Probabilidad

Lista de Problemas 5

Los cuatro primeros problemas son para entregar el martes 12/9/17.

Por favor, entrega por separado los problemas pares y los impares.

1. Demuestre la siguiente generalización de la desigualdad de Chebyshev. Sea $\phi : (0, \infty) \to (0, \infty)$ creciente y suponga que $\mathrm{E}[\phi(|X|)] = M < \infty$. Demuestre que $P(|X| \geq c) \leq M/\phi(c)$.

2. Sea $X_1 \sim Pois(\lambda_1)$, $X_2 \sim Pois(\lambda_2)$ y estas variables son independientes. Demuestre que $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$. Extienda el resultado anterior al caso de n variables.

3. Sean $X_1, \dots, X_n$ v.a.i. con distribución de Bernoulli de parámetro p y sea $S_n = X_1 + \cdots + X_n$. Demuestre que $P(S_m = j | S_n = k)$, $m < n$, no depende de p. Identifique esta distribución y obtenga sus dos primeros momentos.

4. Sea $X_n, n \geq 1$ una sucesión de v.a.i.i.d. con f.d. F. Definimos la función de distribución empírica F_n de la siguiente manera
$$F_n(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq z\}}$$
Halle el valor esperado y la varianza de las funciones indicadoras que aparecen en la suma y muestre que por la LDGN se tiene que $\lim_{n \to \infty} P(|F_n(z) - F(z)| > \varepsilon) = 0$.

5. Si $X \sim B(n, p)$ halle mediante un cálculo directo usando la función de probabilidad, el valor de $\mathrm{E}(X)$ y $\mathrm{Var}(X)$.

6. **Método de Monte Carlo** Este ejercicio es una introducción al método de Monte Carlo para el cálculo numérico de integrales definidas. El objetivo es estimar la integral
$$I = \int_0^1 x^2 \, dx.$$
a) Dibuja un gráfico de la función $y = x^2$ para $0 \leq x \leq 1$ y dibuja también un cuadrado de lado 1 con vértices en $(0,0); (0,1); (1,0); (1,1)$. Usando un generador de números aleatorios en una computadora o una tabla de números aleatorios, obtén un valor de x al azar en el intervalo $(0,1)$ y un valor de y, también en $(0,1)$. De esta manera has escogido al azar un punto (x, y) en el cuadrado unitario.

   b) Repite el proceso de seleccionar un número al azar en el cuadrado unitario $n = 100$ veces y cuenta el número N de puntos que caen debajo de la curva $y = x^2$.

   c) Usa N/n para estimar la integral I. Repite el proceso para $n = 1,000; 10,000$ y compara tus resultados con el valor de la integral I

   d) Explica el papel de la distribución binomial en este problema.

   e) Usa el método de Monte Carlo para estimar la longitud de arco de la función $y = \sin x$ para $0 \leq x \leq \pi$. (Observación: escribe la longitud de arco como una integral definida. La fórmula para la longitud de arco la puedes encontrar en algún libro de Cálculo. Esta integral no se puede calcular directamente).

7. Sea $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$ y estas variables son independientes. Demuestre que $X_1 + X_2 \sim Bin(n_1 + n_2, p)$. Extienda el resultado anterior al caso de k variables.

8. Sean X, Y v.a.i. con distribución geométrica. Demuestre que $\min(X, Y)$ y $X - Y$ son independientes.

9. Sean X, Y v.a.i. con distribución binomial negativa con parámetros (k_x, p) y (k_y, p) respectivamente. Halle la distribución condicional de $X = j$ dado que $X + Y = k$. Demuestre que cuando $k_x = k_y = 1$, $P(X = j | X + Y = k) = \frac{1}{k+1}, j = 0, 1, \dots, k, k = 0, 1, 2, \dots$

10. En el problema del error de redondeo que hemos considerado anteriormente, queremos analizar el efecto acumulado del redondeo cuando consideramos la suma de 100 cantidades redondeadas. Recordamos que las cantidades se redondean al entero más cercano, así, tanto 9.68 como 10.24 se redondean a 10. Los errores que se cometen los modelamos con variables uniformes en el intervalo $(-0.5, 0.5)$ y tenemos $X_1, \dots, X_{100}$ v.a.i. con esta distribución. a) Usando la desigualdad de Chebychef halle una cota para la probabilidad de que la suma de estas 100 variables sea mayor que 1. b) Si en lugar de considerar la suma consideramos el error absoluto medio dado por $\frac{1}{n} \sum_{i=1}^{n} |X_i|$ ¿Qué se puede decir sobre este error cuando n crece usando la LDGN?

11. Sea $X_n, n \geq 1$ una sucesión de v.a.i.i.d. con distribución $\mathcal{U}(-1,1)$ y sea $W_n = \frac{1}{n}\sum_{i=1}^n X_i^2$. ¿Es cierto que existe M tal que para cualquier $\varepsilon > 0$ se tiene que $\lim_{n\to\infty} P(|T_n - M| > \varepsilon) = 0$? Explica por qué esta proposición es cierta o falsa y si es cierta, identifica a M.

