Guía ejercicios 3

1. Un ratoncito está al comienzo de un laberinto, y debe escoger una de tres puertas. La primera puerta conduce al queso de premio y se tarda 3 minutos en recorrer. La segunda puerta lo lleva de vuelta al comienzo del laberinto, volviendo por la tercera puerta, y se tarda 5 minutos en recorrer. Suponga que cada vez que el ratoncito está al comienzo, escoge una puerta al azar e independiente de las elecciones pasadas. ¿Cuál es el tiempo esperado que tarda el ratoncito en encontrar su queso?

2. Sea $U_1, U_2, \ldots$ una secuencia de variables unif$(0,1)$ independientes. Dado $x > 0$ fijo, definimos la variable aleatoria $N(x) = \text{mín}\{n : \sum_{i=1}^{n} U_i > x\}$, es decir, la mínima cantidad de variables U_i que debemos sumar para exceder x. Calcule $m(x) = \mathbb{E}(N(x))$. *Indicación:* condicionando en U_1, encuentre una ecuación para $m(x)$.

3. Se lanza n veces de manera independiente una moneda con probabilidad p de cara, donde p es el resultado de la realización de otra variable aleatoria U con distribución unif$(0,1)$, independiente de los lanzamientos. Sea X la cantidad de caras que se obtienen. Demuestre que para todo $i = 0, \ldots, n$ se tiene que $p_X(i) = \frac{1}{n+1}$. *Indicación:* utilizando una propiedad conocida, calcule $\mathbb{P}(X = i)$ condicionando en los posibles resultados de U; utilice sin demostrar el hecho que

$$\int_0^1 p^i (1-p)^{n-i} dp = \frac{i!(n-i)!}{(n+1)!}.$$

4. En una fiesta de año nuevo a la que usted asiste, se lanza al aire una gran cantidad de diminutos papeles blancos y rojos. Los papeles blancos provienen de un contenedor esférico, y los papeles rojos de una caja cúbica, de manera que el diámetro del contenedor esférico es igual a la arista de la caja (ambos poseen igual cantidad de papeles por cm^3). Al llegar a su casa usted se percata que adheridos a su ropa hay b papeles blancos y r papeles rojos. Argumente por qué $6b/r$ es una buena aproximación de π; explicite sus supuestos.

5. Sea $X_1, X_2, \ldots$ una secuencia de variables i.i.d. y positivas, tal que $\lambda = \mathbb{E}\log(X_1)$ es finita. Pruebe que la *media geométrica* de las primeras n variables, es decir $(\prod_{i=1}^{n} X_i)^{1/n}$, converge casi seguramente cuando $n \to \infty$, y calcule su límite.

6. Se sabe que el valor esperado del puntaje que obtiene un alumno en el examen final de un ramo es de 75.

a) Dé una cota superior de la probabilidad que el puntaje sea mayor que 85.

b) Suponga de aquí en adelante que se sabe que la varianza es 25. ¿Qué puede decirse sobre la probabilidad de que el puntaje obtenido por el alumno esté entre 65 y 85?

c) ¿Cuántos alumnos tienen que dar el examen para asegurar que, con probabilidad de al menos un 99 %, el promedio de notas esté entre 70 y 80? Obtenga un resultado sin utilizar el teorema central del límite, y otro utilizándolo.

7. Cincuenta números se aproximan al entero más cercano y después se suman. Los errores de aproximación de cada uno se modelan como variables independientes y uniformes en el intervalo $(-1/2, 1/2)$.

a) Encuentre una cota para la probabilidad de que la suma resultante difiera de la suma exacta por 3 o más.

b) Usando el TCL, aproxime la probabilidad anterior.

8. Se dispone de 40 baterías, cada una de las cuales puede ser de tipo A o B con igual probabilidad e independiente del resto. La duración media de una batería tipo A es de 50, con raíz de la varianza de 15, mientras que para el tipo B la media es 30 y la raíz de la varianza es 6.

a) Aproxime la probabilidad de que la duración total de las 40 baterías exceda 1700.

b) Suponiendo que se sabe que hay 20 baterías de cada tipo, rehaga la parte anterior.

