Practica 2

Gonzalo Cruz Gómez, Daniel Fernandez Magariños

2024-11-26

Introducción

Ejercicio 1

Demostrar que la log-verosimilitud de p basada en n observaciones es:

$$x \sim \text{Bin}(n, p)$$

$$L(x) = \binom{n}{x} p^r (1-p)^{n-r}$$

$$\ell(p|r) = \log \binom{n}{r} + r\log(p) + (n-r)\log(1-p)$$

$$= r\log\left(\frac{p}{1-p}\right) + n\log(1-p)$$

Ejercicio 2

$$\bar{I}(p) = \mathbb{E}\left[\left(\frac{\partial^2}{\partial p^2}\log L(p)\right)\right]$$

$$\frac{\partial \ell}{\partial p} = \frac{r}{p} + \frac{r-n}{1-p}, \quad \frac{\partial^2}{\partial p^2} = -\frac{r}{p^2} - \frac{n-r}{(1-p)^2}.$$

$$\mathbb{E}\left[-\frac{\partial^2}{\partial p^2}\right] = \mathbb{E}\left[\frac{r}{p^2}\right] + \mathbb{E}\left[\frac{n-r}{(1-p)^2}\right] =$$

$$= \frac{n}{p} - \frac{np}{(1-p)^2} + \frac{n}{(1-p)^2} = \frac{n(1-p)^2 - np^2 + np}{p(1-p)^2} =$$

$$\frac{n(1-2p+p^2) - np^2 + np}{p(1-p)^2} = \frac{n-np}{p(1-p)^2} = \frac{n(1-p)}{p(1-p)^2} = \frac{n}{p(1-p)}$$

$$\bar{I}(p) = \frac{n}{p(1-p)}$$

Ejercicio 3

Contraste de hipótesis

Ejercicio 4

```r
#valores dummy
p0 <- 0.3
n <- 1000
r <- 400


calculate <- function(p0, n, r, parametrizacion){
  if(parametrizacion == 1){
    Z <- r - n*p0
    V <- n*p0*(1-p0)
  }
  else if (parametrizacion == 2){
    Z <- (r-n*p0)/(p0*(1-p0))
    V <- n/(p0*(1-p0))
  }
  else if (parametrizacion == 3) {
    Z <- (2 / sqrt(1 - p0)) * (r / sqrt(p0) - n * sqrt(p0))
    V <- 4*n

}

  return (list(Z=Z, V=V))
}

result1 <- calculate(p0, n, r, parametrizacion = 1)
result2 <- calculate(p0, n, r, parametrizacion = 2)
result3 <- calculate(p0, n, r, parametrizacion = 3)

cat("Resultados para parametrización 1:\n")
```

```
## Resultados para parametrización 1:
```

```r
print (result1)
```

```
## $Z
## [1] 100
##
## $V
## [1] 210
```

```r
cat("Resultados para parametrización 2:\n")
```

```
## Resultados para parametrización 2:
```

```r
print (result2)
```

```
## $Z
## [1] 476.1905
##
## $V
## [1] 4761.905
```

```r
cat("Resultados para parametrización 3:\n")
```

```
## Resultados para parametrización 3:
```

```r
print (result3)
```

```
## $Z
## [1] 436.4358
##
## $V
## [1] 4000
```

Ejercicio 5

Cálculo del tamaño muestral

Ejercicio 6

$$V_1 = n\, p_0\, (1 - p_0)$$

$$V_2 = \frac{n}{p_0(1-p_0)}$$

$$V_3 = 4n$$

$$V = \frac{(z_\alpha + z_\beta)^2}{\theta_R}$$

Primera parametrización:

$$n = \frac{(z_\alpha + z_\beta)^2}{\theta_R^2 \cdot p_0(1-p_0)} = \frac{(z_\alpha + z_\beta)^2}{\log\left(\frac{p(1-p_0)}{p_0(1-p)}\right)^2 p_0\,(1-p_0)}$$

Segunda parametrización:

$$n = \frac{(z_\alpha + z_\beta)^2 p_0(1-p_0)}{(p-p_0)^2}$$

Tercera parametrización:

$$n = \frac{(z_\alpha + z_\beta)^2}{4\theta^2} = \frac{(z_\alpha + z_\beta)^2}{4(\arcsin\sqrt{p} - \arcsin\sqrt{p}_0)^2}$$

