

[30] Un agricultor desea saber la productividad de sus cosechas a lo largo de todos sus huertos. Para ello, examina 10 de sus huertos 60, registrando la producción de paltas en gramos de unos cuantos árboles por huerto. La siguiente tabla resume los datos obtenidos, junto al tamaño de árboles por huerto.

Huerto	Tamaño Huerto	N° árboles en Muestra	Producción del Huerto en gr
1	77	6	200, 150, 720, 420, 400, 200
2	45	3	250, 200, 200
3	64	5	320, 470, 125, 375, 120
4	32	2	250, 400
5	43	4	175, 125, 195, 370
6	16	2	300, 315
7	48	4	180, 120, 320, 170
8	67	5	250, 175, 300, 420, 180
9	39	4	370, 480, 200, 150
10	24	2	400, 250

- (a) (10 puntos) Obtenga una estimación puntual para la producción total de paltas en gramos y confeccione un intervalo de 95 % de confianza para esta cantidad. ¿Debe asumir algo?
- (b) (10 puntos) Obtenga una estimación puntual para la producción media de paltas en gramos y obtenga una estimación para su desviación media.
- (c) (10 puntos) Bajo la muestra obtenida. ¿Es posible aseverar que la media poblacional es superior a 300 gr?. Asuma que la producción se distribuye normal. Justifique y comente su respuesta.

**Solución:** Para obtener la estimación insesgada del total utilizamos:

$$\hat{t}_{unb} = \frac{N}{n} \sum_{i \in S} M_i \bar{y}_i$$

En donde  $N = 60$ ,  $n = 10$  y  $M_i$  son los tamaños de los huertos y  $\bar{y}_i$  sus respectivas medias. Así,

$$\hat{t}_{unb} = 755840,5$$

Luego, para obtener su varianza estimada utilizamos:

$$\widehat{\mathbb{V}[\hat{t}_{unb}]} = N^2 \left(1 - \frac{n}{N}\right) \frac{s_t^2}{n} + \frac{N}{n} \sum_{i \in S} \left(1 - \frac{m_i}{M_i}\right) M_i^2 \frac{s_i^2}{m_i}$$

por lo que,

$$\widehat{\mathbb{V}[\hat{t}_{unb}]} = 13061603608 \Rightarrow \sqrt{\widehat{\mathbb{V}[\hat{t}_{unb}]}} = 114287,4$$

Finalmente, el intervalo de confianza pedido, estará dado por:

$$[755840,5 \pm 1,96 * 114287,4]$$

Es claro, que para la resolución se debe asumir que los huertos pueden ser vistos como conglomerados. De igual manera que antes, ajustando las fórmulas respectivas:

$$\hat{\bar{y}}_r = \frac{\sum_{i \in S} M_i \bar{y}_i}{\sum_{i \in S} M_i} = 276,8647$$

y,

$$\widehat{\mathbb{V}[\widehat{y}_r]} = \frac{1}{\overline{M}^2} \left(1 - \frac{n}{N}\right) \frac{s_r^2}{n} + \frac{1}{nN\overline{M}^2} \sum_{i \in \mathcal{S}} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \frac{s_i^2}{m_i} = 363,7939$$

Para responder la interrogantes debemos plantear un test de hipótesis para la media, usando los datos muestrales:

$$H_0 : \mu \leq 300 \quad H_1 : \mu > 300$$

En donde, debido a que no conocemos la varianza poblacional y la poblacional asumimos que proviene desde una distribución normal, nuestro estadístico de prueba bajo  $H_0$  estará dado por:

$$E = \frac{\widehat{y}_r - 300}{\sqrt{\widehat{\mathbb{V}[\widehat{y}_r]}} / \sqrt{n}} \sim t_{n-1}$$

Reemplazando con los datos calculados:

$$E = -7,378179$$

y el cuantil a comparar es:

$$t_{36} = 2,028094$$

Debido a que  $E \not> t_{36}$  existe evidencia estadística suficiente para afirmar con un 95 % de confianza que la productividad media de paltas en los 60 huertos es superior a 300 gramos.