Modelos Estocásticos I

Primer Examen Parcial

Respuestas

1. Cierta especie de plantas produce un número N de semillas, donde N sigue una distribución de Poisson de parámetro λ. La probabilidad de que una semilla germine es p y esto ocurre para cada semilla de manera independiente a lo que ocurre con el resto de las semillas, y también es independiente del número N de semillas.

   a) Calcule la función generadora de probabilidad del número total de semillas que germinan y halle la función de distribución asociada a esta variable. A partir de la función generadora de probabilidad, halle también su media y varianza.

   b) Demuestre que el número de semillas que germinan y el número de semillas que no germinan son variables aleatorias independientes.

   c) Suponga que $\lambda = 500$ y $p = 2/3$. De las semillas que germinan, 3/4 produce matas con flores rojas y el otro cuarto produce matas con flores amarillas. ¿Qué distribución de probabilidad tiene el número de matas con flores rojas?

   Respuesta.

   a) Sea S el número de semillas que germinan. Podemos describir esta variable como

$$S = \sum_{i=1}^{N} X_i$$

   donde $N \sim \mathcal{P}ois(\lambda)$ y las X_i son v.a.i.i.d. con distribución de Bernoulli de parámetro p. Si $\phi_Z(s)$ representa la función generadora de probabilidad para la variable Z, entonces

$$\phi_S(s) = \phi_N(\phi_X(s))$$

   Tenemos que

$$\begin{aligned}
\phi_X(s) &= \mathrm{E}(s^X) = sp + q \\
\phi_N(s) &= \mathrm{E}(s^N) = \sum_{n=0}^{\infty} s^n e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda s)^n}{n!} \\
&= e^{-\lambda(1-s)} \\
\phi_S(s) &= \phi_N(\phi_X(s)) = \exp\{-\lambda(1 - \phi_X(s))\} \\
&= \exp\{-\lambda p(1 - s)\}
\end{aligned}$$

   de modo que S tiene distribución de Poisson de parámetro λp.

   Para hallar la media derivamos esta f.g.p. y evaluamos en $s = 1$:

$$\phi_S'(s) = \lambda p e^{-\lambda p(1-s)}; \qquad \phi_S(1) = \lambda p$$

   Si derivamos una vez más y evaluamos en $s = 1$ obtenemos $\mathrm{E}(S(S-1))$:

$$\phi_S''(s) = (\lambda p)^2 e^{-\lambda p(1-s)}; \qquad E(S(S-1)) = \phi_S(1) = (\lambda p)^2$$

Por lo tanto, usando la linealidad de la esperanza,

$$E(S(S-1)) = \mathrm{E}(S^2) - \mathrm{E}(S) = (\lambda p)^2$$

y en consecuencia

$$\mathrm{E}(S^2) = (\lambda p)^2 + \mathrm{E}(S) = (\lambda p)^2 + \lambda p$$

Finalmente

$$\mathrm{Var}(S) = \mathrm{E}(S^2) - (\mathrm{E}(S))^2 = \lambda p.$$

b) Sea S el número de semillas que germinan y T el número de semillas que no lo hacen. Veamos que para cualesquiera $j, k \geq 0$

$$P(S = j, T = k) = P(S = j)P(T = k).$$

$$P(S = j, T = k) = P(S = j, T = k | N = j + k)P(N = j + k)$$

pero el primer factor es una binomial de parámetros $j + k$ y p

$$\begin{aligned}
&= \binom{j+k}{j} p^j (1-p)^k e^{-\lambda} \frac{\lambda^{j+k}}{(j+k)!} \\
&= e^{-\lambda p} \frac{(\lambda p)^j}{j!} \; e^{-\lambda q} \frac{(\lambda q)^k}{k!} \\
&= P(S = j)P(T = k)
\end{aligned}$$

c) Un razonamiento similar al del inciso (a) muestra que el número de matas con flores rojas tiene distribución de Poisson de parámetro

$$500 \cdot \frac{2}{3} \cdot \frac{3}{4} = 250.$$

2. Un medidor de radioactividad puede detectar si hay o no radiación en el ambiente en periodos de 1 segundo. La probabilidad de que haya actividad de este tipo en el cuarto donde se encuentra el medidor en un segundo cualquiera es de 0.8, pero, por un defecto en el circuito, 40 % de las veces hay un error de transmisión y en lugar de registrar lo que haya ocurrido, el aparato registra con igual probabilidad que ha habido radioactividad o que no la ha habido. Cuando no hay error de transmisión el aparato registra correctamente el evento que ocurrió.

   a) En un periodo de un segundo, ¿Cuál es la probabilidad de que el aparato registre que hubo radiación?

   b) ¿Cuál es la probabilidad de que registre el resultado correcto?

   c) Si el aparato registra que hubo radiación, ¿cuál es la probabilidad de que realmente haya sido así?

   Respuesta.

   Usaremos la siguiente notación

- R: Hay radiación en el cuarto donde se encuentra el medidor.
- E: Hay un error de transmisión.
- D_R: El detector registra que ha habido radiación en el cuarto.

