GUIA DE EJERCICIOS MA3401

Junio de 2010

Profesor: Joaquín Fontbona

0. Sean $\lambda, \alpha_i, i \in \mathbb{N}\backslash\{0\}$ reales estrictamente positivos, y $(X_1, X_2, ...)$ v.a. independientes, con $X_i \sim Gamma(\alpha_i, \lambda)$, es decir X_i tiene densidad

$$f_{X_i}(x) = c(\alpha_i, \lambda)e^{-\lambda x}x^{\alpha_i - 1}\mathbf{1}_{x>0},$$

con $c(\alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)}$ y $\Gamma(\alpha) = \int_0^\infty e^{-y}y^{\alpha-1}dy$ la función gamma.

Pruebe por inducción que $\sum_{j=1}^n X_j \sim Gamma(\sum_{j=1}^n \alpha_j, \lambda)$. Ind.: Puede usar (sin probar) el hecho que si $(Y_1, Y_2, ...)$ son v.a. independientes, entonces para toda función $F : \mathbb{R}^m \to \mathbb{R}$ medible, $(F(Y_1, ..., Y_m), Y_{m+1}, Y_{m+2}, ...)$ son nuevas v.a. e independientes también. Recuerde también que el caso $n = 2$ se vio en clases...

1. **Ley de Poisson y ley multinomial**

   a. Sean $Z_j \sim \mathcal{P}oiss(\lambda_j)$ con $j = 1 \dots, m$ v.a. independientes y $Z := \sum_{j=1}^m Z_j$. Pruebe que la ley de $(Z_1, \dots, Z_m)$ condicional a $Z = n$ es multinomial de parámetros $\left(n; \frac{\lambda_1}{\lambda}, \dots, \frac{\lambda_m}{\lambda}\right)$, con $\lambda = \sum_{j=1}^m \lambda_j$. Es decir

   $$P\left(Z_1 = n_1, \dots, Z_m = n_m \middle| \sum_{j=1}^m Z_j = n\right) = \frac{n!}{n_1! \cdots n_m!}\left(\frac{\lambda_1}{\lambda}\right)^{n_1} \cdots \left(\frac{\lambda_m}{\lambda}\right)^{n_m}$$

   si $n_1 + \cdots + n_m = n$, y 0 si no.

   b. Muestre que, recíprocamente, si $Z \sim \mathcal{P}oiss(\lambda)$ con $\lambda \in]0, \infty[$, y $\mathbf{M} := (\mathbf{M}_1, \dots, \mathbf{M}_m)$ es una v.a. a valores en $\mathbb{N}^m$ que condicionalmente a $Z = n$ tiene ley multinomial de parámetros $(n; p_1, \dots, p_m)$, entonces las coordenadas de $\mathbf{M}$ son independientes entre sí, y $\mathbf{M}_j \sim \mathcal{P}(p_j\lambda)$ para cada j.

   Nota. Un caso particular importante de la parte b. es una variable Y "de Poisson filtrada", es decir, dada $Z \sim \mathcal{P}oiss(\lambda)$, Y es una v.a. que condicionalmente a $Z = m$, tiene ley $Bin(m, p)$, donde $p \in (0, 1)$. (Note que entonces el par $(Y, m - Y)$ es multinomial de parámetros $(m; p, 1 - p)$, por eso es un caso particular de lo anterior)

   Lo de "filtrado" viene de interpretar esto de la manera siguiente: para cada uno de los Z "objetos contados", se lanza de manera independiente una moneda (que tiene probabilidad p de dar cara), y se "guardan en el conteo sólo una cantidad Y", que son aquellos a los que les tocó cara.

