Formulario Muestreo Aleatorio Simple (MAS) “‘ python print(”ASD”) “‘

- Número de muestras distintas de tamaño N sin reemplazo de una población de tamaño N:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (3.01)$$

- Demostración: Se explica en la Sección 3.1 y se ejemplifica en el Ejemplo 3.1. - Número de muestras posibles de tamaño N con reemplazo y sin orden de una población de tamaño N:

$$\binom{N+n-1}{n} = \frac{(N+n-1)!}{n!(N-1)!} \quad (3.02)$$

- Demostración: Se explica en la Sección 3.2 y se ejemplifica en el Ejemplo 3.2. - Número de muestras posibles de tamaño N con reemplazo y con orden de una población de tamaño N:

$$N^n$$

- Demostración: Se explica en la Sección 3.2 y se ejemplifica en el Ejemplo 3.3.

Probabilidad de Selección

- Probabilidad de que una unidad de la población esté en la muestra en MAS sin reemplazo:

$$\pi_i = \frac{n}{N}$$

- Demostración: Se explica en detalle en la Sección 3.3, considerando la probabilidad de selección en cada ocasión. - Probabilidad de selección conjunta de dos unidades muestrales u_i y u_j en MAS sin reemplazo:

$$\pi_{ij} = \frac{n(n-1)}{N(N-1)}$$

- Demostración: Se deriva al considerar la variable aleatoria producto $e_i e_j$ y su esperanza en la Sección 3.3. - Probabilidad de selección con reemplazo:

- Para cualquier elemento específico en cada una de las N ocasiones:

$$\frac{1}{N}$$

- Los resultados de cada ocasión son independientes.

Propiedades de las Probabilidades Simples y Conjuntas

- Propiedad 1:

$$\sum_{i=1}^{N} \pi_i = n$$

- Demostración: Se demuestra en la Sección 3.4.1 utilizando la esperanza de la suma de las variables auxiliares e_i. - Propiedad 2:

$$\sum_{i=1}^{N} \pi_i^2 = n\pi_i - \sum_{i\neq l}^{N} \pi_{il}$$

- Demostración: Se demuestra en la Sección 3.4.1 relacionándola con la Propiedad 1. - Propiedad 3:

$$\sum_{l\neq i}^{N} \pi_{il} = (n-1)\pi_i$$

- Demostración: Se demuestra en la Sección 3.4.1 utilizando la esperanza de un producto de sumas de variables auxiliares. - Propiedad 4:

$$\sum_{l\neq i}^{N} (\pi_{il} - \pi_i\pi_l) = -\pi_i(1-\pi_i)$$

- Demostración: Se verifica en la Sección 3.4.1 utilizando las Propiedades 2 y 3. - Propiedad 5:

$$\sum_{i\neq l}^{N} \pi_{il} = n(n-1)$$

- Demostración: Se verifica en la Sección 3.4.1 utilizando la Propiedad 3.

Parámetros y Estimadores

- Estimador general de Horvitz y Thompson para el total θ:

$$\hat{\theta}_{HT} = \sum_{i=1}^{n} w_i X_i$$

donde $w_i = 1/\pi_i$. Para MAS, $w_i = N/n$.

$$\hat{T} = \sum_{i=1}^{n} \frac{N}{n} X_i = N\bar{X} \quad (3.12)$$

- Condición de insesgamiento: $E(\hat{\theta}_{HT}) = \theta$. La demostración se encuentra en la Sección 3.5.1. - Media Poblacional:

$$\mu_X = \frac{1}{N}\sum_{i=1}^{N} X_i \quad (3.13)$$

- Estimador de la media muestral (insesgado):

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (3.14, 3.21)$$

- Demostración de insesgamiento: Se muestra en la Sección 3.6. - Total Poblacional:

$$T_X = \sum_{i=1}^{N} X_i \quad (3.15)$$

- Estimador del total (insesgado):

$$\hat{T}_X = N\bar{X} = \frac{N}{n}\sum_{i=1}^{n} X_i \quad (3.16)$$

- Demostración de insesgamiento: Se sigue del insesgamiento de la media muestral. - Proporción Poblacional:

$$P_X = \frac{A_X}{N} = \frac{\sum_{i=1}^{N} X_i}{N} \quad (3.17)$$

donde X_i es 1 si la unidad tiene el atributo y 0 si no.

