
CUESTIÓN 1

Se quiere estimar la superficie total de olivar de una provincia en el 2023. Se dispone como marco de muestreo el listado de explotaciones agrarias del Censo Agrario 2020. Se forman tres estratos de tamaño y en cada uno de ellos se obtiene una muestra aleatoria simple sin reemplazamiento. Se dispone de la siguiente información:

Estratos	Número de explotaciones		Superficie media muestral (Ha.)	Desviación típica muestral (Ha)
	Población (N_h)	Muestra (n_h)		
1	10200	100	2,5	1
2	5000	1200	10	4,5
3	500	250	50	15

Se pide:

- Para cada uno de los estratos:

(a) Estime de manera insesgada el total de hectáreas (Ha) de Olivar junto a su error de muestreo relativo.

(b) Calcule los tamaños muestrales necesarios para realizar las estimaciones anteriores con unos errores de muestreo estimados inferiores a 1000 Ha en el estrato 1, a 500 Ha en el estrato 2, y a 10 Ha en el estrato 3.

- Para el total provincial:

(c) Estime de manera insesgada el total de hectáreas (Ha) de Olivar junto a su error de muestreo relativo.

(d) Calcule los tamaños muestrales por estrato, necesarios para realizar la estimación anterior con un error de muestreo estimado inferior a 500 Ha para el total provincial.

CUESTIÓN 2

Considere tres variables aleatorias A , B y C tales que su distribución conjunta tiene la forma $p(a, b, c) = p(a) \times p(c) \times p(b|a, c)$. A y C tienen distribuciones de Bernoulli de parámetros θ_A y θ_C , respectivamente. Dadas a y c , B tiene distribución de Bernoulli.

(a) Dadas n observaciones independientes idénticamente distribuidas según A , derive el estimador de máxima verosimilitud de θ_A .

(b) Dadas las 5 observaciones de (A, C, B) siguientes $((0,1,1), (0,0,0), (1,0,1), (0,0,0), (0,1,0))$, calcule el estimador de máxima verosimilitud de θ_A , el de θ_C y, finalmente, el del parámetro de Bernoulli de la distribución de B , dados los antecesores a y c , siempre que sea posible. Comente el resultado.

Suponga que se dispone de información a priori para cada uno de los seis parámetros antes mencionados, modelizada en los seis casos mediante distribuciones uniformes $U(0, 1)$ (e independientes):

(c) Dadas n observaciones independientes idénticamente distribuidas según A , derive la media de la distribución a posteriori de θ_A .

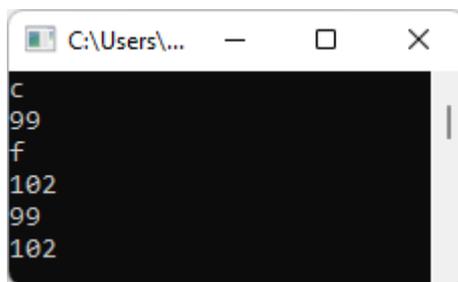
(d) Dadas las mismas 5 observaciones de (A, C, B) , calcule las medias de las distribuciones a posteriori de θ_A , de θ_C y, finalmente, del parámetro de Bernoulli de la distribución de B , dados los antecesores a y c . Comente el resultado.

CUESTIÓN 3

Observe este fragmento de programa en C/C++. Consiste simplemente en una secuencia de instrucciones. Hemos añadido un breve comentario a su derecha explicando su funcionamiento:

```
char car = 'c';           // Sea la variable car, de tipo carácter, con el valor inicial 'c'
printf(" %c \n", car);   // Escribir el valor de car como un carácter
printf(" %d \n", car);   // Escribir el valor de car como un entero
car = car + 3;           // Aumentar el valor de car en 3 unidades
printf(" %c \n", car);   // Escribir el valor de car como un carácter
printf(" %d \n", car);   // Escribir el valor de car como un entero
int n = int('c');        // Sea la variable entera n, que toma el valor inicial del
                          // código ASCII de 'c'
printf(" %d \n", n);     // Escribir el valor de n como un entero
n = n + 3;               // Aumentar el valor de n en 3 unidades
printf(" %d \n", n);     // Escribir el valor de n como un entero
```

Cada instrucción printf contiene el símbolo '\n', que indica un salto de fin de línea, de manera que la ejecución de este fragmento de programa genera la salida siguiente, una línea por cada instrucción printf:



Ahora, vamos con la representación interna de la información que registra caracteres y enteros. Sabemos que la variable 'car' se está almacenando en una palabra (1 byte) y que el código ASCII del carácter 'c' es 99.

