Prueba de hipotesis

Yoel Domínguez

2023-12-18

Prueba de hipótesis

Ejemplo

Supongamos que un fabricante de bombillas afirma que la vida útil promedio de una bombialla es de más de 10,000 horas. En una muestra de 30 bombillas, se descubrió que solo duran 9,900 horas en promedio. Suponga que la desviación estándar de la población es de 120 horas. Con un nivel de significancia de alpha=0.05, ¿podemos rechazar la afirmación del fabricante?

Solución

La hipótesis nula es que µ mayor o igual que 10,000.

Comenzamos con el cálculo de estadística de la prueba.

Variables

xbar = 9900 -> Media de la muestra

mu0 = 10000 -> Valor hipotético

sigma = 120 -> Desviación estándar de población

n = 30 -> Tamaño de la muestra

xbar <- 9900
mu0 <- 10000 sigma <- 120 n <- 30 z <- (xbar-mu0)/(sigma/sqrt(n)) z

```r
xbar <- 9900
mu0 <- 10000
sigma <- 120
```

```r
n <- 30
z <- (xbar-mu0)/(sigma/sqrt(n))
z
```

```
## [1] -4.564355
```

Calculamos el valor crítico a un nivel de significación de

alpha = 0.05

alpha = 0.05

z.alpha = qnorm(1-alpha) -> Valor crítico

-z.alpha -> Resultado

alpha = 0.05 z.alpha = qnorm(1-alpha) -z.alpha

```r
alpha = 0.05
z.alpha = qnorm(1-alpha)
-z.alpha
```

```
## [1] -1.644854
```

Respuesta

La estadística de prueba -4.5644 es menor que el valor crítico de -1.6449. Por lo tanto, a un nivel de significación de *alpha=0.05* Rechazamos la afirmación de que la vida media de una bombilla es superior a 10,000 horas.

mean <- 0; sd <- 1 lb <- -4; ub <- -1.644854

x <- seq(-4,4,length=1000)*sd + mean hx <- dnorm(x,mean,sd)

plot(x, hx, type="n", xlab="IQ Values", ylab="", main="Normal Distribution", axes=FALSE)

i <- x >= lb & x <= ub lines(x, hx) polygon(c(lb,x[i],ub), c(0,hx[i],0), col="red")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd) result <- paste("P(",lb,"< IQ <",ub,") =", signif(area, digits=4)) mtext(result,3) axis(1, at=seq(-4, 4, 1), pos=0)

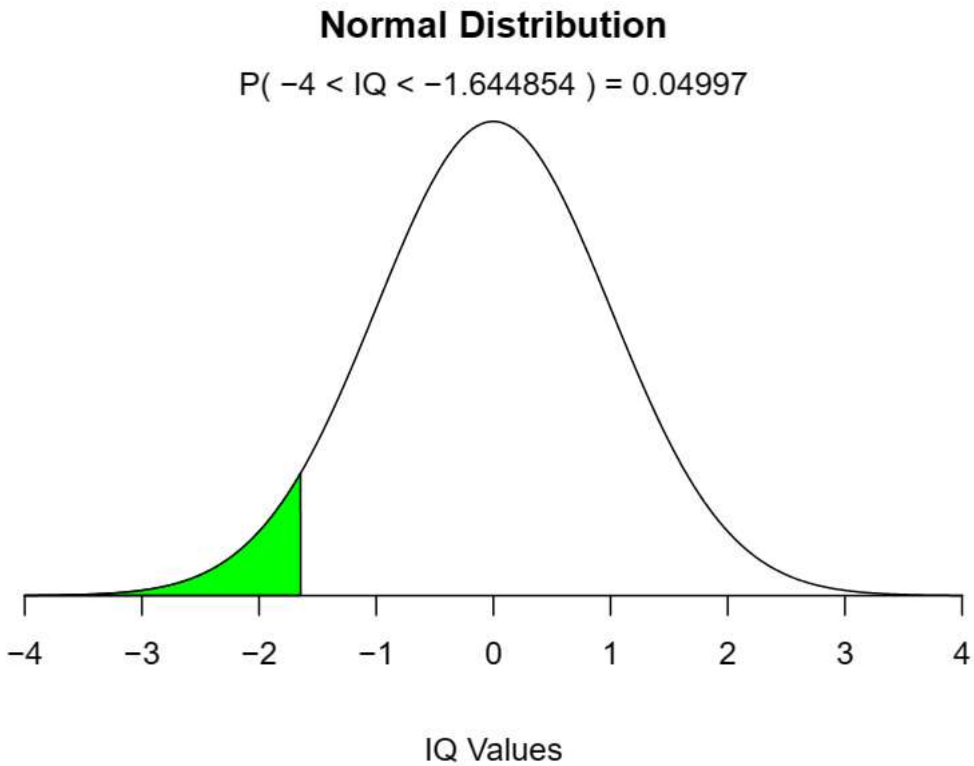
```r
mean <- 0;    sd   <- 1
lb   <- -4;   ub   <- -1.644854

x <- seq(-4,4,length=1000)*sd + mean
hx <- dnorm(x,mean,sd)

plot(x, hx, type="n", xlab="IQ Values", ylab="",
     main="Normal Distribution", axes=FALSE)
```

```
i <- x >= lb & x <= ub
lines(x, hx)
polygon(c(lb,x[i],ub), c(0,hx[i],0), col="green")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd)
result <- paste("P(",lb,"< IQ <",ub,") =",
                signif(area, digits=4))
mtext(result,3)
axis(1, at=seq(-4, 4, 1), pos=0)
```



