Clase Auxiliar 1

6 de Agosto de 2008

P1 Demuestre que la distribución de probabilidad exponencial tiene pérdida de memoria. Recuerde que su función densidad es $f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$.

Sol: Hay que demostrar que:

$$P[x > s+t | x > s] = P[x > t]$$

En efecto,

$$\begin{aligned}
P[x > s+t | x > s] &= \frac{P[x > s+t \wedge x > s]}{P[x > s]} \\
&= \frac{P[x > s+t]}{P[x > s]} \\
&= \frac{1 - P[x \leq s+t]}{1 - P[x \leq s]} \\
&= \frac{1 - [1 - e^{-\lambda(s+t)}]}{1 - [1 - e^{-\lambda(s)}]} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda(s)}} \\
&= e^{-\lambda t} \\
&= P[x > t]
\end{aligned}$$

P2 Demuestre que si T_1, T_2 son 2 v.a. independientes y además

$$T_1 \sim exp(\lambda), \quad T_2 \sim exp(\mu)$$

entonces se tiene que:

$$P(T_1 < T_2) = \frac{\lambda}{\mu + \lambda}$$

Sol:

$$\begin{aligned}
P[T_1 < T_2] &= \int_0^\infty P[T_1 < T_2 | T_2 = t] \cdot f_{T_2}(t) dt \\
&= \int_0^\infty P[T_1 < t] \cdot \mu e^{-\mu t} dt \\
&= \int_0^\infty (1 - e^{-\lambda t}) \cdot \mu e^{-\mu t} dt \\
&= \int_0^\infty \mu e^{-\mu t} dt - \int_0^\infty \mu e^{-(\mu+\lambda)t} dt \\
&= 1 - \tfrac{\mu}{\mu+\lambda} \cdot \int_0^\infty (\mu+\lambda) e^{-(\mu+\lambda)t} dt \\
&= 1 - \frac{\mu}{\mu+\lambda} \\
&= \frac{\lambda}{\mu+\lambda}
\end{aligned}$$

Pues, es claro que

$$P[T_1 < T_2 | T_2 = t] = P[T_1 < t | T_2 = t]$$

Y además, por independencia de T_1 y T_2, "lo que le pasa a T_1 no depende de T_2", entonces

$$P[T_1 < t | T_2 = t] = P[T_1 < t]$$

P3 Sean $X_1, X_2, X_3, \ldots, X_n$, v.a. iid. Encuentre la distribución de probabilidad de la variable aleatoria:

$$X = \text{mín}(X_1, X_2, X_3, \ldots, X_n)$$

`Sol:` Debemos encontrar una expresión para

$$P[X = \text{mín}(X_1, X_2, X_3, \ldots, X_n) \leq t]$$

Calculemos:

$$\begin{array}{rcc} P[X \leq t] & = & 1 - P[X > t] \\ & = & 1 - P[\text{mín}(X_1, X_2, X_3, \ldots, X_n) > t] \\ & = & 1 - P[X_1 > t, X_2 > t, \ldots, X_n > t] \\ & = & 1 - (P[X_1 > t] \cdot P[X_2 > t] \cdot \ldots \cdot P[X_n > t]) \\ & = & 1 - \prod_{i=1}^{n}(1 - F_{X_i(t)}) \end{array}$$

Para el caso exponencial tenemos que:

$$\begin{array}{rcc} P[X \leq t] & = & 1 - \prod_{i=1}^{n}(1 - F_{X_i(t)}) \\ & = & 1 - \prod_{i=1}^{n}(1 - 1 + e^{-\lambda t}) \\ & = & 1 - e^{-(\sum_{i=1}^{n} \lambda t)} \\ & = & 1 - e^{-n\lambda t} \end{array}$$

Así, $X \sim exp(n\lambda)$.