9. Un panel solar está conformado por n celdas fotovoltaicas estándar. La energía total que almacena el panel durante 1 hora es $T = \sum_{i=1}^{n} X_i$, donde $X_1, \ldots, X_n$ representan las energías capturadas por cada celda, las cuales modelamos como variables i.i.d. positivas con media $\mu = 8$ [Wh]. Se desea obtener la cantidad mínima de celdas n^* tal que el panel almacene $E = 112$ [Wh] o más, con probabilidad de al menos $1 - \alpha$, con $\alpha = 2{,}28\,\%$.

a) Sin aproximar con el TCL, encuentre una cota para la probabilidad de que el panel almacene la energía deseada. Obtenga una condición *necesaria* para el valor de n^*.

b) Suponga de aquí en adelante que $\sigma^2 = 4$ es la varianza de cada X_i. Utilizando el TCL, encuentre un valor aproximado para n^*.

c) Sale al mercado un nuevo tipo de celda que en 1 hora captura $\nu = 14$ [Wh] en esperanza, con varianza $\tau^2 = 16$. Usted propone diseñar un nuevo panel compuesto por m de estas celdas, cuyas energías producidas en 1 hora modelamos como variables i.i.d. $Y_1, \ldots, Y_m$, independientes de los X_i. Utilizando el TCL, calcule el mínimo m^* tal que, con probabilidad $1 - \alpha$, la energía total capturada $S = \sum_{i=1}^{m^*} Y_i$ sea al menos un 50 % mayor que la producida por un panel con la misma cantidad de celdas estándar.

10. Sea X variable aleatoria binomial con parámetros n y p.

a) Sean $\hat{p}_1 = X/n$ y $\hat{p}_2 = (X+1)/(n+2)$ estimadores de p. Calcule ECM$(\hat{p}_1)$ y ECM$(\hat{p}_2)$. ¿Para qué valores de p es mejor $\hat{p}_2$ de acuerdo al criterio del error cuadrático medio?

b) Sea $\hat{\sigma}^2 = X(1 - X/n)$ un estimador de $\sigma^2 = \text{var}(X)$. Muestre que $\hat{\sigma}^2$ es sesgado y modifíquelo para obtener un estimador insesgado de σ^2.

11. Sea $X_1, \ldots, X_n$ una m.a.s. con distribución común Poisson(λ). Encuentre el estimador de máxima verosimilitud de λ, calcule su esperanza y varianza, y muestre que este estimador es consistente.

12. Sea $X_1, \ldots, X_n$ una m.a.s. proveniente de una

Gamma(θ, λ), donde $\theta > 1$ es conocido.

a) Muestre que los estimadores de λ del método de los momentos y el de máxima verosimilitud coinciden con $\hat{\lambda} = \theta/\overline{X}$.

b) Concluya que $\hat{\lambda}$ converge casi seguramente a λ cuando el tamaño de la muestra crece indefinidamente.

c) Muestre que la esperanza de $\hat{\lambda}$ es $\lambda n\theta/(n\theta - 1)$ y modifíquelo para obtener un estimador insesgado $\tilde{\lambda}$. Suponiendo $\theta > 2$, calcule la varianza de $\tilde{\lambda}$ y muestre que es un estimador consistente. *Indicación:* utilice el hecho que $\sum_{i=1}^{n} X_i$ tiene distribución Gamma$(n\theta, \lambda)$.

13. Sea $X_1, \ldots, X_n$ una m.a.s. con densidad común dada por $f(x;\theta) = \alpha x^{\alpha-1}\theta^{-\alpha}\mathbb{1}_{[0,\theta]}(x)$, donde $\alpha > 0$ es conocido y θ es desconocido. Obtenga estimadores para θ con los métodos de máxima verosimilitud y de los momentos. Verifique si dichos estimadores son insesgados; en caso que alguno no lo sea, modifíquelo para obtener un estimador insesgado.

14. Se toma una muestra de 24 mediciones independientes para la temperatura de fusión del plomo, obteniendo un promedio de 336 y una varianza muestral insesgada de 441. Encuentre intervalos de confianza para la media μ y la varianza real σ^2 al nivel 95 %.

15. En un laboratorio usted genera 7 cultivos de una determinada bacteria, cuyas biomasas resultaron ser

5,4 8,0 2,0 6,0 5,6 7,0 1,0.

Asumiendo que los datos conforman una m.a.s. proveniente de una normal, obtenga un intervalo de confianza para la varianza σ^2 al nivel 90 %.

16. En base a una m.a.s., se desea encontrar un intervalo de confianza al nivel $1-\alpha$ para la media μ de una variable $\mathcal{N}(\mu, \sigma^2)$, con σ^2 conocido. Si se desea que el largo del intervalo sea a lo más L, ¿cuál es el tamaño mínimo de la muestra?

