Riesgo Atribuible

Christian Badillo, Luis Nuñez, Luz Maria Santana, Sealtiel Pichardo & Liz

Tabla de contenidos

1 Estimación Puntual.

El ***riesgo atribuible*** representa el exceso de riesgo atrbuible al factor de exposición para el desarrollo del evento de interés y se define como:

$$AR = \pi_1 - \pi_2.$$

Donde π_1 representa la proporción de la población que fue expuesta y que desarrollo el evento de interés y π_2 la proporción de la población que no fue expuesta y que presenta el evento de interés.

En el caso de un estudio de casos y controles se puede usar la tabla de contingencia de 2x2.

Tabla 1: Tabla de Contingencia 2x2.

	Evento de Interés		
Exposición	SI	NO	Total
SI	a	b	a + b
NO	c	d	c + d
Total	a + c	b + d	n

Entonces podemos realizar la estimación puntual del Riesgo Atribuible como:

$$\hat{\pi}_1 = \frac{a}{a+b} = \frac{a}{n_{1\bullet}} \quad \text{y} \quad \hat{\pi}_2 = \frac{c}{c+d} = \frac{a}{n_{2\bullet}} \quad \therefore$$

$$\hat{AR} = \hat{\pi}_1 - \hat{\pi}_2$$

2 Estimación por Intervalos.

Para obtener el intervalo de confianza al $100(1-\alpha)\%$, se debe estimar la varianza de nuestro estimador de riesgo atribuible ($\hat{AR}$). Para ello se utilizará el método delta de primer orden para aproximar la varianza del estimador.

$$\mathbb{V}(g(\hat{AR})) \approx \sigma^2 g'(\mu)^2.$$

En este caso $g(\hat{AR}) = \hat{\pi}_1 - \hat{\pi}_2$ por tanto, la derivada parcial respecto a cada estimador es:

$$\frac{\partial \hat{AR}}{\partial \hat{\pi}_1} = 1$$
$$\frac{\partial \hat{AR}}{\partial \hat{\pi}_2} = -1.$$

La varianza de cada una de las proporciones estimadas es:

$$\mathbb{V}(\hat{\pi}_1) = \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_{1\bullet}}$$
$$\mathbb{V}(\hat{\pi}_2) = \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_{2\bullet}}.$$

Entonces podemos aproximar la varianza del riesgo atribuible como:

$$\begin{aligned}\mathbb{V}(\hat{AR}) &\approx (1)^2 \cdot \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_{1\bullet}} + (-1)^2 \cdot \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_{2\bullet}} \\ &\approx \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_{1\bullet}} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_{2\bullet}}\end{aligned}$$

Sustituyendo $\hat{\pi}_1$ y $\hat{\pi}_2$.

$$\begin{aligned}\mathbb{V}(\hat{AR}) &= \frac{\frac{a}{a+b}(1-\frac{a}{a+b})}{a+b} + \frac{\frac{c}{c+d}(1-\frac{c}{c+d})}{c+d} \\ &= \frac{\frac{a}{a+b}(\frac{a+b}{a+b}-\frac{a}{a+b})}{a+b} + \frac{\frac{c}{c+d}(\frac{c+d}{c+d}-\frac{c}{c+d})}{c+d} \\ &= \frac{\frac{a}{a+b}(\frac{b}{a+b})}{a+b} + \frac{\frac{c}{c+d}(\frac{d}{c+d})}{c+d} \\ &= \frac{a}{a+b}\frac{b}{a+b}\frac{1}{a+b} + \frac{c}{c+d}\frac{d}{c+d}\frac{1}{c+d} \\ &= \frac{ab}{(a+b)^3} + \frac{cd}{(c+d)^3} \\ &= \frac{ab}{n_{1\bullet}^3} + \frac{cd}{n_{2\bullet}^3}. \ \square\end{aligned}$$

Entonces el intervalo de confianza se estima como:

$$\left[\hat{AR} - Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{ab}{n_{1\bullet}^3} + \frac{cd}{n_{2\bullet}^3}},\ \ \hat{AR} + Z_{(1-\frac{\alpha}{2})}\sqrt{\frac{ab}{n_{1\bullet}^3} + \frac{cd}{n_{2\bullet}^3}}\right].$$