P4 Encuentre la distribución de probabilidad de una variable aleatoria N que es la suma de 2 variables aleatorias independientes, N_1 y N_2, que siguen distribuciones de Poisson de parámetros $\lambda_1, \lambda_2 > 0$, ie,

$$P(N_i = k) = \frac{e^{-\lambda_i} \lambda^k}{k!}, \ i = 1, 2$$

`Sol:` Como las distribucón de Poisson es una v.a. discreta y positiva, claramente la suma también lo será, luego nos debemos calcular $P[N = k]$, $\forall k \in \mathbb{N}$. Calculemos:

$$
\begin{aligned}
P[N=k] &= P[N_1+N=k] \\
&= \sum_{i=0}^{\infty} P[N_1+N_2=k|N_2=i]\cdot P[N_2=i] \\
&= \sum_{i=0}^{k} P[N_1=k-i]\cdot P[N_2=i] \\
&= \sum_{i=0}^{k} \frac{\lambda_1^{k-i}e^{-\lambda_1}}{(k-i)!}\cdot\frac{\lambda_2^{i}e^{-\lambda_2}}{i!} \\
&= \tfrac{e^{-(\lambda_1+\lambda_2)}}{k!}\cdot\sum_{i=0}^{k}\frac{k!}{(k-i)!i!}\lambda_1^{k-i}\lambda_2 i \\
&= \tfrac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k}{k!}
\end{aligned}
$$

Es conclusión, N es una v.a. de Poisson de parámetro $(\lambda_1+\lambda_2)$.

P5 Usted se encuentra dando una prueba de su ramo preferido el cual está siendo cuidado por dos ayudantes. El primer ayudante dice la verdad $\frac{3}{4}$ de las veces. Por su parte el segundo de los ayudantes dice la verdad $\frac{4}{5}$ de las veces. La pregunta a una pregunta particular tiene 9 alternativas, y usted no sabe cuál es la respuesta correcta (pero los ayudantes sí la conocen). Por esto, usted ha decidido preguntarle a los 2 ayudantes su opinión respecto a la alternativa correcta.

Si ambos afirman que la alternativa correcta es la **(d)**, ¿Cuál es la probabilidad que efectivamente esa sea la respuesta correcta a la pregunta (es decir, que ambos ayudantes estén diciendo la verdad)?.

Hint: La probabilidad *a priori* que cualquiera de las alternativas sea la correcta es $\frac{1}{9}$.

`Sol:` Sean las variables:

$$
\begin{aligned}
R &= \text{Respuesta correcta a la pregunta.} \\
A_1 &= \text{Respuesta dada por el ayudante 1 a la pregunta.} \\
A_2 &= \text{Respuesta dada por el ayudante 2 a la pregunta.}
\end{aligned}
$$

Tenemos que calcular $P(R=d|A_1=d\wedge A_2=d)$. Usando el Teorema de Bayes podemos escribir:

$$
P(R=d|A_1=d\wedge A_2=d) = \frac{P(A_1=d\wedge A_2=d|R=d)\cdot P(R=d)}{P(A_1=d\wedge A_2=d)}
$$

Calculemos cada término de lado derecho de forma explícita:

I.-

$$
\begin{aligned}
P(A_1=d\wedge A_2=d|R=d) &= P(A_1 \text{ diga la verdad})\cdot P(A_2 \text{ diga la verdad}) \\
&= \frac{3}{4}\cdot\frac{4}{5}
\end{aligned}
$$

II.-

$$
P(R=d) = \frac{1}{9}
$$

III.-

$$\begin{aligned} P(A_1 = d \wedge A_2 = d) &= \sum_{r=a}^{i} P(A_1 = d \wedge A_2 = d | R = r) \cdot P(R = r) \\ &= \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{1}{9} + \sum_{\substack{r = a \\ i \neq d}}^{r} \frac{1}{8} \cdot \frac{1}{8} \cdot \frac{1}{4} \cdot \frac{1}{5} \cdot \frac{1}{9} \end{aligned}$$