Potencia y error de tipo I

Ejercicio 7

```r
# Parámetros enunciado
p0 <- 0.003
p <- 0.006
alpha <- 0.025
beta <- 0.20
```

```r
# Percentiles z_alpha/2 y z_beta
z_alpha <- qnorm(1 - alpha/2)
z_beta <- qnorm(1 - beta)

#   = log((p(1-p0)) / (p0(1-p)))
n1 <- (z_alpha + z_beta)^2 / (log((p * (1 - p0)) / (p0 * (1 - p)))^2 * p0 * (1 - p0))

#   = p - p0
n2 <- (z_alpha + z_beta)^2 * p0 * (1 - p0) / ((p-p0)^2)

#   = arcsin(sqrt(p)) - arcsin(sqrt(p0))
n3 <- (z_alpha + z_beta)^2 / (4*(asin(sqrt(p))-asin(sqrt(p0)))^2)

# usamos ceiling(n) para redondear al siguiente entero
cat("Tamaño muestral para cada parametrización:\n")
```

```
## Tamaño muestral para cada parametrización:
```

```r
cat("  = log((p(1-p0)) / (p0(1-p))): n =", ceiling(n1), "\n")
```

```
##   = log((p(1-p0)) / (p0(1-p))): n = 6558
```

```r
cat("  = p - p0: n =", ceiling(n2), "\n")
```

```
##   = p - p0: n = 3159
```

```r
cat("  = arcsin(sqrt(p)) - arcsin(sqrt(p0)): n =", ceiling(n3), "\n")
```

```
##   = arcsin(sqrt(p)) - arcsin(sqrt(p0)): n = 4597
```

Ejercicio 8

```r
p0 <- 0.003
p <- 0.006
n1 <- ceiling(n1) # log
n2 <- ceiling(n2) # diferencia
n3 <- ceiling(n3) # arcsin

n <- c(n1, n2, n3)

calcular_errores<- function(n, p0, p) {
  # error estándar de H0
  stderror_h0 <- sqrt(p0 * (1 - p0) / n)
  v_crit <- p0 + 1.96 * stderror_h0

  # error tipo 1
  error_1 <- pnorm(v_crit, mean = p0, sd = stderror_h0, lower.tail = FALSE)

  # error estándar de H1
  stderror_h1 <- sqrt(p * (1 - p) / n)

  # potencia
  potencia <- pnorm(v_crit, mean = p, sd = stderror_h1, lower.tail = FALSE)
```

```r
  cat(sprintf("n = %d:\n", n))
  cat(sprintf("Valor crítico: %f\n", v_crit))
  cat(sprintf("Error Tipo I ( ): %f\n", error_1))
  cat(sprintf("Potencia (1- ): %f\n\n", potencia))
}
# calculamos para cada una de las parametrizaciones
for (size in n) {
  calcular_errores(size, p0, p)
}
```

```
## n = 6558:
## Valor crítico: 0.004324
## Error Tipo I ( ): 0.024998
## Potencia (1- ): 0.960612
##
## n = 3159:
## Valor crítico: 0.004907
## Error Tipo I ( ): 0.024998
## Potencia (1- ): 0.786795
##
## n = 4597:
## Valor crítico: 0.004581
## Error Tipo I ( ): 0.024998
## Potencia (1- ): 0.893585
```

Ejercicio 9

Ejercicio 10

Para comparar una proporción con el valor de referencia, una prueba adecuada sería la prueba del signo. Nuestras hipótesis para el contraste serían:

H0 : el estadístico de la prueba es mayor que el valor crítico H1 : el estadístico de la prueba es menor o igual que el valor crítico

El código en R para dicha prueba sería:

```r
n <- 1000
p0 <- 0.3
p <- 0.4
r <- p*n

prueba <- binom.test(x = r, n = n, p = p0, alternative = "greater")

print(prueba)
```

```
##
##  Exact binomial test
##
## data:  r and n
## number of successes = 400, number of trials = 1000, p-value = 1.104e-11
## alternative hypothesis: true probability of success is greater than 0.3
## 95 percent confidence interval:
##  0.3742609 1.0000000
## sample estimates:
## probability of success
```

```
##                     0.4
```

```r
# Para comprobar si rechazamos o no H0 (elegimos nivel de significancia del 5%):

if(prueba$p.value < 0.05){
  print("Rechazamos H0")
}else{
    print("No existen evidencias significativas para rechazar H0")
}
```

```
## [1] "Rechazamos H0"
```