estado Miranda: $Z = 8.622832$
- Estado Zulia: $Z = 11.03225$

Los estadísticos de ambos estados se encuentran dentro de la región de rechazo, por lo que para las dos pruebas de hipótesis se rechaza la hipótesis nula. Por lo tanto, existe suficiente evidencia con un 7% de significancia para decir que el número medio de hijos para una mujer, tanto en Miranda como en Zulia, es mayor a dos.

Luego, para calcular el p-valor de la prueba, buscamos la probabilidad de que el estadístico Z sea mayor que 8.622832 en el caso del estado Miranda, y mayor que 11.03225 en el estado Zulia. Los valores obtenidos para cada estado fueron los siguientes:

- Estado Miranda: p-valor = 3.265908×10^{-18}
- Estado Zulia: p-valor = 1.33553×10^{-28}

Con los resultados de ambos p-valor podemos decir que la hipótesis nula se rechaza para un nivel de significancia del 7%. Además, como los p-valor de los dos estados dieron un número que es casi cero, se puede decir que la hipótesis nula siempre se rechaza. Por lo tanto, analizando el p-valor de cada prueba concluimos lo mismo que con la región de rechazo, que existe suficiente evidencia para afirmar que el número medio de hijos para una mujer, tanto en Miranda como en Zulia, es mayor a 2.

A continuación se presenta el script de R utilizado para hacer todos los cálculos mencionados en el presente informe:

```
# Laboratorio 4 de Estadística
# Hecho por: María Victoria Jorge 11-10495
# Entidades utilizadas: Miranda y Zulia

miranda = scan('mirmus.txt') # Guardamos en un vector las muestras del Edo. Miranda
zulia = scan('zulmus.txt') # Guardamos en un vector las muestras del Edo. Zulia

nM = length(miranda)
nZ = length(zulia)

# Medias y desviación estándar muestrales para ambos estados
muM = mean(miranda)
sM = sd(miranda)

muZ = mean(zulia)
sZ = sd(zulia)

# Determinar si las medias poblacionales del número de hijos son iguales o no entre los estados
# Miranda y Zulia,
# con significancia de 3%.
# Usando una prueba de hipótesis
# Ho:  $\mu_Z - \mu_M = 0$ , Ha:  $\mu_M \neq \mu_Z$ 

alfa = 0.03
Zalfa = qnorm(alfa/2, lower.tail = F)

Z = (muZ - muM)/sqrt((var(miranda)/nM) + (var(zulia)/nZ))

# Calculando el p-valor

p_valor = 2*pnorm(Z, lower.tail = F)

# Determinar si el número medio de hijos para una mujer en Miranda y Zulia es mayor a dos, con
# significancia de 7%.
# Usando una prueba de hipótesis
# Ho:  $\mu = 2$ , Ha:  $\mu > 2$ 

alfa2 = 0.07 # Nivel de significancia
Z2alfa = qnorm(alfa2, lower.tail = F)
```

```
Z2Miranda = (muM - 2)/(sM/sqrt(nM))  
Z2Zulia = (muZ - 2)/(sZ/sqrt(nZ))
```

```
# Calculando el p-valor
```

```
p_valor2Miranda = pnorm(Z2Miranda,0,1,lower.tail = F)  
p_valor2Zulia = pnorm(Z2Zulia,0,1,lower.tail = F)
```