Entonces, de los cálculos anteriores:

$$\begin{aligned} P(R = d | A_1 = d \wedge A_2 = d) &= \frac{\frac{1}{15}}{\frac{1}{1440} + \frac{1}{15}} \\ &= \frac{96}{97} \end{aligned}$$

P6 Sean $X_1, X_2, ..., X_n$ variables aleatorias independientes continuas de densidad común f. Sea $X_{(i)}$ el i-ésimo menor valor entre $X_1, X_2, ..., X_n$.

a) Demuestre que la densidad de $X_{(i)}$ está dada por:

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1} \overline{F}(x)^{n-i} f(x)$$

b) $X_{(i)}$ será menor que x si, y sólo si, cuántas de las variables X son menores que x?

c) Use *b)* para obtener una expresión para $P[X_{(i)} \leq x]$.

d) Utilice *a)* y *c)* para establecer la siguiente identidad para $0 \leq y \leq 1$:

$$\sum_{k=i}^{n} \binom{n}{k} y^k (1-y)^{(n-k)} = \int_0^y \frac{n!}{(i-1)!(n-i)!} s^{i-1} (1-s)^{n-i} ds$$

Sol: *a*) Para que $X_{(i)}$ sea igual a x obligatoriamente $i-1$ de las variables X_i tienen que ser menores a x y $n-i$ mayores a x. Como existen $\binom{n}{i-1}$ formas de escoger que variables X_i serán menores que x y de las $(n-(i-1))$ restantes existen $\binom{n-(i-1)}{1}$ formas de escoger cuál será la i-ésima, tenemos que:

$$f_{X_{(i)}}(x) = \binom{n}{i-1} \binom{n-(i-1)}{1} F(x)^{i-1} \overline{F}(x)^{n-i} f(x) = \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1} \overline{F}(x)^{n-i} f(x)$$

b) Esto ocurrirá si i o más variables X_i son menores que x.

c) Claramente:

$$P[X_{(i)} \leq x] = \sum_{j=i}^{n} \binom{n}{j} F(x)^j \overline{F}(x)^{n-j}$$

d) Calculamos $P[X_{(i)} \leq x]$, integrando la funci'on de densidad encontrada en *a)* desde 0 hasta x.Y usamos *c)*para encontrar la identidad:

$$\sum_{j=i}^{n} \binom{n}{j} F(x)^j \overline{F}(x)^{n-j} = \int_0^x \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1} \overline{F}(x)^{n-i} f(x) dx$$

Realizando el cambio de variable $y = F(x)$ y usando que $dy = dF(x) = f(x)dx$ tendremos que:

$$\sum_{j=i}^{n} \binom{n}{j} y^j (1-y)^{(n-j)} = \int_0^y \frac{n!}{(i-1)!(n-i)!} s^{i-1}(1-s)^{n-i} ds$$

P7 Suponga que el dueño de un centro de dental quiere definir la política de citaciones de pacientes. Cada dentista puede atender 2 pacientes en el período en estudio. El primer paciente se cita al comienzo de la sesión $t_1 = 0$ y el segundo se cita en el instante $t_2 = x$. Se ha observado que todos los pacientes son puntuales.

Los tiempos de cada atención son variables aleatorias i.i.d. con distribuciones Uniformes entre a y b $[hrs]$. Así, cuando el segundo paciente llega al centro dental existen dos posibilidades:

- El dentista está ocupado atendiendo al primer paciente. En este caso, el segundo cliente debe esperar hasta que termine la atención del primero para ser atendido.
- El dentista está ocioso, ya que terminó de atender al primer paciente. El segundo cliente se atiende de inmediato.

El dentista cobra $M[\$/hr]$ independientemente si está trabajando o no, por lo que a Ud. le interesa que este se desocupe en el menor tiempo posible de la atención de los 2 pacientes. Sin embargo, si un paciente llega y su dentista está ocupado el centro percibe un costo de $P[\$/hr]$ mientras éste espera.