Solución alternativa

En lugar de utilizar el valor crítico, aplicamos la función pnorm para calcular el p-valor de la cola inferior de la prueba de estadística. Como resulta ser menor que el nivel de significación alpha=0.05, rechazamos la hipótesis nula de que mu es mayor o igual a 10,000

p-valor mayor o igual que alpha

2.505166e-06 menor que 0.05

pval=pnorm(z) pval

```
pval=pnorm(z)
pval
```

```
## [1] 2.505166e-06
```

Se grafica con un código similar al anterior, para graficar el p valor, dándonos cuenta de que ya no se muestra dentro de nuestra gráfica.

mean <- 0; sd <- 1 lb <- -100000; ub <- -4.5644

x <- seq(-4,4,length=1000)*sd + mean hx <- dnorm(x,mean,sd)

plot(x, hx, type="n", xlab="IQ Values", ylab="", main="Normal Distribution", axes=FALSE) points(-4.5644,0) i <- x >= lb & x <= ub lines(x, hx) polygon(c(lb,x[i],ub), c(0,hx[i],0), col="red")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd) result <- paste("P(",lb,"< IQ <",ub,") =", signif(area, digits=4)) mtext(result,3) axis(1, at=seq(-4, 4, 1), pos=0)

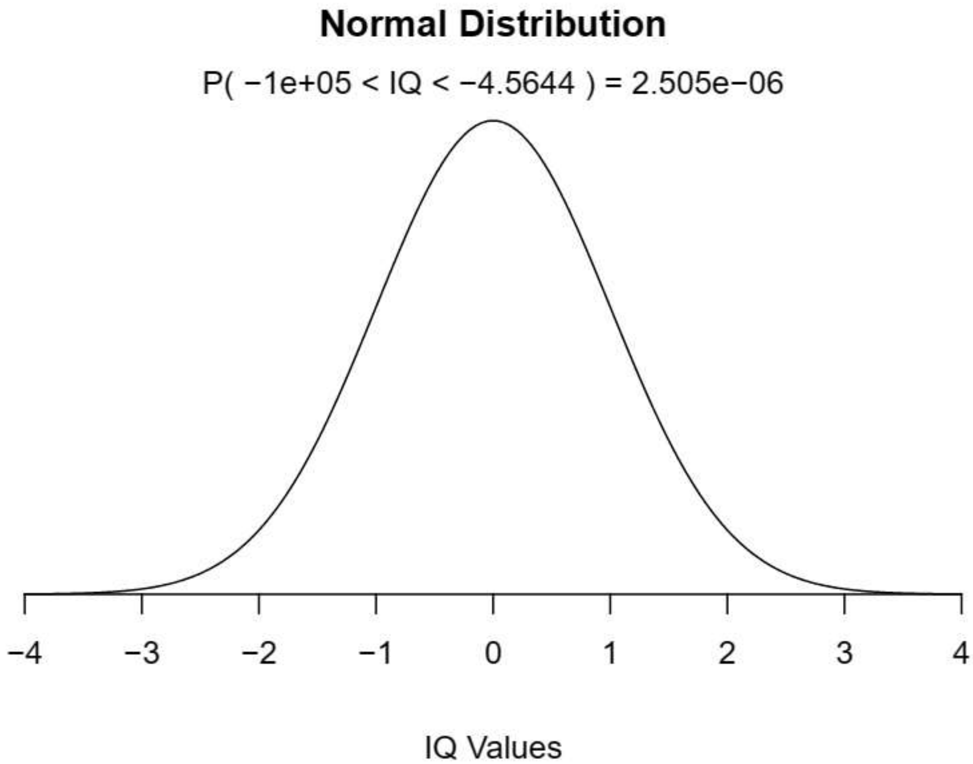
```
mean <- 0;    sd   <- 1
lb   <- -100000;   ub   <- -4.5644

x <- seq(-4,4,length=1000)*sd + mean
hx <- dnorm(x,mean,sd)

plot(x, hx, type="n", xlab="IQ Values", ylab="",
     main="Normal Distribution", axes=FALSE)
points(-4.5644,0)
i <- x >= lb & x <= ub
lines(x, hx)
polygon(c(lb,x[i],ub), c(0,hx[i],0), col="green")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd)
```

```r
result <- paste("P(",lb,"< IQ <",ub,") =",
                signif(area, digits=4))
mtext(result,3)
axis(1, at=seq(-4, 4, 1), pos=0)
```



