**EVALUACIÓN DEL CAPÍTULO 6 DEL MASTER EN ESTADÍSTICA APLICADA**

**TÉCNICAS DE MUESTREO Y DISEÑO DE ENCUESTAS**

**EJERCICIO PRÁCTICO 1**

Se plantea la realización de un estudio sobre el hábito de fumar en la población mayor de 65 años en una determinada región compuesta por tres distritos, en donde se ha realizado una encuesta por muestreo.

El tamaño muestral se ha establecido en 900 personas, igualmente repartido en cada distrito.

En cada distrito se seleccionan 15 secciones censales, empleando el método de Sampford, siendo obviamente el número de personas mayores de 65 años de cada sección censal la variable tamaño.

Una vez seleccionadas las correspondientes secciones censales, se reparte uniformemente el tamaño muestral entre todas, y en cada una de ellas se obtiene una muestra de varones y otra de mujeres, empleando muestreo aleatorio simple en cada grupo de sexo, y repartiendo el tamaño de forma proporcional.

1. Describir la estructura del diseño muestral utilizado, especificando cada una de las fases y describiendo las variables de diseño utilizadas, es decir, estratos y conglomerados.
2. Hallar las probabilidades de inclusión de cada elemento de la población.
3. Si definimos *πij* como la probabilidad de inclusión de segundo orden derivada de la aplicación del método de Sampford, hallar las probabilidades de inclusión para:

1. Dos individuos pertenecientes a distritos diferentes.

2. Dos individuos pertenecientes al mismo distrito pero a diferentes secciones censales.

1. Estimar la proporción de mayores de 65 años que tienen el hábito de fumar, suponiendo que la población objeto de estudio en las provincias es de 25.000, 30.000 y 20.000 respectivamente, y que en las muestras obtenidas hemos observado 80, 44 y 62 mayores que fuman.
2. Supongamos que dentro de cada distrito tenemos un hábitat rural y otro urbano, y que existen diferencias significativas en el hábito de fumar dependiendo de dicho entorno. ¿Podría mejorarse el diseño muestral utilizado en este caso? Si así fuera, explicar cómo.

**EJERCICIO PRÁCTICO 2**

Se ha realizado un estudio por muestreo sobre el uso de Internet en 3 provincias de una determinada región. La población se ha estratificado atendiendo a la provincia y al hábitat (rural / urbano).

La muestra se ha seleccionado en dos etapas, siendo las unidades de primera etapa las secciones censales de cada provincia y las de segunda etapa la población residente de 16 y más años.

La selección de secciones se ha efectuado de forma aleatoria y dentro de cada sección censal, se ha obtenido una muestra de forma sistemática de 20 personas, obteniendo por tanto un total de 900 encuestas.

Los datos correspondientes a la población y a la muestra aparecen en la hoja de cálculo “Ejemplo Uso de Internet.xls” incluida en la carpeta Datos del DVD.

Se pide:

1. Obtener las estimaciones del uso de Internet por provincias, hábitat y sexo, atendiendo al diseño muestral establecido.
2. Como puede observarse, la variable sexo no ha sido tenida en cuenta en el diseño de la muestra. Si quisiéramos que las estimaciones reproduzcan adecuadamente la población, tanto por provincias, como por hábitat y sexo, ¿qué procedimiento/s podríamos utilizar?