Guía 12

12.1 Una urna contiene 6 bolas. La distribución *a priori* de la cantidad b de bolas blancas es equiprobable sobre $\{1, 2, 3, 4, 5, 6\}$. Se extraen dos bolas de la urna: una es blanca, la otra es negra. Basándose en esa información muestral

(**a**) hallar la distribución *a posteriori* de la cantidad de bolas blancas que había inicialmente en la urna;

(**b**) calcular el estimador máximo *a posteriori* de b;

(**c**) calcular el estimador bayesiano de b.

12.2 La longitud (en cm.) de ciertas varillas es una variable aleatoria con distribución normal de media μ y desvío estándar 2, donde μ es una variable aleatoria discreta cuya función de probabilidad *a priori* es $\mathbf{P}(\mu = 10) = 0.25$ y $\mathbf{P}(\mu = 14) = 0.75$. Se observa que la longitud de una varilla es 12.1 cm. En virtud de la información muestral, estimar la probabilidad de que la longitud de otra varilla del mismo tipo sea mayor que 13 cm.

12.3 Se recibe un lote de termos provenientes de China. *A priori*, la proporción p de termos defectuosos se distribuye de acuerdo con la función de probabilidad $\mathbf{P}(p = 0.1) = \mathbf{P}(p = 0.9) = 0.5$. Se extraen termos del lote hasta encontrar el primero defectuoso, que se obtiene en la tercera extracción. En base a esa información muestral estimar la media de la cantidad de termos defectuosos que se encontrarán en otros 100000 termos del mismo lote.

12.4 Se arrojará sucesivas veces una moneda cuya probabilidad de salir cara es p. *A priori*, p es una variable aleatoria con distribución $\text{Beta}(\nu_1, \nu_2)$; y en los primeros n tiros se observaron x caras.

(**a**) Hallar la media *a posteriori* de p y mostrar que su valor se encuentra comprendido entre la media *a priori* de p y la frecuencia relativa con que se observa cara.

(**b**) Mostrar que la varianza *a posteriori* de p se comporta asintóticamente de la misma forma que $\frac{(x/n)(1-(x/n))}{n}$.

(**c**) Estimar la probabilidad *a posteriori* de que en el siguiente tiro de la moneda salga cara.

(**d**) Si la moneda "parece equilibrada" y como resultado de los primeros n tiros se observaron $n - 1$ caras y 1 ceca, ¿cómo debe ser n para que en el siguiente tiro las probabilidades sean 2 a 1 a favor de cara?

12.5 El Gobierno de la Ciudad Autónoma de Buenos Aires contrata a una empresa encuestadora para "determinar" la proporción, p, de la población de la Ciudad irritada por las polémicas declaraciones de uno de sus flamantes ministros. Para analizar los datos la encuestadora cuenta con dos especialistas: el primero opina que la densidad *a priori* para p es $f_1(p) = 10(1-p)^9 \mathbf{1}\{0 < p < 1\}$, pero el segundo opina que sería más apropiado usar una *a priori* de la forma $f_2(p) = 10p^9 \mathbf{1}\{0 < p < 1\}$.

(**a**) Analizar qué significado tienen esas dos opiniones y qué consecuencias tienen sobre el análisis de los datos si se decide que para estimar p se utilizará la media de la distribución *a posteriori* basada en el resultado de una encuesta realizada a 10 vecinos de la Ciudad.

(**b**) ¿Qué cantidad de vecinos de la Ciudad deben encuestarse si se quiere que las medias de las dos distribuciones *a posteriori* de p difieran en menos de 0.001?

12.6 Un canal de comunicación binario emite un 1 con probabilidad p. *A priori*, p es una variable aleatoria con distribución Beta$(3,3)$. El receptor indica que hay un 1 cuando efectivamente se ha enviado un 1 con probabilidad $9/10$ e indica que hay un 0 cuando efectivamente se ha enviado un 0 con probabilidad 1. En un mensaje de 5 dígitos se recibieron exactamente 4 unos. En base a esa información muestral calcular la probabilidad de que en un nuevo mensaje de 2 dígitos se reciban exactamente 2 unos.

