

Ejercicios Propuestos

Juan Rubio Cobeta

January 16, 2026

Contents

1	Apartado a)	2
1.1	Solución	2
2	Apartado b)	3
2.1	Solución	3
3	Apartado c)	5
3.1	Solución	5

1 Apartado a)

La cantidad $\theta = p(1 - p)$ donde p es una proporción de la población, se puede estimar mediante $\hat{\theta} = \hat{p}(1 - \hat{p})$. Suponiendo que \hat{p} es un estimador insesgado de p , utiliza la técnica de linealización para calcular un estimador de la varianza de $\hat{\theta}$.

1.1 Solución

Para resolver este problema, aplicaremos el método de Linealización de Taylor (Series de Taylor de primer orden), adecuado para aproximar la varianza de estimadores que son funciones no lineales de otros parámetros.

1. Definición de la función

Sea $\hat{\theta}$ una función de \hat{p} , definida como:

$$f(\hat{p}) = \hat{p}(1 - \hat{p}) = \hat{p} - \hat{p}^2 \quad (1)$$

Esta es una función no lineal. Para estimar su varianza, necesitamos linealizarla alrededor del verdadero valor del parámetro p .

2. Cálculo de la derivada

Calculamos la primera derivada de la función $f(x) = x - x^2$ respecto a x :

$$f'(x) = \frac{d}{dx}(x - x^2) = 1 - 2x \quad (2)$$

3. Aplicación de la fórmula de linealización

Según la técnica de linealización, la varianza aproximada de un estimador transformado $f(\hat{p})$ viene dada por:

$$V(\hat{\theta}) \approx [f'(p)]^2 V(\hat{p}) \quad (3)$$

Sustituyendo nuestra derivada:

$$V(\hat{\theta}) \approx (1 - 2p)^2 V(\hat{p}) \quad (4)$$

4. Estimador de la varianza

Para obtener el *estimador* de la varianza (que podemos calcular con los datos de la muestra), sustituimos el parámetro poblacional p por su estimador \hat{p} y la varianza poblacional $V(\hat{p})$ por su estimador $\hat{V}(\hat{p})$:

$$\hat{V}(\hat{\theta}) = [f'(\hat{p})]^2 \hat{V}(\hat{p}) \quad (5)$$

Sustituyendo la expresión de la derivada obtenida en el paso 2:

$$\boxed{\hat{V}(\hat{\theta}) = (1 - 2\hat{p})^2 \hat{V}(\hat{p})} \quad (6)$$

Donde $\hat{V}(\hat{p})$ es el estimador de la varianza de la proporción, cuyo cálculo dependerá del diseño muestral específico utilizado.

2 Apartado b)

Los siguientes datos corresponden a una muestra aleatoria simple obtenida de una población de 1230 individuos:

Individuo	1	2	3	4	5	6	7	8
Edad	16	12	28	32	19	54	21	48
Viven	SI	SI	NO	NO	NO	NO	SI	NO

donde *viven* indica si viven con ambos padres o no.

Utiliza el método de linealización para estimar la varianza del parámetro $\hat{p}(1 - \hat{p})$ siendo p :

1. La proporción de individuos que tienen menos de 30 años.
2. La proporción de individuos que viven con ambos padres.

2.1 Solución

Se dispone de una Muestra Aleatoria Simple (MAS) de tamaño $n = 8$ extraída de una población finita de tamaño $N = 1230$. El objetivo es estimar la varianza del estimador no lineal $\hat{\theta} = \hat{p}(1 - \hat{p})$ utilizando la técnica de linealización de Taylor de primer orden.

Fórmulas y Datos Previos

Del apartado anterior, sabemos que el estimador de la varianza linealizada viene dado por:

$$\widehat{V}(\hat{\theta}) \approx (1 - 2\hat{p})^2 \widehat{V}(\hat{p}) \quad (7)$$

Dado que el diseño es un Muestreo Aleatorio Simple sin reemplazamiento, la varianza estimada de la proporción $\widehat{V}(\hat{p})$ se calcula como:

$$\widehat{V}(\hat{p}) = (1 - f) \frac{\hat{p}(1 - \hat{p})}{n - 1} \quad (8)$$

donde el factor de corrección por población finita es:

$$1 - f = 1 - \frac{n}{N} = 1 - \frac{8}{1230} = \frac{1222}{1230} \approx 0,9935$$

Caso 1: Proporción de individuos con menos de 30 años

1. Estimación puntual de la proporción (\hat{p}):

Analizando la muestra, los individuos con edad inferior a 30 años son los individuos 1, 2, 3, 5 y 7.

$$n_{<30} = 5 \quad \Rightarrow \quad \hat{p} = \frac{5}{8} = 0,625$$

2. Estimación de la varianza de la proporción $\widehat{V}(\hat{p})$:

Sustituimos los valores en la fórmula del MAS:

$$\begin{aligned} \widehat{V}(\hat{p}) &= \left(\frac{1222}{1230} \right) \cdot \frac{0,625(1 - 0,625)}{7} \\ &= 0,9935 \cdot \frac{0,234375}{7} \\ &= 0,9935 \cdot 0,033482 \\ &\approx 0,033264 \end{aligned}$$