- Estimador de la proporción muestral (insesgado):

$$\hat{P}_X = p_x = \frac{a_x}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \quad (3.18)$$

- Demostración de insesgamiento: Se menciona en la Sección 3.6, ya que la proporción es una media para valores dicotómicos. - Razón Poblacional:

$$R_0 = \frac{T_A}{T_B} = \frac{\sum_{i=1}^{N} A_i}{\sum_{i=1}^{N} B_i}$$

- Estimador de la razón muestral (sesgado):

$$R_y = \frac{\bar{y}}{\bar{x}} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}$$

- Se indica que es sesgado en la Sección 3.6.

Varianza de los Estimadores

- Varianza del estimador del total $\hat{T}_X$:

$$V(\hat{T}_X) = \sum_{i=1}^{N} (\frac{X_i}{\pi_i})^2 V(e_i) + \sum_{i=1}^{N} \sum_{j \neq i}^{N} (\frac{X_i}{\pi_i})(\frac{X_j}{\pi_j}) Cov(e_i, e_j) \quad (3.23)$$

Para MAS ($\pi_i = n/N$, $\pi_{ij} = n(n-1)/(N(N-1))$):

$$V(\hat{T}_X) = N^2(1-f)\frac{S_X^2}{n} \quad (3.27)$$

donde $f = n/N$ y $S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \mu_X)^2$ (cuasivarianza poblacional).

- Demostración: Se desarrolla en la Sección 3.7 utilizando la varianza y covarianza de las variables auxiliares.
- Varianza del estimador de la media $\bar{X}$:

- Sin reemplazo:

$$V(\bar{X}) = (1-f)\frac{S_X^2}{n} = \frac{N-n}{N}\frac{S_X^2}{n} \quad (3.29)$$

- Con reemplazo:

$$V(\bar{X}) = \frac{\sigma_X^2}{n} = \frac{N-1}{N}\frac{S_X^2}{n} \quad (3.30)$$

donde $\sigma_X^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu_X)^2$ (varianza poblacional).
- Demostración: Se detalla en la Sección 3.8. - Varianza del estimador de la proporción $\hat{P}_X = p_x$:

$$V(p_x) = (1-f)\frac{P_X Q_X}{n-1}\frac{N}{N-1} \approx (1-f)\frac{P_X Q_X}{n} \quad (3.42)$$

donde $Q_X = 1 - P_X$.

- Demostración: Se sigue un procedimiento similar al de la media en la Sección 3.9. - Varianza del estimador del total de casos $\hat{A}_X = Np_x$:

$$V(\hat{A}_X) = N^2(1-f)\frac{P_X Q_X}{n-1}\frac{N}{N-1} \approx N^2(1-f)\frac{P_X Q_X}{n}$$

- Derivación: Se obtiene multiplicando por N^2 la varianza del estimador de la proporción.

Estimador de Varianza

- Estimador general de varianza para el total $\hat{V}(\hat{T}_X)$:

$$\hat{V}(\hat{T}_X) = \sum_{i=1}^{n} (\frac{X_i}{\pi_i})^2 (1-\pi_i) + \sum_{i=1}^{n} \sum_{j \neq i}^{n} (\frac{X_i}{\pi_i})(\frac{X_j}{\pi_j})(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}})$$

Para MAS:

$$\hat{V}(\hat{T}_X) = N^2(1-f)\frac{s_x^2}{n} \quad (3.48)$$

donde $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ (cuasivarianza muestral).

