Un guardabosque quiere estimar el número promedio de hectáreas plantadas de árboles en los ranchos de un estado. Ya que el número de hectáreas de árboles varía considerablemente con respecto al tamaño del rancho, decide estratificar con base en el tamaño de los ranchos. Los 240 ranchos en el estado son puestos en una de 4 categorías de acuerdo con el tamaño. Una muestra aleatoria estratificada de 40 ranchos, seleccionados mediante asignación proporcional, produce los resultados del número de hectáreas plantadas de árboles que se muestran en la siguiente tabla:

Estrato 1 0 – 200 hectáreas		Estrato 2 201 – 400 hectáreas		Estrato 3 401 – 600 hectáreas		Estrato 4 Más de 600 hectáreas
$N_1=86$		$N_2=72$		$N_3=52$		$N_4=30$
$n_1=14$		$n_2=12$		$n_3=9$		$n_4=5$
97	67	125	155	142	256	167
42	125	67	96	310	440	220
25	92	256	47	495	510	780
105	86	310	326	320	396	655
27	43	220	352	196	-	540
45	59	142	190	-		-
53	21	-		-		-

1. Realice una estimación para el número promedio de hectáreas plantadas de árboles en los ranchos de los estratos 1 y 2, y realice su respectivo intervalo de confianza del 95 %.

El enunciado pide $\bar{Y}_1$ y $\bar{Y}_2$ y sus IC con $\alpha = 0,05$

$$\bar{Y}_1 = \frac{887}{14} \approx 63,3571 \; ; \; s_1^2 = 1071,786$$

$$\hat{Var}(\bar{Y}_1) = \left(\frac{86-14}{86}\right)\frac{s_1^2}{14} = 64,0935$$

$\therefore$ (46.06157 80.65271)

$$\bar{Y}_2 \approx 190,5 \; ; \; s_2^2 = 10596,45$$

$$\hat{Var}(\bar{Y}_2) = \left(\frac{72-12}{72}\right)\frac{s_2^2}{12} = 735,8699$$

$\therefore$ (130.7943 250.2057)

2. Al realizar una estimación para el número promedio de hectáreas plantadas de árboles en los ranchos del estado, y su respectivo intervalo de confianza del 95 %, se obtiene:
   a. 212.69 y (205.6, 312.8)
   b. 207.4399 y (173.8382, 259.45)
   c. 212.942 y (177.4371, 249.62)
   ~~d.~~ Ninguna de las anteriores

Estrato 1 Estrato 2 Estrato 3 Estrato 4

$\bar{Y}_1 = 63,3571$ $\bar{Y}_2 = 190,5$ $\bar{Y}_3 = 340,5556$ $\bar{Y}_4 = 472,4$

$\hat{var}(\bar{Y}_1) = 69,0935$ $\hat{var}(\bar{Y}_2) = 735,8649$ $\hat{var}(\bar{Y}_3) = 1593,064$ $\hat{var}(\bar{Y}_4) = 12062,72$

Piden IC con $\alpha = 0,05$ para $\bar{Y}_{st}$

$$\bar{Y}_{st} = \left(\frac{86}{240}\right)\bar{Y}_1 + \left(\frac{72}{240}\right)\bar{Y}_2 + \left(\frac{52}{240}\right)\bar{Y}_3 + \left(\frac{30}{240}\right)\bar{Y}_4$$

$$\approx 212,69$$

IC con $\alpha = 0,05$ para $\bar{Y}_{st}$:

$$\bar{Y}_{st} \pm t_{\frac{0,05}{2}, 40-4} \cdot \hat{se}(\bar{Y}_{st})$$

$$\hat{se}(\bar{Y}_{st}) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{4} N_i^2 \hat{var}(\bar{Y}_i)} \approx 18,31327$$

$\therefore$ (175.549 249.831)

3. Encuentre un intervalo de confianza del 90 % para el total de hectáreas plantadas en los ranchos del estado.

IC con $\alpha = 0,1$ para $\hat{\tau}_{st}$

$$N \cdot \bar{Y}_{st} \pm t_{\frac{0,1}{2}, 40-4} \cdot N \cdot \hat{se}(\bar{Y}_{st})$$

(43625.22 58465.98)

Se pretende determinar la prevalencia de una peligrosa enfermedad (Maedi), en una explotación de $N = 250$ ovejas: La probabilidad de que una oveja esté infectada de Maedi está

directamente relacionada con su edad, por lo que decide estratificarse con base en la edad. Se tomó una muestra de $n = 98$ ovejas mediante asignación proporcional. La siguiente tabla resume los resultados.