3. Cálculo del término de linealización:

$$[f'(\hat{p})]^2 = (1 - 2(0,625))^2 = (1 - 1,25)^2 = (-0,25)^2 = 0,0625$$

4. Varianza estimada de $\hat{\theta}$:

$$\hat{V}(\hat{\theta}) = 0,0625 \cdot 0,033264 = \boxed{0,002079} \quad (9)$$

Caso 2: Proporción de individuos que viven con ambos padres**1. Estimación puntual de la proporción (\hat{p}):**

Contando los casos afirmativos ("SI") en la muestra, observamos 3 individuos.

$$n_{SI} = 3 \Rightarrow \hat{p} = \frac{3}{8} = 0,375$$

2. Estimación de la varianza de la proporción $\hat{V}(\hat{p})$:

Observamos que $p(1-p)$ es simétrico respecto a $p = 0,5$. Dado que $0,375 = 1 - 0,625$, el producto $\hat{p}(1-\hat{p})$ es idéntico al Caso 1:

$$0,375(1 - 0,375) = 0,234375$$

Por tanto, la varianza de la proporción se mantiene constante:

$$\hat{V}(\hat{p}) \approx 0,033264$$

3. Cálculo del término de linealización:

$$[f'(\hat{p})]^2 = (1 - 2(0,375))^2 = (1 - 0,75)^2 = (0,25)^2 = 0,0625$$

4. Varianza estimada de $\hat{\theta}$:

Al ser idénticos tanto la varianza base como el término de linealización cuadrático, el resultado final es el mismo:

$$\hat{V}(\hat{\theta}) = 0,0625 \cdot 0,033264 = \boxed{0,002079} \quad (10)$$

3 Apartado c)

Un estimador usual de la media poblacional cuando el tamaño de la población es desconocido es el estimador de Hájek (1971) dado por:

$$t_h = \frac{\sum_s y_i / \pi_i}{\sum_s 1 / \pi_i} \quad (11)$$

siendo π_i la probabilidad de selección de la unidad i .

Determina un estimador de la $V(t_h)$ a partir del método de la aproximación lineal.

3.1 Solución

Para obtener el estimador de la varianza, utilizaremos la técnica de linealización aplicada a estimadores de razón, siguiendo la sugerencia del enunciado.

1. Caracterización del Estimador

El estimador de Hájek se define como:

$$t_h = \frac{\hat{Y}_{HT}}{\hat{N}_{HT}} = \frac{\sum_{k \in s} y_k / \pi_k}{\sum_{k \in s} 1 / \pi_k} \quad (12)$$

Si definimos una variable auxiliar $x_k = 1$ para todo elemento k de la población, entonces el denominador es el estimador de Horvitz-Thompson del total de x (que estima el tamaño poblacional N). Por tanto, t_h es estructuralmente un Estimador de Razón de la forma $\hat{R} = \frac{\hat{Y}_{HT}}{\hat{X}_{HT}}$.

2. Variable Linealizada (Transformación)

La aproximación lineal de Taylor para la varianza de un estimador de razón establece que esta es aproximadamente igual a la varianza del estimador del total de una variable transformada (residuo), escalada por el inverso del cuadrado del tamaño poblacional estimado.

La variable linealizada estimada e_k para un estimador de razón es:

$$e_k = y_k - \hat{R}x_k \quad (13)$$

Sustituyendo nuestros valores específicos:

- El estimador de la razón \hat{R} es t_h .
- La variable auxiliar es $x_k = 1$.

Obtenemos el residuo específico para el estimador de Hájek:

$$e_k = y_k - t_h \cdot 1 = y_k - t_h \quad (14)$$

3. Estimador de la Varianza

El estimador de la varianza de t_h se construye calculando la varianza de Horvitz-Thompson sobre estos residuos e_k y dividiendo por el cuadrado del denominador estimado (\hat{N}_{HT}^2).

La expresión general es:

$$\hat{V}(t_h) = \frac{1}{(\hat{N}_{HT})^2} \hat{V}_{HT}(\hat{E}_{HT}) \quad (15)$$

Desarrollando la fórmula de la varianza de Horvitz-Thompson:

$$\widehat{V}(t_h) = \frac{1}{(\widehat{N}_{HT})^2} \sum_{i \in s} \sum_{j \in s} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \frac{(y_i - t_h)}{\pi_i} \frac{(y_j - t_h)}{\pi_j} \quad (16)$$

Donde $\widehat{N}_{HT} = \sum_{k \in s} \frac{1}{\pi_k}$ es el tamaño estimado de la población.

Esta expresión creo que es válida para cualquier diseño muestral con probabilidades de inclusión de segundo orden π_{ij} conocidas.