3 Programación de las Estimaciones.

3.1 Aproximación Teórica.

Se define una función en el lenguaje de programación `R` para la estimación puntual y por intervalo del riesgo atribuible.

```r
# Se define una función
ar_estimation <- function(df, alpha = 0.05){
    if(!all(dim(df) == c(2, 2))) {
        # Check the dim of the data
        stop("The data must be a 2 x 2 data.frame")
    }

    # Specify the values of the 2x2 table
    a <- df[1, 1]; b <- df[1, 2]; c <- df[2, 1]; d <- df[2, 2]

    # Estimate the AR
    ar_est <- a / (a + b) - c / (c + d)

    # Estimate the SE of the AR
    se_ar <- sqrt(a*b/(a+b)^3 + c*d/(c+d)^3)

    # EStimate the Interval of Confidence for the AR.
    ci_ar <- ar_est + c(-1, 1) * qnorm(1 - alpha) * se_ar

    # Return a list with the Estimations.
    return(list(puntual.estimation = ar_est, SE = se_ar, CI = ci_ar))
}
```

Se pone a prueba con la tabla de Hulley y Cummings (1993).

```r
# Data
hm_data <- data.frame(
    infarto = c(40, 10),
    no_infarto = c(460, 490),
    row.names = c("Toma Café", "No Toma Café")
)

ar.theo <- ar_estimation(hm_data)
```

```r
ar.theo$puntual.estimation
```

```
[1] 0.06
```

```r
ar.theo$SE
```

```
[1] 0.01365284
```

```r
ar.theo$CI
```

```
[1] 0.03754308 0.08245692
```

3.2 Bootstrapping

Se define una función en R para la estimación puntual y por intervalos por medio de bootstrapping.

```r
# Definimos una función para muestreo con bootstrapping
ar_bootstrap <- function(data, sims_num = 1000, seed = 140801){
    # Data
    a <- data[1, 1]; b <- data[1, 2]; c <- data[2, 1]; d <- data[2, 2]

    # Sample Size
    n <- a + b + c + d

    # Define where we consider a sample bootstrap be a, b, c and, d
    define_a <- 1:a
    define_b <- (a+1):(b + a)
    define_c <- (b+a+1):(c + a + b)
    define_d <- (a+b+a+c+1):(d + a + b + c)

    # Define the vector to store the estimations and its size
    ar_est <- numeric(length = sims_num)

    #Set the seed
    set.seed(seed)

    # Run the Bootstrap
    for (i in 1:sims_num) {

        # Sample ID
        idx_samples <- sample(1:n, replace = T, size = n)

        # How many a, b, c, d are in the sample?
        sample_a <- sum(define_a[1] <= idx_samples & idx_samples <= tail(define_a, n=1))
        sample_b <- sum(define_b[1] <= idx_samples & idx_samples <= tail(define_b, n=1))
        sample_c <- sum(define_c[1] <= idx_samples & idx_samples <= tail(define_c, n=1))
        sample_d <- sum(define_d[1] <= idx_samples & idx_samples <= tail(define_d, n=1))

        # Store the estimated AR of the i-th simulation.
        ar_est[i] <- sample_a / (sample_a + sample_b) - sample_c / (sample_c + sample_d)
    }
    return(ar_est)
}

# Run Bootstrapping
ar.boot <- ar_bootstrap(hm_data, 15000, seed = 123456789)
```

La estimación puntual es:

```r
mean(ar.boot)
```

```
[1] 0.05826044
```

Y la estimación por intervalos es:

```r
quantile(ar.boot, c(0.025, 0.975))
```

```
      2.5%      97.5%
0.03112889 0.08565380
```

La estimación puntual y por intervalos usando máxima verosimilitud es:

```r
ar.theo$puntual.estimation
```

```
[1] 0.06
```

```r
ar.theo$CI
```

```
[1] 0.03754308 0.08245692
```