2. **Sobre el proceso de Poisson.**

   Sean $(S_n)_{n\in\mathbb{N}\backslash\{0\}}$ v.a. independientes de ley $\mathbf{e}xp(\lambda)$ con $\lambda \in]0, \infty[$. Nos referiremos a estas variables como los tiempos "entre llegadas". Llamaremos "instantes de llegadas" a la sucesión de v.a. definidas para $\omega \in \Omega$ por

   $$T_0(\omega) = 0, \quad T_1(\omega) = S_1(\omega), \quad T_n(\omega) = S_1(\omega) + \cdots + S_n(\omega).$$

El **proceso de Poisson en $\mathbb{R}_+$ de parámetro λ** es la familia $(N_t)_{t\in\mathbb{R}_+}$ de variables aleatorias (indexada por $\mathbb{R}_+$) definida por

$$N_t(\omega) = \sup\{n \in \mathbb{N} : T_n(\omega) \leq t\}$$

Note que para cada $\omega \in \Omega$, la variable aleatoria $N_t(\omega)$ *cuenta* el número de llegadas ocurridas en el intervalo $]0, t]$. (Puede pensar que las v.a. T_n representan por ejemplo, los instantes en que llegan clientes a un banco).

a. Pruebe que para cada $t \in \mathbb{R}_+$, $N_t < \infty$ c.s.

   Ind.: Muestre que $\mathbb{P}(T_n \leq t) \leq \mathbb{P}(S_1 \leq t, \ldots, S_n \leq t) \to 0$ cuando $n \to \infty$. Relacione el evento $\{N_t = \infty\}$ con los eventos $\{T_n \leq t\}$ y deduzca que $\mathbb{P}(N_t = \infty) = 0$.

b. Pruebe que para todo $t \in \mathbb{R}_+$ el número de llegadas ocurridas hasta el instante t tiene ley de Poisson de parámetro λt, es decir,

$$N_t \sim \mathcal{P}oiss(\lambda t).$$

   Ind.: Muestre que $\{N_t = n\} = \{T_n \leq t < T_n + S_{n+1}\}$. Luego, explique porqué T_n y S_{n+1} son independientes (no se pide probarlo en detalle) e identifique la ley de T_n. Usando todo esto, calcule $\mathbb{P}(N_t = n)$. Le será útil el hecho que $\Gamma(n) = \int_{\mathbb{R}_+} x^{n-1}e^{-x}dx = (n-1)!$, lo que se muestra por inducción.

c. Usando lo anterior, pruebe que para cada $t \geq 0$,

$$\mathbb{P}(T_{n+1} - t > s | N_t = n) = \mathbb{P}(T_1 > s) = e^{-\lambda s}$$

   Para cada tiempo $t > 0$, diga entonces cuál es la densidad condicional del tiempo restante hasta la siguiente llegada, dado el número de llegadas ya ocurridas.

3. **Suma de una cantidad geométrica de exponenciales**

Sean $(S_n)_{n\in\mathbb{N}\setminus\{0\}}$ v.a. independientes de ley $\mathbf{e}xp(\lambda)$ con $\lambda \in]0, \infty[$. Sea G una v.a. geométrica de parámetro $p \in (0, 1)$, independiente de las variables (S_n).

Definimos la variable

$$S := \sum_{j=1}^{G} S_j$$

es decir, en el evento $\{G = m\}$, se tiene $S = \sum_{j=1}^{m} S_j$.

a. Calcule para cada $m \in \mathbb{N}\setminus\{0\}$ y $t > 0$, $\mathbb{P}(S > t | G = m)$.

   (Ind: tenga presente la ley de una suma de v.a. exponenciales independientes de mismo parámetro)

b. Usando lo anterior deduzca que S tiene ley exponencial de parámetro λp.

c. De un ejemplo de S en términos de "tiempos de pasada de buses" y "lanzamientos de un moneda".

d. Relacione este resultado con la parte b. de la pregunta 2 y con la nota de la pregunta 1.

4. **Una relación entre leyes Gamma y Beta**

   Sean X_1, X_2 v.a. independientes de ley Gamma(α, λ), es decir con densidad $ce^{-\lambda x}x^{\alpha-1}\mathbf{1}_{x>0}$, con $\alpha, \lambda > 0$ y la constante $c > 0$ adecuada.