Se quiere buscar cuál es el instante x óptimo para citar al 2° paciente si el tomador de decisiones es neutral al riesgo. Para ello conteste las siguientes preguntas.

a) Si el segundo paciente es citado en el instante x y por lo tanto llega al sistema en el instante x (recuerde que es puntual), ¿Cuál es la probabilidad que encuentre al dentista ocupado atendiendo al primer paciente y por lo tanto tenga que esperar para recibir su atención?.

b) Si el segundo cliente es citado a la consulta en el instante x. ¿Cuál es la esperanza del tiempo de espera si encuentra al dentista ocupado?¿Cuál es la esperanza si lo encuentra desocupado? Deduzca de lo anterior la esperanza del tiempo de espera.

c) ¿Cuál es la esperanza del tiempo que el dentista estará en el centro dental?.

d) Utilizando las partes anteriores, defina el problema de optimización que debe resolver el dueño de la consulta. Encuentre el momento óptimo para citar al segundo paciente.

Sol: Para determinar x^*, el momento óptimo para citar al segundo paciente, seguiremos los pasos dados enumerados en el enunciado del problema.

Sean

- T_i: tiempo de atención del paciente i, con $i = 1, 2$.
- W_2: el tiempo de espera del paciente 2.

Notar que T_1, T_2, W_2 son variables aleatorias.

a) Primero debemos calcular la probabilidad que, dado un x, el segundo paciente encuentre al dentista ocupado. Esto es equivalente a que la atención del paciente 1 se demore más de x unidades de tiempo, es decir $T_1 \geq x$. Hagamos el calculillo que nos piden entonces:

$$P_x[\text{Encontrar al dentista ocupado}] = P_x[T_1 \geq x] = \int_x^b \frac{dt}{(b-a)} = \frac{b-x}{(b-a)}$$

pues T_1 es una v.a. uniforme en $[a, b]$.

Comentarios: Vemos que no tiene sentido que $x > b$ dado que implica que con certeza el dentista estará libre. Tampoco tiene sentido un $x < a$ dado que con certeza el dentista estará ocupado. Entonces, el dominio de x, para el problema de optimización, será $x \in [a, b]$.

b) Primero, note que el tiempo de espera está dado por la fórmula:

$$W_2 = T_1 - x$$

Separemos los casos:

- Si encuentra al *dentista ocupado* la esperanza del tiempo restante de espera será:

$$E_x[W_2|\text{dentista ocupado}] = \int_x^b \frac{(t-x)}{(b-x)} dt = \frac{b-x}{2}$$

  Lo anterior se tiene pues, dado que *el dentista está ocupado*, T_1 **no** es una variable con distribución uniforme en $[a, b]$, sino que tiene una distribución condicional al evento *dentista ocupado*. Como en este caso $T_1 \geq x$, la distribución condicional de T_1 es entonces una uniforme en $[x, b]$.

- Si encuentra al *dentista desocupado* la esperanza del tiempo restante de espera será:

$$E_x[W_2|\text{dentista desocupado}] = 0$$

  Pues en este caso el paciente es atendido apenas llega. No espera!!.

Entonces, combinando los dos resultados anteriores, "condicionando" y "descondicionando", tenemos que:

$$\begin{aligned} E_x[W_2] &= E_x[W_2|\text{dentista desocupado}] \cdot P_x[\text{dentista desocupado}] \\ &\quad + E_x[W_2|\text{dentista ocupado}] \cdot P_x[\text{dentista ocupado}] \\ E_x[W_2] &= \frac{b-x}{b-a} \cdot \frac{b-x}{2} = \frac{(b-x)^2}{2 \cdot (b-a)} \end{aligned}$$

c) Claramente el dentista estará en el consultorio hasta que termine la atención del segundo paciente:

$$\begin{aligned} E_x[\text{Tiempo del dentista en consulta}] &= x + E_x[W_2] + E_x[T_2] \\ &= x + \frac{(b-x)^2}{2 \cdot (b-a)} + \frac{a+b}{2} \end{aligned}$$

d) Entonces el problema de minimización a enfrentar será el siguiente:

$$\min_{a \leq x \leq b} \left(M \cdot E_x[\text{Tiempo del dentista en consulta}] + P \cdot E_x[W_2] \right)$$

es decir

$$\min_{a \leq x \leq b} \left(M \cdot [x + \frac{(b-x)^2}{2 \cdot (b-a)} + \frac{a+b}{2}] + P \cdot [\frac{(b-x)^2}{2 \cdot (b-a)}] \right)$$

Si derivamos e igualamos a 0 tenemos que:

$$x^* = b - \frac{M \cdot (b-a)}{M+P}$$

También hay que evaluar en los extremos del intervalo pues como es un problema de optimización con restricciones, no sirve sólo derivar (en control derive no más `:D`).

La otra opción es argumentar de que la función objetivo es una parábola y encontrar el vértice con la fórmula "conocida" del colegio.

P8 Suponga que $\{N_1(t)\}_{t\geq 0}$ y $\{N_2(t)\}_{t\geq 0}$ son 2 procesos de poisson independientes con tasas de llegada λ_1 y λ_2, respectivamente. Pruebe que $\{N_1(t) + N_2(t)\}_{t\geq 0}$ es un proceso de Poisson con tasa $\lambda_1 + \lambda_2$. Muestre además que la probabilidad de que el primer evento del proceso combinado provenga de $\{N_1(t)\}_{t\geq 0}$ es $\frac{\lambda_1}{\lambda_1+\lambda_2}$, independiente del tiempo en que ocurra el evento.

Sol: Llamemos $\{N(t)\}_{t\geq 0}$ al proceso combinado, es decir, $N(t) = N_1(t) + N_2(t), \ \forall\ t \geq 0$. Para probar que $\{N(t)\}_{t\geq 0}$ es un proceso de Poisson debemos verificar las siguientes condiciones:

I.- $N(0) = 0$

II.- El proceso posee incrementos independientes.

III.- $P(N(t+s) - N(s) = k) = \frac{e^{-(\lambda t)}(\lambda t)^k}{k!},\ i = 1, 2$. En este caso para $\lambda = \lambda_1 + \lambda_2$.

Veamos que todas se cumplen:

I.- $N(0) = N_1(0) + N_2(0) = 0 + 0 = 0$

II.- Como $N_1(t), N_2(t)$ poseen incrementos independientes, es claro que "lo que pase con la suma" en un intervalo será independiente de "lo que pase con la suma" en otro intervalo disjunto al primero. Lo anterior pues la suma depende explícitamente de $N_1(t), N_2(t)$.

III.- Se demuestra de la misma manera que en la pregunta P4. Hacerlo como buen ejercicio para practicar "condicionamiento/descondicionamiento".

Como se cumplen las tres condiciones anteriores, $\{N_1(t)+N_2(t)\}_{t\geq 0}$ es efectivamente un proceso de Poisson con tasa $\lambda_1 + \lambda_2$.

Para calcular la probabilidad que nos piden definamos:

- X_1: tiempo de la primera llegada del proceso $N_1(t)$.
- Y_1: tiempo de la primera llegada del proceso $N_2(t)$.

De clases sabemos que $X_1 \sim exp(\lambda_1)$, $Y_1 \sim exp(\lambda_2)$.

Con estas deficniciones en mente, lo que debemos calcular es:

$$P[X_1 < Y_2]$$

Que es justamente lo que se hizo en la P2. Hacerlo de nuevo como ejercicio.

El resultado que se obtiene es

$$P[X_1 < Y_2] = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Gustavo Angulo
gangulo@dim.uchile.cl

Diego Morán
dmoran@dim.uchile.cl