- Se espera que sea insesgado: $E[\hat{V}(\hat{T}_X)] = V(\hat{T}_X)$. La verificación para MAS se encuentra en la Sección 3.10.
- Estimador de varianza para la media $\hat{V}(\bar{X})$:

- Sin reemplazo (insesgado):

$$\hat{V}(\bar{X}) = (1-f)\frac{s_x^2}{n} = \frac{N-n}{Nn}s_x^2 \quad (3.52)$$

- Demostración de insesgamiento: Se presenta en la Sección 3.11. - Con reemplazo (sesgado):

$$\hat{V}(\bar{X}) = \frac{s_x^2}{n}$$

- Se demuestra que es sesgado en la Sección 3.11. - Estimador de varianza para la proporción $\hat{V}(p_x)$:

$$\hat{V}(p_x) = (1-f)\frac{p_x q_x}{n-1} \quad (3.65)$$

donde $q_x = 1 - p_x$.

- Estimador de varianza para el total de casos $\hat{V}(\hat{A}_X) = N^2\hat{V}(p_x)$:

$$\hat{V}(\hat{A}_X) = N^2(1-f)\frac{p_x q_x}{n-1}$$

Precisión y Acuracidad

- Error Cuadrático Medio (ECM) de un estimador $\hat{\theta}$:

$$ECM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = V(\hat{\theta}) + [Sesgo(\hat{\theta})]^2 \quad (3.67)$$

- Precisión: Medida por la varianza $V(\hat{\theta})$. - Acuracidad: Medida por el ECM, que considera tanto la varianza como el sesgo.

Estimación por Intervalos

- Intervalo de Confianza General para un parámetro θ:

$$\hat{\theta} \pm k\sqrt{\hat{V}(\hat{\theta})}$$

donde k depende del nivel de confianza (generalmente se usa $z_{\alpha/2}$ para muestras grandes o distribución normal asumida, y $t_{n-1,\alpha/2}$ para muestras pequeñas con varianza desconocida estimada).

Tamaño de Muestra (MAS)

- Para estimar la media μ_X con margen de error absoluto e:

$$n \geq \frac{k^2 S_X^2}{e^2 + \frac{k^2 S_X^2}{N}} \quad (3.77)$$

Si N es grande: $n_0 = \frac{k^2 S_X^2}{e^2}$. Entonces $n = \frac{n_0}{1+\frac{n_0}{N}}$ (3.80).

- Para estimar el total T_X con margen de error absoluto E:

$$n \geq \frac{k^2 N^2 S_X^2}{E^2 + k^2 N S_X^2} \quad (3.83)$$

- Para estimar la media μ_X con margen de error relativo e':

$$n \geq \frac{k^2 S_X^2}{e'^2 \mu_X^2 + \frac{k^2 S_X^2}{N}} \quad (3.87)$$

Si N es grande: $n_0 = \frac{k^2 CV_X^2}{e'^2}$, donde $CV_X = S_X/\mu_X$. Entonces $n = \frac{n_0}{1+\frac{n_0}{N}}$ (3.88).

- Para estimar el total T_X con margen de error relativo e': La fórmula es la misma que para la media con margen de error relativo.

- Para estimar la proporción P_X con margen de error absoluto e:

$$n \geq \frac{N k^2 P_X Q_X}{e^2(N-1) + k^2 P_X Q_X} \quad (3.93)$$

Si N es grande: $n_0 = \frac{k^2 P_X Q_X}{e^2}$. Entonces $n = \frac{n_0}{1+\frac{n_0}{N}}$ (3.94). Usar $P_X = 0.5$ para máxima varianza si no se conoce P_X.

Errores de Muestreo y No de Muestreo

- Error de Muestreo Relativo (Coeficiente de Variación) de un estimador $\hat{\theta}$:

$$CV(\hat{\theta}) = \frac{\sqrt{\hat{V}(\hat{\theta})}}{|\hat{\theta}|} \times 100\%$$