	Estrato1 Ovejas que tienen hasta dos años	Estrato2 Ovejas que tienen entre tres y cuatro años	Estrato3. Ovejas que tienen entre cinco y seis años	Estrato4. Ovejas que tienen más de seis años
Número de ovejas	110 N_1	70 N_2	45 N_3	25 N_4
Número de ovejas muestreadas	40 n_1	28 n_2	19 n_3	11 n_4
Proporción de ovejas infectadas	15% $\hat{p}_1$	50% $\hat{p}_2$	52% $\hat{p}_3$	63% $\hat{p}_4$

4. Realice una estimación para la proporción de ovejas con Maedi en la explotación, así como su respectivo intervalo de confianza del 95 %. De lo anterior se puede concluir:

a. Con una confianza del 95 % la proporción de ovejas en la explotación con Maedi se encuentra entre 73,2481 y 108,0519.

~~b~~. Con una confianza del 95 %, se puede afirmar que la proporción de ovejas con Maedi en la explotación se encuentra entre 0,293 y 0,432.

c. Con una confianza del 95 % la proporción de ovejas en la explotación con Maedise encuentra entre 73,24 % y 108,05 %.

d. El porcentaje de ovejas en la explotación con Maedi se encuentra entre 0,0293 y 0,432, a un nivel de confianza del 95 %.

Primero, se debe hallar $\hat{p}_{st}$

$$\hat{p}_{st} = \frac{110 \cdot 0{,}15 + 70 \cdot 0{,}5 + 45 \cdot 0{,}52 + 25 \cdot 0{,}63}{250} \approx 0{,}3626$$

$$\hat{var}(\hat{p}_{st}) = \frac{1}{250^2} \cdot \left(\sum_{i=1}^{4} N_i^2 \, \hat{var}(\hat{p}_i) \right)$$

donde $\hat{var}(\hat{p}_i) = \left(\frac{N_i - n_i}{N_i} \right) \frac{\hat{p}_i(1-\hat{p}_i)}{n_i - 1}$

$$\sqrt{\hat{var}(\hat{p}_{st})} \approx 0{,}03509918$$

Ahora, un IC con $\alpha = 0{,}05$ para $\hat{p}_{st}$

$$\hat{p}_{st} \pm t_{\frac{0{,}05}{2}, 98-4} \cdot \sqrt{\hat{var}(\hat{p}_{st})}$$

(0.293009 0.432191)

5. Se desea estimar la altura media (en centímetros) en una población donde se tomaron tres estratos generados por la etnia. El presupuesto disponible solo alcanza para seleccionar una muestra de $n = 210$ unidades. Los tres estratos estan compuestos por 1000, 2000 y 5000 individuos cada uno, además, estudios previos tienen las siguientes estimaciones de la varianza dentro de cada estrato: $S_1^2 = 100$, $S_2^2 = 81$, $S_3^2 = 36$. ¿Cómo se distribuiría la muestra global haciendo uso de?

a) Afijación de Neyman.
b) Afijación proporcional.

a)
$$\psi_i = \frac{N_i\sigma_i}{\sum_{k=1}^{L} N_k\sigma_k}$$

$N_1 = 1000$
$N_2 = 2000$
$N_3 = 5000$

$$\psi_1 = \frac{1000\cdot\sqrt{100}}{1000\cdot\sqrt{100}+2000\cdot\sqrt{81}+5000\cdot\sqrt{36}} = \frac{5}{29} \approx 0,1724$$

$$\psi_2 = \frac{2000\cdot\sqrt{81}}{1000\cdot\sqrt{100}+2000\cdot\sqrt{81}+5000\cdot\sqrt{36}} = \frac{9}{29} \approx 0,3103$$

$$\psi_3 = \frac{5000\cdot\sqrt{36}}{1000\cdot\sqrt{100}+2000\cdot\sqrt{81}+5000\cdot\sqrt{36}} = \frac{15}{29} \approx 0,5172$$

$n_1 = \psi_1 n \approx 36,2 = 36$
$n_2 = \psi_2 n \approx 65,2 = 65$
$n_3 = \psi_3 n \approx 108,6 = 109$

b) $\psi_i = \frac{N_i}{N}$

$$\psi_1 = \frac{1000}{8000} = \frac{1}{8} \approx 0,125$$

$$\psi_2 = \frac{2000}{8000} = \frac{1}{4} \approx 0,25$$

$$\psi_3 = \frac{5000}{8000} = \frac{5}{8} \approx 0,625$$

$n_1 = \psi_1 n \approx 26,25 = 26$
$n_2 = \psi_2 n \approx 52,5 = 52$
$n_3 = \psi_3 n \approx 131,25 = 132$

