**Estimación del tamaño poblacional (Prof. Jimmy Reyes R.)**

## 1.- Muestreo Directo

Se tiene una población de tamaño N desconocido y para estimarlo se elige una muestra aleatoria de *t* elementos, se marcan y se devuelven a la población.

De esta forma $p = \frac{t}{N}$ es la proporción de marcados en la población, luego

**Parámetro**: $N = \frac{t}{p}$

**Estimador**: Se elige una segunda muestra de tamaño $n$ en la cual se encuentran $s$ marcados, luego $\widehat{p} = \frac{s}{n}$ es la proporción estimada de marcados en la segunda muestra, así un estimador puntual del tamaño de la población es $\widehat{N} = \frac{t}{\widehat{p}} = \frac{nt}{s}$

**Varianza estimada**: $V(\widehat{N}) = \frac{t^2 n(n-s)}{s^3}$

Un intervalo de confianza de nivel $(1-\alpha)$ para el tamaño de la población está dado por

$$\left[\widehat{N} - z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})};\ \widehat{N} + z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})}\right]$$

**Ejemplo 1**:

Antes de abrir la temporada de caza se desea estimar la población de venados, para lo cual se captura una muestra de 300 venados, se marcan y se regresan al bosque. Dos semanas después se eligen 200 venados, de los cuales 62 venían marcados. Estime el total de venados en el bosque y determine un intervalo de 95% de confianza para dicho parámetro.

En este caso se tiene $t = 300$ ; $n = 200$ ; $s = 62$

Luego el número estimado de venados en el bosque es $\widehat{N} = \frac{200 * 300}{62} = 967.74 \approx 968$

$$V(\widehat{N}) = \frac{(300)^2(200)(138)}{(62)^3} = 10422.61086$$

$$d = 1.96\sqrt{10422.61086} = 200.098 \approx 200$$

Luego un intervalo de 95% de confianza para el total de venados en el bosque es

$$[968 - 200; 968 + 200] = [768; 1168]$$

## 2.- Muestreo Inverso

Se selecciona una muestra de *t* elementos de la población, se marcan y se regresan a la población. Luego se realiza una selección aleatoria hasta obtener exactamente $s$ elementos marcados. Si la segunda muestra es necesario seleccionar $n$ elementos para obtener los $s$ marcados, entonces $\widehat{p} = \frac{s}{n}$ es una estimación de la proporción de marcados en la población

**(Note que en este caso *s* es fijo y *n* es aleatoria)**

**Estimador**: $\widehat{N} = \frac{t}{\widehat{p}} = \frac{nt}{s}$

**Varianza estimada**: $V(\widehat{N})=\dfrac{t^2n(n-s)}{s^2(s+1)}$

En este caso un intervalo de confianza de nivel $(1-\alpha)$ para el tamaño de la población está dado por

$$\left[\widehat{N}-z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})};\ \widehat{N}+z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})}\right]$$

**Ejemplo 2**:

En una gran reservación de animales es de interés estimar el número total de pájaros de una determinada especie que allí habitan, para lo cual se atrapa una muestra inicial de 150 pájaros, se marcan y luego se dejan libres.
En el mismo mes se capturan pájaros hasta encontrar 35 marcados, logrando capturar 100 pájaros para lograr esto. Estime el número total de pájaros en la reservación y un intervalo de 95% de confianza para dicho total.

En este caso $t = 150$; $n = 100$ y $s = 35$
Luego el número estimado de pájaros en la reservación es
$\widehat{N}=\dfrac{100*150}{35}=428.57\approx 429$

$V(\widehat{N})=\dfrac{(150)^2(100)(65)}{(35)^2 36}=3316.3265$

$d=1.96\sqrt{3316.3265}=112.87\approx 113$

Luego un intervalo de 95% de confianza para el total de pájaros en la reservación es
$[429-113;429+113]=[316;542]$

**Determinación del tamaño de muestra para el caso de muestreo directo y muestreo inverso.**

**1.- Tamaño de muestra para muestreo directo.**

Sea $p_1=\dfrac{t}{N}$ y $p_2=\dfrac{n}{N}$ las fracciones de muestreo de la 1° y 2° captura respectivamente..