Problema 1

Un fabricante de cerveza artesanal afirma que el volumen de producción es mayor que 1,000 litros. En una muestra de tres producciones, se descubrió que la producción es de 991 litros en promedio. Suponga que la desviación estándar de la población es de 15. Con un nivel de significancia de alpha=0.1, ¿podemos rechazar la afirmación del fabricante?

Solución

La hipótesis nula es que µ mayor o igual que 1,000. Comenzamos con el cálculo estadístico de la prueba.

Variables

xbar = 991 -> Media de la muestra

mu0 = 1000 -> Valor hipotético

sigma = 15 -> Desviación estándar de población

n = 3 -> Tamaño de la muestra

xbar <- 991
mu0 <- 1000
sigma <- 15
n <- 3
z <- (xbar-mu0)/(sigma/sqrt(n)) z

```
xbar  <- 991
mu0   <- 1000
sigma <- 15
n     <- 3
z <- (xbar-mu0)/(sigma/sqrt(n))
z
```

```
## [1] -1.03923
```

Solución

Entonces calculamos el valor crítico a un nivel de significación de alpha=0.01

alpha = 1

z.alpha = qnorm(1-alpha) -> Valor crítico

-z.alpha -> Resultado

alpha = 0.1 z.alpha = qnorm(1-alpha) -z.alpha

```r
alpha = 0.1
z.alpha = qnorm(1-alpha)
-z.alpha
```

```
## [1] -1.281552
```

Se procede a hacer la comparación y ver si se cumple la condición o no.

Respuesta

La prueba estadística de z = -1.03923 es mayor que el valor crítico de z(alpha) = -1.281552. Por lo tanto, a un nivel de significación de alpha=0.1, aceptamos la afirmación de que la producción medua de cerveza es mayor a 1,000 litros.

Realizamos el gráfico.

mean <- 0; sd <- 1 lb <- -4; ub <- -1.281552

x <- seq(-4,4,length=1000)*sd + mean hx <- dnorm(x,mean,sd)

plot(x, hx, type="n", xlab="IQ Values", ylab="", main="Normal Distribution", axes=FALSE) points(-1.03923,0) i <- x >= lb & x <= ub lines(x, hx) polygon(c(lb,x[i],ub), c(0,hx[i],0), col="red")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd) result <- paste("P(",lb,"< IQ <",ub,") =", signif(area, digits=4)) mtext(result,3) axis(1, at=seq(-4, 4, 1), pos=0)

```r
mean <- 0;     sd   <- 1
lb   <- -4;    ub   <- -1.281552

x <- seq(-4,4,length=1000)*sd + mean
hx <- dnorm(x,mean,sd)
```

```r
plot(x, hx, type="n", xlab="IQ Values", ylab="",
     main="Normal Distribution", axes=FALSE)
points(-1.03923,0)
i <- x >= lb & x <= ub
lines(x, hx)
polygon(c(lb,x[i],ub), c(0,hx[i],0), col="green")

area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd)
result <- paste("P(",lb,"< IQ <",ub,") =",
                signif(area, digits=4))
mtext(result,3)
axis(1, at=seq(-4, 4, 1), pos=0)
```