12.7 En 1898 el matemático ruso Ladislaus Bortkiewicz observó que la cantidad de muertes de soldados por año causadas por patadas de caballos o mulas en los cuerpos de caballería del ejército prusiano obedecía a una distribución de Poisson(μ). Bortkiewicz registró la cantidad de muertos por año en cada uno de 10 cuerpos de caballería durante 20 años, obteniendo así 200 registros. Publicó la siguiente tabla que describe los registros:

Muertes	0	1	2	3	4
Frecuencia	109	65	22	3	1

Utilizando como distribución *a priori* para μ una Gamma de media $1/2$ y varianza $1/8$, calcular la media de la distribución *a posteriori* de μ basada en los datos publicados por Bortkiewicz.

12.8 La cantidad de accidentes semanales en una planta industrial tiene una distribución de Poisson de media μ. En una muestra de 100 semanas se observaron las frecuencias:

Cantidad de accidentes	0	1	2	3	4	5
Frecuencia	10	29	25	17	13	6

A priori, μ tiene una distribución exponencial de media 2. En virtud de la información muestral:

(**a**) estimar la probabilidad de que en la semana del 18 de diciembre de 2017 no ocurra ningún accidente en la mencionada planta;

(**b**) hallar un intervalo de confianza de nivel 0.95 para estimar μ.

12.9 El tiempo de espera (medido en horas) hasta que se produce la segunda falla en cierto tipo de sistemas es una variable aleatoria con distribución Gamma$(2,\lambda)$. *A priori* se supone que λ tiene una distribución exponencial de media 1. Se observó que en uno de tales sistemas el tiempo hasta la segunda falla fue $3/4$ de hora. En virtud de la información muestral hallar la *distribución a posteriori* de λ, calcular su moda y su media.

12.10 El tiempo de realización (en minutos) de una determinada tarea en un proceso industrial es una variable aleatoria con distribución uniforme sobre el intervalo $[1, 1+\theta]$. *A priori*, θ se distribuye de acuerdo con la densidad

$$f(\theta) = \frac{192}{\theta^4}\mathbf{1}\{\theta \geqslant 4\}.$$

Se observan los siguientes tiempos de realización de la tarea: 3, 5 y 8. En base a esa información muestral hallar la distribución a posteriori de θ, y estimar la media del tiempo de realización de la tarea.

12.11 El tamaño, X (en GB), de ciertos archivos obedece a una distribución uniforme sobre el intervalo $(0, \theta]$. *A priori*, θ se distribuye de acuerdo con la densidad

$$f(\theta) = \frac{3}{2}\,\theta^{-5/2}\mathbf{1}\{\theta > 1\}.$$

Se observa que los tamaños de 10 archivos son:

$$0.93,\ 1.55,\ 2.50,\ 1.12,\ 1.11,\ 3.00,\ 1.99,\ 0.20,\ 2.61,\ 0.73.$$

En base a esta información muestral estimar la probabilidad de que un archivo del mismo tipo tenga un tamaño superior a 2 GB.

12.12 El tamaño, X (en GB), de ciertos archivos es una variable aleatoria cuya función de densidad es

$$f_{X|T=\theta}(x) = \theta\, x^{-2}\mathbf{1}\{x \geqslant \theta\}.$$

A priori, T tiene distribución uniforme sobre el intervalo $(1, 2)$. Se observa que los tamaños de dos archivos son: 1.75, 2.35. En virtud de la información muestral estimar la probabilidad de que un archivo del mismo tipo tenga un tamaño superior a 2 GB.

12.13 La longitud (en cm.) de las varillas de un lote es una variable aleatoria con distribución normal de media μ y desvío estándar 2, donde μ es una variable aleatoria con distribución *a priori* normal de media 13 y desvío estándar 1. Se observó una muestra aleatoria de n varillas pertenecientes al lote, y el el promedio de sus longitudes resultó ser 12.1 cm. En base a esa información

(**a**) hallar la distribución *a posteriori* de μ;

(**b**) mostrar que la media *a posteriori* puede expresarse como un promedio ponderado de la forma $\gamma_n 13 + (1-\gamma_n)12.1$, y analizar el comportamiento de γ_n cuando $n \to \infty$.

(**c**) Se muestrea al azar una nueva varilla de lote y mide X cm. Hallar la distribución predictiva de X.

(**d**) Para $n = 10$, hallar un intervalo de confianza de nivel 0.95 para μ y un intervalo predictivo de nivel 0.95 para X.

(**e**) Repetir el inciso anterior para $n = 100$.