Ejercicios de estimación por intervalos

Ricardo Mayer

Octubre de 2019

Fórmulas relevantes (ver Anderson y Sweeny, cap 8, página 326)

Estimación por intervalo de una media poblacional, σ conocida:

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Estimación por intervalo de una media poblacional, σ desconocida:

$$\bar{x} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$$

Tamaño de la muestra para una estimación de la media poblacional:

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2}$$

Estimación por intervalo de una proporción poblacional:

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Tamaño de la muestra para una estimación de la proporción poblacional:

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2}$$

Donde $z_{\alpha/2}$ es igual a -1,645 o -1,96 o -2,576 para α igual a 10, 5 y 1 por ciento, respectivamente

Ejercicios

El siguiente vector contiene un muestra con observaciones de precios de departamentos, denominados en U.F.

```r
pd <- c(2500, 2600, 2700, 2850, 3200, 3300, 2300, 3000, 2800, 2500, 2600, 2400, 3100, 3000,
        2700, 2900, 2200, 3000, 2660, 2700, 2500, 2680, 2700, 2400, 2480, 2300, 3200, 2900,
        2900, 2800, 2700, 3000, 2800, 2600, 2700, 2900, 2900, 3000, 2850, 2690, 2700, 2700)
```

Como vemos, el tamaño de la muestra es 42, la media muestral es igual a 2747,857 U.F. y la desviación estóndar muestral es 251,789 U.F. (doscientos cincuenta y uno coma setecientos ochenta y nueve U.F.)

```r
n <- length(pd)
xbar <- mean(pd)
s <- sd(pd)

n
```

```
## [1] 42
```

```r
round(xbar, digits = 3)
```

```
## [1] 2747.857
```

```r
round(s, digits = 3)
```

```
## [1] 251.789
```

1. *Estimación por intervalo de una media poblacional: σ conocida*

Asuma que la desviación estándar poblacional es igual a 250 U.F. Encuentre el margen de error y el intervalo de confianza al 90%. Luego repita el cálculo para un 95 y 99 por ciento de confianza

Utilizando las fórmulas escritas más arriba y tomando en cuenta que $\sigma = 250, n = 42, z_{\alpha/2} = 1,645$:

$$IC : \bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$IC : 2747,857 \pm 1,645\frac{250}{42}$$
$$IC : 2747,857 \pm 63,45725$$
$$IC : (2684,4, 2811,314)$$

Y el margen de error es igual a 63,45725 U.F.

Los cálculos para el 95 y 99 por ciento de confianza son muy similares y solamente cambia el valor de $z_{\alpha/2}$

2. *Estimación por intervalo de una media poblacional: σ conocida*

Calcule los intervalos de confianza y los márgenes de error para los casos de 90, 95 y 99 por ciento de confianza. Para dicho cálculo necesitará los cuantiles relevantes de la distribución t con 41 grados de libertad: -1.683, -2.02 y -2,701 respectivamente

Como indica la fórmula respectiva, este caso es similar al primer ejecicio, excepto que debemos cambiar $z_{\alpha/2}$ por $t_{41,\alpha/2}$, y cambiar $\sigma = 250$ por $s = 251,789$ (respuesta parcial: para $\alpha = 10\%$, el margen de error es 65,38772 U.F.) 3. *Tamaño de la muestra para una estimación de la media poblacional* Como la muestra aún no se ha recogido, no es posible usar la desviación estándar muestral propiamente tal, por lo que se usa un valor σ como si fuera el poblacional (ver el libro para una discusión al respecto). Asuma que $\sigma = 250$ y encuentre el n necesario para que el margen de error (o error muestral máximo) sea igual a 50 U.F., con un nivel de confianza de 95 por ciento. Luego repita su estimación para un margen de error de 5 U.F. Finalmente repita el ejercicio para ambos márgenes de error, pero cambie el nivel de confianza a 90 por ciento (note que el margen de error siempre está en las mismas unidades de lo que estamos estimando, que en estos casos es el precio promedio poblacional de los departamentos y está en U.F.)

Para este caso, tenemos que

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2}$$
$$n = \frac{(-1,96)^2 250^2}{50^2}$$
$$n = 96.04$$

Note que n es ligeramente mayor a 96 y el número de observaciones obviamente debe ser un número entero, por lo tanto si queremos que el margen de error sea menor a las 50 U.F. necesitamos un tamaño muestral de al menos 97.

3. *Estimación por intervalo de una proporción poblacional*

Supongamos que estamos interesados en estimar que porcentaje de todos los departamentos en la población permiten vivir con mascotas. En la muestra de 42 departamentos que logramos levantar, el 15 por ciento de los departamentos son pet-friendly. Utilice dicha información para construir un intervalo de confianza, al 99 por ciento, para la proporción poblacional. Reporte el margen de error asociado a dicha estimación. Luego, repita el ejercicio para niveles de confianza de 90 y 95 por ciento

De las fórmulas listadas arriba, la relevante para este caso es

$$\bar{p} \pm 2,576\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Y con los datos del enunciado tenemos que

$$0,15 \pm 2,576\sqrt{\frac{0,15 * 0,85}{42}}$$
$$0,15 \pm 0,1419307$$
$$I.C.: (0.008069313, 0.2919307)$$

Es decir, con un 99% de confianza la proporción poblacional estarìa entre 0,81% y 29,19% de los departamentos y el margen de error para este nivel de confianza es igual a 14,19307% . Note que està expresado en las mismas unidades de lo que queremos estimar, que es un porcentaje o proporciòn.

4. *Tamaño de la muestra para una estimación de la proporción poblacional*

Ahora, sin tener ninguna informaciòn previa de cuanto podría ser la proporción poblacional determine el tamaño muestral necesario para tener un error muestral menor o igual a 4 por ciento, con un nivel de confianza de 99 por ciento. Luego repita el cálculo, pero esta vez utilice una proporcion poblacional sugerida por un estudio anterior que indica que la proporción sería igual a 30 por ciento. Finalmente, repita todos los cálculos para niveles de confianza de 90 y 95 por ciento

Bueno, en el primer caso, sin otra información, debemos utilizar como $p^* = 0,5$, que equivale a nuestro peor escenario:

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2}$$
$$n = \frac{2,576^2 * 0,5 * 0,5}{0.04^2}$$
$$n = 402,5$$

Por lo tanto necesitamos un tamaño muestral de 403 para que nuestro margen de error no sea mayor a 4 por ciento, con este nivel de confianza.

El segundo caso se calcula exactamente igual pero usando $p^* = 0,3$ y la respuesta final es 338.1 y por lo tanto debemos tener 339 observaciones.

Para los otros niveles de confianza solamente cambia el valor de z a utilizar