

# Muestreo Estadístico: Muestreo Aleatorio Simple con y sin Reemplazamiento \*

García Prado, Sergio  
sergio@garciparedes.me

8 de octubre de 2017

## 1. Demostraciones

A continuación se demuestran distintas propiedades relacionadas con los estadísticos referidos al muestreo aleatorio simple con y sin reemplazamiento.

### 1.1. Probabilidad de inclusión de primer y segundo nivel en muestreo aleatorio simple sin reemplazamiento (m.a.s)

La probabilidad de inclusión en el muestreo aleatorio simple sin reemplazamiento viene condicionada por dos factores diferentes: (a) el número de posibles muestras que incluyan el elemento  $k$ -ésimo determinado por  $\binom{N-1}{n-1}$  y el de posibles muestras que incluyan los elementos  $k$  y  $l$  tales que  $k \neq l$  determinado por  $\binom{N-2}{n-2}$  así como, (b) la probabilidad de elegir una determinada muestra de tamaño  $n$  de entre todas las posibles, definida como  $\frac{1}{\binom{N}{n}}$ .

Nótese que dichos valores provienen del área de teoría combinatoria, donde se cumple la siguiente propiedad: “Sean  $n$  elementos distintos y  $k \leq n$ . A las distintas agrupaciones no ordenadas de  $k$  elementos elegidos entre  $n$  distintos se las denomina combinaciones de  $n$  elementos tomados en grupos de  $k$ , y el número de dichas combinaciones es  $\binom{n}{k}$ ” [LMPMSB14]

Por tanto, la probabilidad de inclusión queda determinada por el sumatorio de las probabilidades de todas aquellas muestras que incluyen al elemento  $k$  para el caso del primer nivel y de los elementos  $\{k, l\}$ , con  $k \neq l$  para el segundo nivel:

$$\pi_k = \sum_{k \in s} p(s) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \frac{n}{N} \quad (1)$$

$$\pi_{kl} = \sum_{\{k, l\} \in s} p(s) = \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)} \quad (2)$$

### 1.2. Probabilidad de inclusión de primer nivel en muestreo aleatorio con reemplazamiento

Para el caso del muestreo aleatorio simple con reemplazamiento, la probabilidad de inclusión de primer nivel se define de manera diferente respecto de la del caso sin reemplazamiento. Esto se debe a que una observación  $k$  puede aparecer más de una vez en una determinada muestra  $s$ .

---

\*URL: <https://github.com/garciparedes/statistical-sampling-simple-random-sample>

En este caso, para probar la probabilidad de inclusión del elemento  $k$ -ésimo en la muestra nos apoyaremos en la distribución *Binomial*, que modeliza la probabilidad de selección con reemplazamiento de un elemento de entre  $N$  posibles con probabilidad  $p$ . En este caso escogeremos  $p = \frac{1}{N}$ , que es la probabilidad de selección de una observación en la muestra.

Por tanto, definiremos la variable  $X \sim B(n, \frac{1}{N})$ . Para probar la probabilidad de que el elemento sea seleccionado al menos una vez, necesitamos conocer  $P(X \geq 1) = P(X = 1) + P(X = 2) + \dots + P(X = n)$ . Sin embargo, esta probabilidad puede ser calculada de manera sencilla como  $P(X \geq 1) = 1 - P(X = 0)$ . Puesto que  $P(X = 0) = \binom{n}{0} (\frac{1}{N})^0 (1 - \frac{1}{N})^{n-0} = (1 - \frac{1}{N})^n$ . Dado que la probabilidad de inclusión  $\pi_k = P(X \geq 1)$  por definición, entonces:

$$\pi_k = 1 - \left(1 - \frac{1}{N}\right)^n \quad (3)$$

### 1.3. Varianza del $\pi$ -estimador para proporciones en muestreo aleatorio simple sin reemplazamiento (m.a.s)

$$\widehat{Var}(\widehat{P}_\pi) = \frac{1-f}{n-1} \widehat{P}_\pi (1 - \widehat{P}_\pi) \quad (4)$$

### 1.4. Intervalo de confianza del $\pi$ -estimador para el total poblacional en muestreo aleatorio simple sin reemplazamiento (m.a.s)

$$Pr\left(\tau \in \left[\widehat{\tau}_\pi \pm z_{1-\alpha/2} \sqrt{\widehat{Var}[\widehat{\tau}_\pi]}\right]\right) = 1 - \alpha \quad (5)$$

## Referencias

- [LMPMSB14] Esperanza Larrinaga Miner, María Felisa Pérez Martínez, and Araceli Suárez Barrio. Grado en ingeniería informática, uva: Matemática discreta. 2014.
- [TG18] Jesús Alberto Tapia García. Muestreo Estadístico 1, 2017/18.