6. Suponga que se tiene una población dividida en cuatro estratos, con afijaciones $\psi_1 = 0.3$, $\psi_2 = 0.2$, $\psi_3 = 0.25$ y costos unitarios de muestreo por estrato $C_1 = 10\text{USD}$, $C_2 = 25\text{USD}$, $C_3 = 50\text{USD}$ y $C_4 = 5\text{USD}$. Estime el tamaño de muestra sabiendo que el presupuesto total del que se dispone para hacer el muestreo es de 1500USD

$$\psi_4 = 1-(0,2+0,3+0,25) = 0,25$$

Note que:

$$C - C_0 = \sum_{i=1}^{4} c_i n_i = 1500 \quad \wedge \quad n_i = \psi_i n$$

$$\therefore n \sum_{i=1}^{4} c_i \psi_i = 1500$$

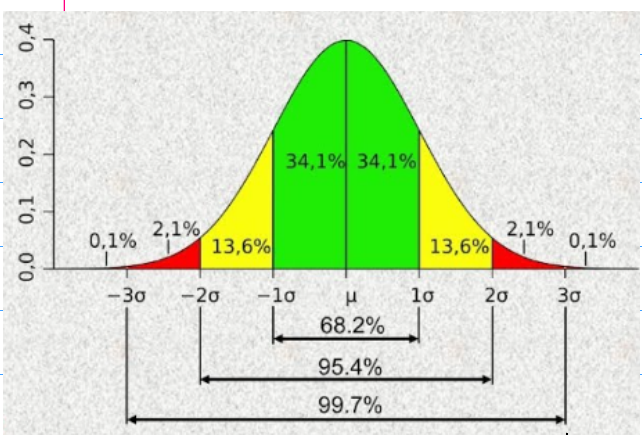
$$\therefore n = \frac{1500}{\sum_{i=1}^{4} c_i \psi_i} \approx 68,9655$$

Como hay restricción presupuestal,

$$\boxed{n = 68}$$

7. El jefe de personal de una empresa desea estimar para un año el número total de días utilizados de permiso de enfermedad entre las 46 plantas de la empresa. Las 46 plantas se dividen en 20 plantas pequeñas y 26 plantas grandes. Por experiencia se sabe que el jefe cree que las plantas pequeñas pueden utilizar entre 0 y 100 días de permiso de enfermedad, mientras que las plantas grandes pueden utilizar entre 10 y 200 días de permiso por enfermedad. Si desea estimar con un error no mayor a 3 días.

a) Encuentre la afijación apropiada de la muestra para los estratos.
b) Determine el tamaño de muestra adecuado.



min

max

$$\text{Max} - \text{min} = (\mu + 3\sigma) - (\mu - 3\sigma) = 6\sigma$$

$$\sigma = \frac{\text{Max} - \text{min}}{6}$$

$$\sigma_1 = \frac{100}{6} = \frac{50}{3} \approx 16,67$$

$$\sigma_2 = \frac{190}{6} = \frac{95}{3} \approx 31,67$$

$$N = 46, \quad N_1 = 20, \quad N_2 = 26$$

a) No se conocen costos pero sí N_i y σ_i, por lo que afijación de Neyman

$$\psi_1 = \frac{20 \cdot \frac{50}{3}}{20 \cdot \frac{50}{3} + 26 \cdot \frac{95}{3}} = \frac{100}{347} \approx 0{,}288$$

$$\psi_2 = \frac{26 \cdot \frac{95}{3}}{20 \cdot \frac{50}{3} + 26 \cdot \frac{95}{3}} = \frac{247}{347} \approx 0{,}712$$

b)

$$n = \frac{\sum_{i=1}^{L} \frac{N_i^2 \sigma_i^2}{\psi_i}}{N^2 D + \sum_{i=1}^{L} N_i \sigma_i^2}$$

$$D = \frac{\delta^2}{Z_{\alpha/2}^2}$$

$$D = \frac{3^2}{Z_{\alpha/2}^2} = \frac{3^2}{(1{,}95996)^2}$$

$$n = \frac{\frac{20^2 \cdot \left(\frac{50}{3}\right)^2}{\frac{100}{347}} + \frac{26^2 \cdot \left(\frac{95}{3}\right)^2}{\frac{247}{347}}}{46^2 \cdot \frac{3^2}{(1{,}95996)^2} + 20 \cdot \left(\frac{50}{3}\right)^2 + 26 \cdot \left(\frac{95}{3}\right)^2} \approx 36{,}5687$$

$n_1 \approx 10{,}5385$
$n_2 \approx 26{,}0301$

No hay restricción presupuestal

$\therefore n_1 = 11$
$n_2 = 27$

$$\boxed{n = 38}$$