Estimación puntual

Christian Limbert Paredes Aguilera

2/12/2021

La media muestral

Sea $X_1, \ldots X_n$ una muestra aleatoria simple de tamaño n de una v.a. X de esperanza μ_X y desviación típica σ_X La media muestral es:

$$\overline{X} = \frac{X_1 + \ldots + X_n}{n}$$

En estas condiciones,

$$E(\overline{X}) = \mu_X, \qquad \sigma_X = \frac{\sigma_X}{\sqrt{n}}$$

donde $\sigma_{\overline{X}}$ es el error estándar de $\overline{X}$

- Es un estimador puntual de μ_X
- $E(\overline{X}) = \mu_X$: el valor esperado de de $\overline{X}$ es μ_X.
- Si tomamos muchas veces una m.a.s. y calculamos la media muestral, el valor medio de estas medias tiende con mucha probabilidad a ser μ_X
- $\sigma_{\overline{X}} = \sigma_X/\sqrt{n}$: la variabilidad de los resultados de $\overline{X}$ tiende a 0 a medida que tomamos muestras más grandes.

Ejercicio

1. Generar 10000 muestra de tamaño 40 con repocisión de las longitudes del pétalo.
2. A continuación hallaremos los valores medios de cada muestra.
3. Consideraremos la media y la desviación típica de dichos valores medios y los compararemos con los valores exactos dados por las propiedades de la media muestral.

```
set.seed(1001)
#1
valores.medios.long.pétalo <- replicate(10000,mean(sample(iris$Petal.Length,
                              40,
                              replace = TRUE)))
head(valores.medios.long.pétalo,10)
```

```
##  [1] 3.5975 3.5150 3.9400 3.2650 3.9125 3.9650 4.2825 3.2950 3.8500 3.7850
```

```
#2
mean(valores.medios.long.pétalo)
```

```
## [1] 3.754478
```

```
sd(valores.medios.long.pétalo)
```

```
## [1] 0.2796513
```

```r
#3
mean(iris$Petal.Length)
```

```
## [1] 3.758
```

```r
sd(iris$Petal.Length)/sqrt(40)
```

```
## [1] 0.2791182
```

poblaciones normales

Combinación lineal de distribuciones normales

La combinación lineal de distribuciones normlaes es normal. es decir, si $Y_i, \ldots Y_n$ son v.a. normales independientes, cada $Y_i \sim N(\mu_i, \sigma_i)$ y $a_1, \ldots, a_n, b \in \mathbb{R}$ entonces

$$Y = a_i Y_i + \ldots a_n Y_n + b$$

es una v.a. $N(\mu, \sigma)$ con μ y σ las que correspondan:

- $E(Y) a_i \cdot \mu_i + \ldots a_n \cdot \mu_n + b$
- $\sigma\left(Y\right)^2 = a_1^2 \cdot \sigma_1^2 + \ldots + a_n^2 \cdot \sigma_n^2$

Distribución de la media muestral

Sea $X_1, \ldots, X_n$ una m.a.s de una v.a. X de esperanza μ_X y desviación típica σ_X Si X es $N(\mu_X, \sigma_X)$ entonces

$$\overline{X} \text{ es } N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$$

y por tanto

$$Z = \frac{\overline{X}}{\frac{\sigma_X}{\sqrt{n}}} \text{ es } N(0,1)$$

Teorema central del límite

Sea $SX_1, \ldots, X_n$ una m.a.s. de una v.a. X cualquiera, de esperanza μ_X y desviación típica σ_X. Cuando $n \to \infty$

$$\overline{X} \to N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$$

y por tanto

$$Z = \frac{\overline{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} \to N(0,1)$$

Esta convergencia se refiere a las distribuciones

Caso n grande: Si n es grande $n \geq 30$, $\overline{X}$ es aproximadamente normal, con esperanza μ_X y desviación típica $\frac{\sigma_X}{\sqrt{n}}$