Tenemos que $P(R) = 0.8,\ P(E) = 0.4$.

a) La probabilidad de que el detector registre que hay radiación es

$$P(D_R) = P(D_R|E^c)P(E^c) + P(D_R|E)P(E) = 0.8 \cdot 0.6 + 0.5 \cdot 0.4 = 0.68$$

b) Llamemos C el evento de que el detector registre el resultado correcto. Su probabilidad es:

$$P(C) = P(C|E)P(E) + P(C|E^c)P(E^c) = 0.5 \cdot 0.4 + 1 \cdot 0.6 = 0.8$$

c) Tenemos

$$P(R|D_R) = \frac{P(D_R \cap R)}{P(D_R)} = \frac{P(D_R \cap R)}{0.68}$$

por el resultado del inciso (a). Veamos cuánto vale el numerador, podemos descomponer la intersección de los dos eventos según haya habido un error de transmisión o no:

$$\begin{aligned} P(R \cap D_R) &= P(R \cap D_R \cap E) + P(R \cap D_R \cap E^c) \\ &= P(D_R|R \cap E)P(E|R)P(R) + P(D_R|R \cap E^c)P(E^c|R)P(R) \end{aligned}$$

Ahora bien, si hubo error, independientemente de si ha habido radiación en el cuarto o no, el aparato registra que la ha habido con probabilidad 0.5, de modo que $P(D_R|R \cap E) = 0.5$. Por otro lado, la probabilidad de que haya un error es independiente de que haya habido o no radiación: $P(E|R) = 0.4$ y $P(E^c|R) = 0.6$. Por último, si no hay error y hay radiación, la probabilidad de que el detector registre radiación es $P(D_R|R \cap E^c) = 1$. Por lo tanto

$$P(R \cap D_R) = 0.5 \cdot 0.4 \cdot 0.8 + 1 \cdot 0.6 \cdot 0.8 = 0.64$$

Finalmente

$$P(R|D_R) = \frac{0.64}{0.68} = 0.9411$$

3. Sean X, Y v.a. con densidad conjunta

$$f(x,y) = \begin{cases} \frac{x}{\sqrt{2\pi}} e^{-\frac{1}{2}x(x+y)}, & 0 \le x < \infty, 0 \le y < \infty \\ 0, & \text{en otro caso.} \end{cases}$$

a) Halle la densidad condicional de Y dado que $X = x$.

b) ¿Son independientes estas variables?

c) Calcule $\mathrm{E}[Y|X = x]$ y $\mathrm{E}[Y|X]$.

Respuesta.

a) La densidad condicional está dada por la fórmula

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

y por lo tanto tenemos que calcular la densidad marginal de X: $f_X(x)$. Para esto, integramos la densidad conjunta respecto a y:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{1}{2}x(x+y)} \\ &= \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \int_0^{\infty} e^{-\frac{1}{2}xy} = \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big(-\frac{2}{x} e^{-\frac{1}{2}xy} \Big|_0^{\infty} \\ &= \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

para $x \geq 0$, que es la densidad del valor absoluto de una variable N con distribución normal típica. Ahora podemos calcular la desnidad condicional:

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{x}{2} e^{-\frac{x}{2}y}.$$

Por lo tanto la densidad de Y dado que $X = x$ es exponencial de parámetro $x/2$.

b) No lo son porque la densidad condicional de Y dado $X = x$ depende del valor de x.

c) Usando los resultados que conocemos para la densidad exponencial tenemos que

$$\mathrm{E}(Y|X = x) = \frac{2}{x}$$

y en consecuencia

$$\mathrm{E}(Y|X) = \frac{2}{X}.$$

4. (a) Explique cómo usaría el método de la transformada inversa para generar variables aleatorias con densidad $f(x) = 2(1-x)$ para $0 \leq x \leq 1$.

   (b) Describa un algoritmo para simular valores de una variable aleatoria con la densidad del inciso anterior usando el método del rechazo.

   (c) ¿Qué criterio usaría para seleccionar uno de los dos métodos, si tuviera que generar una muestra realmente grande?

   Respuesta.

   a) Para usar el método de la transformada inversa tenemos que hallar la función de distribución de esta densidad e invertirla. La función de distribución es

$$\begin{aligned} F(x) &= \int_0^x f(t)\,dt = \int_0^x (2-2t)\,dt \\ &= 2x - x^2 \end{aligned}$$

para $0 \leq x \leq 1$.

Para hallar la inversa observamos que F es estrictamente creciente en $0 \leq x \leq 1$ y por lo tanto tiene una inversa F^{-1}. Si $x = F^{-1}(y)$, es decir, $y = F(x)$, entonces $F(x) = 2x - x^2 = y$, es decir $x^2 - 2x + y = 0$, que tiene como raíces $1 \pm \sqrt{1-y}$. Pero en $0 \leq y < 1$, $1 + \sqrt{1-y} > 1$ y no corresponde a los valores posibles de una función de distribución. Por lo tanto

$$F^{-1}(y) = 1 - \sqrt{1-y}.$$

Es fácil verificar que $F(F^{-1}(y)) = F^{-1}(F(y)) = y$.

Por lo tanto, para generar una variable con esta distribución generamos una variable $U \sim U[0,1]$ y ponemos $X = 1 - \sqrt{1-U}$. Si tenemos en cuenta que $1-U$ también tiene distribución uniforme en $[0,1]$ podemos poner $X = 1 - \sqrt{U}$.

b) Como X toma valores en $[0,1]$ podemos usar la distribucuón uniforme para implementar el método del rechazo. En este caso tenemos $f(x) = 2(1-x)$, $g(x) = 1$ para $0 \leq x \leq 1$ y

$$\frac{f(x)}{g(x)} = \frac{2(1-x)}{1} \leq 2, \qquad \text{para } 0 \leq x \leq 1$$

Por lo tanto el algoritmo sería:

1. Generamos U_1, U_2 uniformes en $[0,1]$.
2. Si $U_2 \leq \frac{f(U_1)}{2g(U_1)} = 1 - U_1$ ponemos $X = U_1$ y paramos. Si no, volvemos al paso 1.

c) El tiempo promedio de CPU empleado para generar una variable aleatoria.