(a) Dé la representación, en binario y en hexadecimal, del entero 99.

(b) Dé la representación, en binario y en hexadecimal, del carácter 'c'.

(c) ¿Por qué la instrucción siguiente no genera un error?

`car = car + 3;`

(d) En una máquina cuyas palabras de memoria son de 1 byte, tenemos una palabra con el siguiente valor almacenado: "01100001". ¿Qué puede representar?

CUESTIÓN 4

Dada la población:

u_i	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
x_i	2	7	2	3	5	6	4	1

se desea obtener una muestra sistemática de tamaño 3 para lo cual se determina un periodo $k = N/n$, redondeado al entero más próximo y un arranque aleatorio entre 1 y K .

Se pide:

- (a) Determine el espacio muestral y la función de probabilidad.
- (b) Calcule la probabilidad π_i que tiene la unidad u_i de pertenecer a la muestra.
- (c) Calcule la $V(\hat{X})$ siendo \hat{X} el estimador insesgado del total.
- (d) A partir de la distribución en el muestreo de \hat{X} , calcule $E(\hat{X})$ y $V(\hat{X})$.

CUESTIÓN 5

Se tienen datos $\{x_1, x_2, \dots, x_n\}$ independientes idénticamente distribuidos de una normal de media μ y varianza conocida e igual a 9. Se determina que, a priori, μ se distribuye según una distribución normal de media 10 y varianza 4.

(a) Calcule la verosimilitud del modelo.

(b) Calcule el estimador de máxima verosimilitud de μ .

(c) Calcule la distribución a posteriori de μ .

(d) Calcule la moda a posteriori de μ .

(e) Proporcione la expresión de un intervalo de probabilidad 0.8 a posteriori para μ , suponiendo que se tiene acceso a una rutina $qnorm(p)$ que, dada una probabilidad p , nos da el correspondiente cuantil de la distribución normal estándar.

CUESTIÓN 6

En una localidad turística hay 12 empresas de alojamiento dedicadas a la hostelería, de las que se conoce su cifra de negocio del último mes, expresada en u.m.

<i>Empresa</i>	<i>Cifra de negocios</i>	<i>Grupos de actividad CNAE</i>
A	25.000	551 Hoteles y alojamientos similares
B	30.000	552 Alojamientos turísticos y otros alojamientos de corta estancia
C	15.000	553 Campings y aparcamientos para caravanas
D	12.000	559 Otros alojamientos
E	18.000	551 Hoteles y alojamientos similares
F	13.000	551 Hoteles y alojamientos similares
G	3.000	552 Alojamientos turísticos y otros alojamientos de corta estancia
H	2.000	552 Alojamientos turísticos y otros alojamientos de corta estancia
I	11.000	559 Otros alojamientos
J	2.000	552 Alojamientos turísticos y otros alojamientos de corta estancia
K	2.000	552 Alojamientos turísticos y otros alojamientos de corta estancia

Se pretende publicar una tabla que refleje la cifra de negocio de dichas empresas atendiendo al grupo CNAE en el que se engloba su actividad y respetando la confidencialidad de los informantes. La tabla propuesta es la siguiente:

<i>Grupos de actividad CNAE</i>	<i>Cifra de negocios</i>
551 Hoteles y alojamientos similares	56.000
552 Alojamientos turísticos y otros alojamientos de corta estancia	39.000
553 Campings y aparcamientos para caravanas	15.000
559 Otros alojamientos	23.000

Justifique razonadamente si la tabla propuesta resulta segura en un contexto de control de secreto estadístico atendiendo a los siguientes criterios:

(a) Según la regla de la frecuencia (o valor límite), para $k=2$.

(b) Según la regla de dominancia, para $n=1$, $k=75\%$.

(c) Según la regla del porcentaje, para $p=10\%$.

(d) Proponga alguna alternativa concreta para solucionar los problemas de identificación encontrados.