17. Se observó que 10 de un total de 100 veces, el tiempo de espera del bus fue mayor a 10 minutos. Entregue un intervalo de confianza al nivel 90 % para la probabilidad p de que un bus tarde más de 10 minutos.

18. El tiempo de vida de un cierto microorganismo se modela como una variable normal con parámetros μ y σ^2 desconocidos. En un laboratorio, usted toma una muestra aleatoria simple con $n = 16$ realizaciones de esta variable, y obtiene un promedio de 20,0 y un estimador insesgado de la varianza de 8,1.

a) Obtenga un intervalo de confianza para μ al nivel 95 %.

b) Obtenga un intervalo de confianza para σ^2 al nivel 95 %.

c) Una compañera de laboratorio toma una muestra de tamaño $m = 20$, obteniendo el mismo promedio de 20,0 y un estimador insesgado de la varianza de 10,0. Si esta muestra se *consolida* con la suya (es decir, se trabaja con el total de $n+m = 36$ datos), ¿cuál es el nuevo intervalo de confianza para μ al mismo nivel 95 %?

d) Un reciente estudio realizado por otro prestigioso laboratorio muestra que para todo efecto práctico, la varianza real del tiempo de vida del microorganismo puede asumirse igual a $\sigma^2 = 9{,}0$. Usando la muestra consolidada (con $n + m = 36$ datos), ¿cuál es ahora el intervalo de confianza para μ al mismo nivel 95 %?

19. Como parte de un informe para un ramo de su universidad, usted realiza un pequeño estudio evaluando la calidad de un cierto producto, para lo cual usted toma una muestra de tamaño 13 y obtiene una desviación estándar (es decir, la raíz del estimador insesgado de la varianza) de 12,0.

a) Una de las características más importantes de la variable en estudio es su variabilidad σ^2 (es decir, la varianza). Encuentre un intervalo de confianza para σ^2 al nivel 90 %.

b) Un compañero de clase realizó el mismo estudio con el mismo tamaño de muestra y el mismo nivel de confianza, obteniendo un intervalo cuyo largo es 43,75 % menor que el obtenido por usted. ¿Cuál es el valor de la desviación estándar obtenida por su compañero?

c) Al comparar las muestras, con su compañero se dan cuenta que ambos obtuvieron el mismo promedio. Para mejorar la calidad de los resultados, ustedes deciden trabajar con el total de 26 datos obtenidos en conjunto por ambos. Manteniendo el nivel de confianza, ¿cuál es ahora el intervalo para σ^2?

20. Sea X variable aleatoria con densidad $f_X(x) = \frac{x}{\theta}e^{-x^2/(2\theta)}\mathbb{1}_{[0,\infty)}(x)$, donde $\theta > 0$ es un parámetro desconocido. Sea $X_1, \ldots, X_n$ una m.a.s. con densidad común f_X.

a) Obtenga $\hat{\theta}$, el estimador de máxima verosimilitud de θ.

b) Muestre que $X^2/2 \sim \exp(1/\theta)$.

c) Mustre que $\hat{\theta}$ converge casi seguramente a θ cuando $n \to \infty$.

d) Muestre que el estadístico $Z = (\hat{\theta} - \theta)/(\theta/\sqrt{n})$ posee distribución aproximadamente normal estándar.

e) Utilice el estadístico Z para encontrar un intervalo de confianza aproximado genérico para θ al nivel $1 - \alpha$. Para $n = 49$, $\alpha = 5\,\%$ y suponiendo que el valor observado de $\hat{\theta}$ es de 20,0, obtenga el intervalo explícitamente.

21. El gerente comercial de una marca de detergentes afirma que el 40 % del público prefiere dicha marca. Usted trabaja para la competencia, y cree que dicha información es exagerada. Para verificar quién tiene razón usted encarga realizar un estudio de mercado, el cual indica que 54 de 150 personas encuestadas prefiere la marca. Plantee hipótesis H_0 y H_1 adecuadas e indique su conclusión para $\alpha = 5\,\%$.

22. Un productor afirma que al menos el 20 % del público prefiere su producto. Se toma una muestra de 100 personas para verificar su afirmación. Con $\alpha = 0{,}05$, ¿cuál es la mínima cantidad de personas que prefieren el producto de manera que no haya suficiente evidencia para rechazar la afirmación del productor? Si solo 10 personas prefieren el producto, ¿cuál es el resultado del test?