

afectar en la percepción del servicio. Para trabajar la heterogeneidad de las comunas es posible dividir Santiago en macro-zonas de comunas homogéneas. Los argumentos pueden variar de acuerdo a los supuestos. Importante señalar que el muestreo debe ser aleatorio, es decir, que cada cliente tenga la misma probabilidad de ser seleccionado, para que el error asociado sea válido.

4. Diseño Muestral

El Ministerio de Desarrollo Social desea recopilar información sobre el impacto del terremoto en los microempresarios de las distintas comunas afectadas. Sabiendo que el número de microempresarios afectados es muy grande y heterogéneo en las N comunas, desea aplicar alguna técnica de muestreo para obtener esta información. Para ello responda:

- a) En este caso particular, ¿recomendaría utilizar un muestreo aleatorio simple o sugeriría un muestreo más sofisticado? ¿de qué depende esta decisión?

R: Se recomienda utilizar una técnica de muestreo que permita por un lado manejar la heterogeneidad (obtener muestras representativas de microempresarios en cada comuna) y por otro lado reducir los altos costos que implicaría realizar un muestreo aleatorio simple en las zonas afectadas.

- b) Diseñe un método de muestreo que permita cumplir con el objetivo del Ministerio, realizando los supuestos que estime necesario y utilizando toda la información disponible. ¿Qué puede decir del método propuesto cuando es estrictamente necesario contar con representantes de cada comuna?

R: Considerando la información disponible, se recomienda considerar las comunas como unidades muestrales y utilizar un muestreo por etapas para reducir costos. El muestreo por etapas, sin embargo, NO asegura que todas las comunas se encuentren representadas, ya que realiza un muestreo aleatorio de los conglomerados. Se sugiere modificar el método para que el muestreo sea estratificado a nivel comunas (de esta manera se garantiza que todas las comunas se encuentren representadas) y por etapas dentro de las comunas, con el fin de reducir costos.

5. Cálculos Muestrales

En una empresa productora de manzanas la producción se clasifica en cuatro categorías de calidad: 1 (exportación), 2 (Premium nacional), 3 (regular nacional), 4 (baja nacional). Las personas que inspeccionan las manzanas para su clasificación realizan adecuadamente su trabajo, sin embargo, se equivocan una fracción p de las veces. La clasificación la realizan dos personas por

separado y si ambas están en desacuerdo, se llama a un experto que nunca se equivoca. Se asume que la proporción real de manzanas en cada clase es $\frac{1}{4}$. Las manzanas para exportación deben pasar por un proceso de selección antes de su venta, el cual funciona de la siguiente manera: el cliente extranjero toma aleatoriamente 5 manzanas de un cajón seleccionado también aleatoriamente. Si el cliente encuentra 2 o más manzanas que no corresponden a la categoría (bajo los mismos criterios que sigue la empresa), el cliente rechaza el cargamento. El gerente de la empresa productora, preocupado, lo contrata a usted para que lo ayude a medir el riesgo del cargamento.

- Si se asume que la probabilidad de equivocarse en cualquiera de las categorías es la misma, ¿cuál es la probabilidad de rechazo del cargamento?
- Con el fin de reducir el riesgo de rechazo, usted sugiere tomar una muestra de tamaño n de las manzanas clasificadas para exportación. Asumiendo varianza máxima, un error máximo permitido de 0,05 y para un nivel de significancia del 5%, ¿Cuál es el valor de n ? (*)
- Siguiendo su consejo, el gerente junto a su experto realiza un muestreo de n manzanas, con n el tamaño sugerido por usted, de las cuales sólo 1 manzana resultó mal clasificada para exportación. Si el gerente considera que un 1% de manzanas mal clasificadas como exportación es muy riesgoso para el negocio, en cuyo caso habría que venderlas a menor precio junto con las Premium en el mercado nacional, ¿qué le recomienda usted? Fundamente su respuesta con un test de hipótesis ad hoc, con nivel de significancia del 1% y definiendo claramente sus hipótesis (*). Entregue adicionalmente el nivel de significancia más bajo al cual se puede rechazar la hipótesis nula.

(*) Asuma que la población es grande.

Solución:

- R: La probabilidad de rechazo se distribuye binomial con parámetro $p_{err}=P(\text{manzana no es de exportación} \mid \text{manzana clasificada como exportación})$.**

$$P(\text{rechazo}) = \sum_{j=2}^5 \binom{5}{j} (p_{err})^j (1 - p_{err})^{5-j}$$

Para obtener p_{err} necesitamos el Teorema de Bayes. Se definen los sucesos:

- **Ne:** Manzana seleccionada no es de exportación
- **E:** Manzana seleccionada es de exportación
- **Ce:** Manzana clasificada como exportación

$$p_{err} = P(Ne|Ce) = \frac{P(Ce|Ne)P(Ne)}{P(Ce|Ne)P(Ne) + P(Ce|E)P(E)} = \frac{\frac{p^2}{9} \cdot \frac{3}{4}}{\frac{p^2}{9} \cdot \frac{3}{4} + (1 - p^2) \cdot \frac{1}{4}} = \frac{p^2}{3 - 2p^2}$$

b)
$$n = \frac{Z^2 p(1-p)}{e^2} = \frac{1.96^2 0.5(1-0.5)}{0.05^2} = 384 \text{ manzanas}$$

c) R: $H_0: \pi \geq 0.01, H_a: \pi < 0.01, p = \frac{1}{384} = 0.0026$

$$z = \frac{p - \pi}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.0026 - 0.01}{0.0026} = -2.844$$

Como $z^* = -2.33$, se rechaza la hipótesis nula. La proporción poblacional del error de clasificación es estadísticamente menor al 1%. El gerente puede estar tranquilo. Finalmente, el valor p es 0.0023.