<u>Muestreo</u>

1. En una ciudad se quiere hacer una encuesta para conocer el porcentaje de ciudadanos que aprueban la gestión del ayuntamiento en cuestiones medioambientales (limpieza de calles, contaminación, cuidado de parques…). Se pretende que la muestra sea representativa por sexo y edad; para la edad se establecen tres estratos: 10 a 25 años (jóvenes), 25 a 60 años (adultos) y mayores de 60. El número de personas de cada grupo es: jóvenes, 3000; adultos, 8500; mayores de 60, 2500. Por sexo, la distribución es: 6800 hombres y 7200 mujeres, que se suponen proporcionales a cada grupo de edad.
Si el tamaño de la muestra es de 500 personas, determina, redondeando si es necesario, el tamaño muestral correspondiente a cada estrato.
<u>Solución</u>:
El total de ciudadanos que pueden ser encuestados es de 3000 + 8500 + 2500 = 14000. De ellos, el $\frac{6800}{14000}\cdot 100 = 48,571\%$ son hombres. Por tanto, en cada estrato se tienen:

	Jóvenes	Adultos	Mayores
Hombres	1457	4129	1214
Mujeres	1543	4371	1286

Se pretende encuestar a 500 personas de las 14000 existentes; esto supone 1 por cada 28 (14.000/500 = 28), luego dividiendo entre 28 el número de componentes de cada estrato, se obtienen los tamaños muestrales respectivos, que son los dados en la siguiente tabla.

	Jóvenes	Adultos	Mayores	Total
Hombres	52	147 + 1	43	243
Mujeres	55	156	46	257

En hombres adultos se ha añadido 1 con el fin de completar la muestra, pues en los redondeos se pierde un individuo. Se hace en ese estrato por ser el que deja más resto.

2. Supongamos que en un centro escolar los alumnos y docentes se distribuyen de acuerdo con la tabla:

Alumnos	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	110	95	115	20
Mujeres	130	120	130	25

Si se quiere realizar una encuesta entre ellos de tamaño $n = 50$, por el método de muestreo estratificado por sexo y nivel de trabajo, ¿a cuántas persona de cada *clase* hay que preguntar?
<u>Solución</u>:
El total de personas que pueden ser encuestados es de 240 + 215 + 245 + 45 = 745, que es la suma de los cuatro niveles. Como se pretende encuestar a 50 personas de las 745 posibles, y 745 : 50 = 14,9, hay que preguntar a 1 de cada 14,9 personas de cada grupo.
Los valores que se obtienen son:

Muestra	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	7,38	6,38	7,93	1,34
Mujeres	8,72	8,05	8,72	1,68

Hay que redondear atendiendo a los restos.
Puede optarse por la siguiente elección:

Muestra	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	7	6	8	1
Mujeres	9	8	9	2

3. Supongamos que en tu ciudad (o en tu comarca) hay 4 institutos o colegios (IES), cuyo alumnado se distribuye como sigue:

	IES 1	IES 2	IES 3	IES 4
Alumnos	370	460	620	520
Alumnas	340	500	650	540

Si se quiere realizar un muestreo de tamaño 100, estratificado por sexo y centro escolar, ¿cuáles serían los tamaños muestrales correspondientes a cada IES y a cada estrato?

Solución:

El total de alumnos/as por IES se indica en la siguiente tabla:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	370	460	620	520	1970
Alumnas	340	500	650	540	2030
Total	710	960	1270	1060	4000

Como hay que elegir una muestra de tamaño 100 habrá que tomar 1 de cada 40 estudiantes. Los valores que se obtienen son:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	9,25	11,5	15,5	13	49,25
Alumnas	8,5	12,5	16,25	13,5	50,75
Total	17,75	24	31,75	26,5	100

Hay que redondear atendiendo a los restos. En este caso puede optarse por subir los representantes del IES 1 y 3 y bajar los el IES 4; eligiendo a 49 alumnos y 51 alumnas. Por tanto, el tamaño muestral correspondiente a cada estrato es el siguiente:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	9	11	16	13	49
Alumnas	9	13	16	13	51
Total	18	24	32	26	100

4. (Selectividad, Galicia 2016).

a) Se desea tomar una muestra estratificada de las personas mayores de edad de un municipio, cuyos estratos son los siguientes intervalos de edades, en años: de 18 a 30, de 31 a 45, de 45 a 60 y mayores de 60. En el primer intervalo hay 7500 personas, en el segundo 8400, en el tercero 5700 y en el cuarto 3000. Calcula el tamaño de la muestra total y su composición, sabiendo que el muestreo se hace con afijación proporcional y se han elegido 375 personas del primer estrato.

b) Dada la población {2, 4, 6}, construye todas las muestras posibles de tamaño 2, que puedan formar mediante muestreo aleatorio simple, y halla la varianza de las medias muestrales de todas las muestras.

Solución:

a) Como la afijación es proporcional, el peso de cada estrato en la muestra es directamente proporcional a los individuos de la población correspondiente.

Si en el primer estrato, formado por 7500 personal se han elegido $m_1 = 375$;

del segundo estrato: 8400 → se elegirán m_2 $\Rightarrow m_2 = \dfrac{8400 \cdot 375}{7500} = 420$;

del tercer estrato: 5700 → se elegirán m_3 $\Rightarrow m_3 = \dfrac{5700 \cdot 375}{7500} = 285$;

del cuarto estrato: 3000 → se elegirán m_4 $\Rightarrow m_4 = \dfrac{3000 \cdot 375}{7500} = 150$;

El tamaño total de la muestra es 375 + 420 + 285 + 150 = 1230.
Se elige 1 de cada 20 personas en cada estrato.

b) En el muestreo aleatorio simple se mantiene la probabilidad de extracción en cada caso.
Por tanto, las extracciones deben hacerse con reemplazamiento.
El número de muestras de tamaño 2 que pueden obtenerse de la población {2, 4, 6} son 9:
{2, 2}; {2, 4}; {2, 6}; {4, 2}; {4, 4}; {4, 6}; {6, 2}; {6, 4} y {6, 6}
(Son las variaciones con repetición de 3 elementos tomados 2 a 2).

Muestra	Elementos	Media de la muestras: $\bar{x}_i$	$\left(\bar{x}_i - \bar{x}\right)^2$
$M1$	{2, 2}	2	4
$M2$	{2, 4}	3	1
$M3$	{2, 6}	4	0
$M4$	{4, 2}	3	1
$M5$	{4, 4}	4	0
$M6$	{4, 6}	5	1
$M7$	{6, 2}	4	0
$M8$	{6, 4}	5	1
$M9$	{6, 6}	6	4
Sumas		36	12

Recuerda que la media y la varianza valen:

$$\mu = \frac{\sum_{i=1}^{n} x_i}{N}; \quad \sigma^2 = \frac{\sum_{i=1}^{n} \left(x_i - \mu\right)^2}{N}$$

La media y la varianza de la población son:

$$\mu = \frac{2+4+6}{3} = 4;$$

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = \frac{8}{3}.$$

Media de las muestras: $\bar{x} = \dfrac{36}{9} = 4$;

Varianza de las muestras: $\sigma_{\bar{x}}{}^2 = \dfrac{12}{9} = \dfrac{4}{3}$ → Desviación típica: $\sigma_{\bar{x}} = \sqrt{\dfrac{4}{3}} = \dfrac{2}{\sqrt{3}}$.

Puede comprobarse que: $\mu = \bar{x} = 4$; y que $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{\sqrt{8/3}}{\sqrt{2}} = \dfrac{2}{\sqrt{3}}$.