12. Se quiere determinar la cantidad de agua presente en una fruta, de la cual tenemos un número grande de muestras. En cada medición se obtiene una valor cercano al verdadero valor pero alterado por un error de medición. Modelamos esta situación a través de las variables aleatorias $M_i = H + U_i$ donde M_i es el valor porcentual de agua que se mide, H es el verdadero valor y U_i es el error aleatorio, del cual sabemos que $\mathrm{E}(U_i) = 0$ y $\mathrm{Var}(U_i) = 2$. Suponiendo que las variables U_i son independientes, ¿cuántas muestras hacen falta para estar 90 % seguro de que el promedio de las mediciones está a menos de 1 % del verdadero valor?

13. Para $t > 1$ dado, sea X una v.a. con valores 0 y t con probabilidades respectivas $1 - (1/t)$ y $1/t$. Para esta variable $\mathrm{E}(X) = 1$ y $\mathrm{Var}(X) = t - 1$. Considere la probabilidad $P(|X-1| > a)$. Si $t = 10$ y $a = 8$ entonces $P(|X-1| > a) = 1/10$ mientras que la desigualdad de Chebychef acota esta probabilidad por $9/64$. La diferencia es aproximadamente 0.04. a) Halle la diferencia para $t = 10$, $a = 5$ y $a = 10$. b) ¿Puedes hallar valores de t y a para los cuales la diferencia sea menor que 0.01, menor que 0.001, menor que 0.0001? c) Con base en los resultados que has obtenido, ¿Crees que es posible obtener una desigualdad que valga para toda variable aleatoria y que sea mejor que la de Chebychef?

14. Supón que puedes jugar un juego en el cual el valor esperado de tus ganancias es 0.01 unidades. ¿Cuál de las siguientes proposiciones es cierta? a) Si juegas un número n suficientemente grande de veces ganarás n unidades con probabilidad mayor a 0.99. b) Si juegas un número n suficientemente grande de veces ganarás $(0.009)n$ unidades con probabilidad mayor a 0.99. Explica tus respuestas.

15. Un jugador apuesta de la siguiente manera en una sucesión de volados de una moneda balanceada: Antes de cada lanzamiento escoge al azar si apuesta Aguila o Sol, con probabilidad p de escoger Aguila. Si apuesta a Aguila, gana o pierde \$1 mientras que si apuesta a Sol, gana o pierde \$2. Halle la distribución de su ganancia (o pérdida) en cada volado.

16. Una fábrica produce tornillos que se agrupan en lotes cada uno de los cuales es revisado por un inspector antes de ser despachados. El inspector revisa 15 tornillos y verifica si satisfacen o no las especificaciones. El número de objetos defectuosos es 3 o más el lote es rechazado. En otro caso se acepta.

    a) Suponga que la proporción de defectuosos es p y sea A el evento que el lote sea aceptado. Demuestre que $P(A) = (1-p)^{13}(1 + 13p + 91p^2)$.

    b) Haga la gráfica de $P(A)$ como función de p. Esta curva se conoce como la característica operativa o curva característica del proceso de revisión.

    c) Estime el valor de p para el cual $P(A) = 0.95$. Este valor se conoce como el nivel aceptable de calidad, porque representa la calidad de un lote que con alta probabilidad (0.95) pasaría una prueba de calidad.

    Considere ahora el siguiente plan alternativo de inspección. Se revisan 15 tornillos de cada lote. Si el número de defectuosos es 3 o más se rechaza el lote. Si este número es 0, el lote es aceptado. Si hay uno o dos tornillos defectuosos, se toma un segundo lote de cinco tornillos. Si el total de objetos defectuosos en las dos muestras es 2 o menos, el lote es aceptado y en caso contrario se rechaza.

    d) Demuestre que para este nuevo plan de inspección, $P(A) = (1-p)^{15} + 15p(1-p)^{19} + 180p^2(1-p)^{18}$.

    e) Haga la gráfica de la curva característica para este proceso de revisión y compárela con la del proceso anterior.

    d) Estime el nivel aceptable de calidad para este proceso y compárelo con el anterior.

17. Sean X, Y dos variables aleatorias con f.d. $F_{XY}(x,y)$. Demuestre que estas variables son independientes si y sólo si para cualquier par de intervalos $(a,b]$ y $(c,d]$ de números reales se cumple que

$$P(X \in (a,b], Y \in (c,d]) = P(X \in (a,b])P(Y \in (c,d]).$$

18. Juan lanza un dardo a una diana y la probabilidad de que acierte la diana en cada disparo es $1/2$. Dado que acierta la diana, la probabilidad de que acierte el centro de la diana es p. Juan lanza mientras acierte a la diana. Sea X el total de aciertos al centro de la diana cuando Juan termina de lanzar. Halle su distribución.

19. Julia tiene una moneda asimétrica que con probabilidad p cae águila. Ella lanza la moneda hasta obtener un águila y luego lanza una moneda simétrica tantas veces como lanzó la primera moneda. Por cada águila que obtiene con la moneda simétrica lanza un dado balanceado. Determine el valor esperado y la varianza del total de puntos que obtuvo lanzando los dados.