Incertidumbre o Error (de una estimación)

Basado en la clase de Estadística para químicos de Mariela Sued

Estadística

POBLACION $\leftrightarrow F$	MUESTRA $X_1, \ldots X_n$ i.i.d. $X_i \sim F$
Parámetro: Valor asociado de F	Estimador:estadístico para estimar θ
$\theta = \theta(F)$	$\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$
θ: valor poblacional	$\widehat{\theta}_n$ NUEVA VARIABLE ALEATORIA

Estimación

Point estimation refers to providing a single "best guess" of some quantity of interest.
All of statistics. Wasserman

- $\theta =$ Objeto de interés.
- best guess: Estimador:*cuenta hecha con la muestra*
- Estimador: Función de la muestra

$$\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$$

- Estimación: Valor del estimador en un conjunto de datos:

$$\widehat{\theta}_{n\ obs} = \widehat{\theta}_n(x_1, \ldots, x_n)$$

Notemos que el estimador ...

$$\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$$

- $\widehat{\theta}_n$ es una variable aleatoria.
- $\widehat{\theta}_n$ tiene distribución.

Sampling distribution of $\widehat{\theta}_n$*:* $f_{\widehat{\theta}_n}$

- $\widehat{\theta}_n$ tiene esperanza: $\mathbb{E}(\widehat{\theta}_n) = \int u f_{\widehat{\theta}_n}(u) du$
- $\widehat{\theta}_n$ tiene varianza: $\mathbb{V}(\widehat{\theta}_n)$
- $\widehat{\theta}_n$ tiene (en general) desvío estandar.

$$\mathsf{se} = \mathsf{se}(\widehat{\theta}_n) = \sqrt{\mathbb{V}(\widehat{\theta}_n)} \quad \textit{Standard error of } \widehat{\theta}_n.$$

Cómo se informa: $\pm$

> On the other hand, the invariance of retrograde speeds distribution suggests that the endogenous and KHC576-mCherry-Pex26 motors contribute similarly to retrograde transport. In addition, KHC576-mCherry-Pex26 did not affect the median values of plus and minus-end directed run lengths (Table 1) and the number of reversions in the trajectories (1.8 $\pm$ 0.1 reversions/trajectory).

$$\text{estimación } (\pm \text{ incertidumbre}),$$

$$\text{estimación } (\pm \text{ error})$$

- Se informa poniendo el valor de la estimación y **"su error"** - **uncertainty**.
- ¿A qué llamamos ahora **"error"** (uncertainty)?

Incertidumbre - Error de Estimación para $\widehat{\mu}_n$

- Parámetro de interés: $\mu = \mathbb{E}(X)$. Estimador: $\widehat{\mu}_n = \overline{X}_n$.
- Estimación: Promedio de mis datos; en R: `mean(datos)`
- Varianza y desvío estandar del estimador:

$$V(\widehat{\mu}_n) = \frac{V(X)}{n}\,, \quad \text{se} = \sqrt{V(\widehat{\mu}_n)} = \sqrt{\frac{V(X)}{n}}$$

- Incertidumbre -Error de Estimación
  1. Si $V(X) = \sigma_0^2$ es un valor conocido, el error de estimación es

     $$\text{Incertidumbre - Error de estimación : se} = \sigma_0/\sqrt{n}$$

  2. Si $V(X)$ es desconocida, estimamos se:

     $$\text{Incertidumbre - Error de estimación : } \widehat{\text{se}}_{\text{obs}} = S_{\text{obs}}/\sqrt{n}$$

     en R : `sd(datos)/sqrt(length(datos))`

Incertidumbre - Error de Estimación para $\widehat{p}_n$

- $Y_i \sim \mathcal{B}(1,p)$. Parámetro de interés: $p = \mathbb{E}(Y)$.
- Estimador: $\widehat{p}_n = \overline{Y}_n$.
- Estimación: Promedio de mis datos; en R: `mean(datos)`
- Varianza y desvío estandar del estimador:

$$V(\widehat{p}_n) = \frac{V(Y)}{n}\,, \quad \text{se} = \sqrt{V(\widehat{p}_n)} = \sqrt{\frac{V(Y)}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

Incertidumbre - Error de estimación :

$$\widehat{\text{se}}_{\text{obs}} = \sqrt{\widehat{p}_{n,\text{obs}}(1-\widehat{p}_{n,\text{obs}})/n}$$

en R : `mean(datos)*(1-mean(datos))/sqrt(length(datos))`

Error de estimación - Uncertainty - Incertidumbre

Definición: llamamos incertidumbre (de la estimación) o error de una estimación a la estimación del desvío (exacto o aproximado) del estimador con el cual estimamos.

- Estimador: $\widehat{\theta}_n$
- Estimación: $\widehat{\theta}_{n,\text{obs}}$
- $\text{se} = \sqrt{V(\widehat{\theta}_n)}$ o $\text{se} \approx \sqrt{V(\widehat{\theta}_n)}$
- Incertidumbre o Error de Estimación: $\widehat{\text{se}}_{\text{obs}}$

Resumiendo: usamos datos para obtener observados de

$$\widehat{\mu}_n \; , \quad \mathsf{se}^2 = V(\widehat{\mu}_n) = \sigma^2/n \; , \;\; \widehat{\mathsf{se}} = \sqrt{S^2/n} \; ,$$
$$\widehat{p}_n \; , \;\; \mathsf{se}^2 = V(\widehat{p}_n) = p(1-p)/n \; , \; \widehat{\mathsf{se}} = \sqrt{\widehat{p}_n(1-\widehat{p}_n)/n} \; ,$$