Considerando que $\dfrac{V(\widehat{N})}{N}\approx\dfrac{1-p_1}{p_1p_2}$ y un error de estimación dado por $d_0=z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})}$

se obtiene que $V(\widehat{N})=\left(\dfrac{d_0}{z_{1-\frac{\alpha}{2}}}\right)^2$ y multiplicando por $\dfrac{1}{N}$ en ambos lados se obtiene

$$\frac{V(\widehat{N})}{N}=\left(\frac{d_0}{z_{1-\frac{\alpha}{2}}}\right)^2\frac{1}{N}=\frac{1-p_1}{p_1p_2}$$

Si fijamos $p_1$ se tiene que $p_2=\dfrac{N(1-p_1)}{p_1}\left(\dfrac{z_{1-\frac{\alpha}{2}}}{d_0}\right)^2$

Luego el tamaño de muestra para la 1° captura es $t = p_1N$
Y el tamaño de muestra para la 2° captura es $n = p_2N$
Donde $N$ se puede obtener de estudios anteriores similares.

**Ejemplo 3**:

Para estimar el tamaño de la población de venados con un error de estimación no mayor a 220 venados y una confianza de 95%, sabiendo que el año pasado el número de venados estaba entre los 800 y 1000 venados. ¿Cuántos venados se deben escoger en la 2º muestra si en la 1º muestra se desea escoger la cuarta parte de la población?

En este caso $p_1 = 0.25$ y $N = 1000$ luego $p_2 = \dfrac{1000 * 0.75}{0.25}\left(\dfrac{1.96}{220}\right)^2 = 0.238$

Como $p_1 = 0.25 \Rightarrow t = 0.25 * 1000 = 250$ venados en la 1º muestra

$p_2 = 0.238 \Rightarrow n = 0.238 * 1000 = 238$ venados en la 2º muestra

**2.- Tamaño de muestra para muestreo inverso.**

Sea $p_1 = \dfrac{t}{N}$ y $p_2 = \dfrac{s}{N}$ las fracciones de muestreo de la 1º y de la cuota de marcados para la 2º captura respectivamente..

Considerando que $\dfrac{V(\widehat{N})}{N} \approx \dfrac{1-p_1}{p_2}$ y un error de estimación dado por $d_0 = z_{1-\frac{\alpha}{2}}\sqrt{V(\widehat{N})}$

se obtiene que $V(\widehat{N}) = \left(\dfrac{d_0}{z_{1-\frac{\alpha}{2}}}\right)^2$ y multiplicando por $\dfrac{1}{N}$ en ambos lados se obtiene

$$\frac{V(\widehat{N})}{N} = \left(\frac{d_0}{z_{1-\frac{\alpha}{2}}}\right)^2 \frac{1}{N} = \frac{1-p_1}{p_2}$$

Si fijamos $p_1$ se tiene que $p_2 = N(1-p_1)\left(\dfrac{z_{1-\frac{\alpha}{2}}}{d_0}\right)^2$

Luego el tamaño de muestra para la 1º captura es $t = p_1 N$

Y la cuota de marcados para la 2º captura es $s = p_2 N$

Donde $N$ se puede obtener de estudios anteriores similares.

**Ejemplo 4**:

Para estimar la población de pájaros con una confianza de 95% y un error de estimación no superior a 115 pájaros, si se sabe que el número de pájaros era aproximadamente de 500 pájaros. ¿Cuántos pájaros marcados se desean obtener en la segunda muestra si la primera se escogió la cuarta parte de la población?

En este caso $p_1 = 0.25$ y $N = 500$ luego $p_2 = 500 * 0.75\left(\dfrac{1.96}{115}\right)^2 = 0.109$

Como $p_1 = 0.25 \Rightarrow t = 0.25 * 500 = 125$ pájaros en la 1º muestra

$p_2 = 0.109 \Rightarrow s = 0.238 * 500 = 54.5 \approx 55$ pájaros marcados en la 2º muestra