   Pruebe que condicionalmete a $X_1 + X_2 = x$, $X_1/(X_1 + X_2)$ tiene densidad Beta, explicitando los parámetros . (Ind.: le será útil calcular primero la ley conjunta del par $(X_1, (X_1 + X_2))$. Puede inspirarse tambinén de un ejemplo hecho en clases)

5. **Convolución de leyes discretas**

   a. Sean X, Y v.a. independientes a valores en $\mathbb{N}$, con $p_i^X = \mathbb{P}(X = i)$ y $p_j^Y = \mathbb{P}(Y = j)$. Defina $Z = X + Y$. Escriba $p_k^Z = \mathbb{P}(Z = k)$ en términos de (p_i^X) y (p_i^Y) y compare la expresión con la densidad de la suma de v.a. continuas independientes.

   b Sean $X_1, X_2, ...$ v.a. discretas independientes, y de manera genérica, sea $g_Z(s)$ la función generadora de momentos de una v.a. discreta Z. Pruebe por inducción que $g_{X_1+...+X_n}(s) = \prod_{j=1}^n g_{X_j}(s)$.

6. **Sobre "Procesos de Ramificación".**

   Se modela la evolución temporal aleatoria del tamaño Z_n de una población de plantas acuáticas de la manera siguiente:

   i) $Z_0 = 1$, es decir la población se inicia en la generación $n = 0$ con un único ancestro.

   ii) Condicionalmente a $Z_n = m$, los m individuos de la $n-$ésima generación tienen cada uno cantidades aleatorias de hijos, $X_1^{(n)}, ..., X_m^{(n)}$ respectivamente, con los $X_i^{(n)}$ independientes y con una misma distribución discreta F en $\mathbb{N}$. (F es la ley del número de hijos que tiene un individuo cualquiera)

   iii) En el evento $Z_n = m$, se define $Z_{n+1} := X_1^{(n)} + \cdots + X_m^{(n)}$, es decir, los individuos de la generación $n + 1$ son todos los hijos de los m individuos de la generación n.

   Sea $g_Z : [0, 1] \to \mathbb{R}_+$ la función generadora de momentos de una v.a. Z de ley F, $g_Z(s) := \mathbb{E}(s^Z)$. Se calculará $\varphi_n(s) := \mathbb{E}(s^{Z_n}), s \in [0, 1]$ la función generadora de momentos de Z_n.

   a. Diga cuál es la ley condicional de Z_{n+1} dado que $Z_n = m$. Pruebe entonces que $\mathbb{E}(s^{Z_{n+1}}|Z_n = m) = (g_Z(s))^m$. (Puede usar la pregunta 5.).

   b. Pruebe por inducción que

   $$\varphi_n(s) = g_Z^{\circ n}(s) := g_Z \circ \cdots \circ g_Z(s) \text{ (convolución } n \text{ veces)},$$

   donde $g_Z^{\circ 0}(s) = s$.

   c. Si Z se ditribuye de manera equiprobable en $\{0, 1, 2\}$, muestre que la probabilidad $\mathbb{P}(Z_3 = 0)$ de que al cabo de $n = 3$ años la población desaparezca es igual a $\frac{1249}{2187}$.

7. **Simulación de variables aleatorias por método de"aceptación-rechazo".[40%]**

Sean f y g dos densidades de probabilidad en $\mathbb{R}^d$. Se cuenta con un lenguaje informático que es capaz de simular realizaciones de variables aleatorias de densidad g. El lenguaje también también puede generar variables aleatorias de ley $U[0,1]$.

Suponga que todas las realizaciones distintas de v.a. (cualesquiera) son independientes. Suponga además que existe una constante $c \in (1, \infty)$ tal que

$$\sup_{x \in \mathbb{R}^d} \frac{f(x)}{g(x)} \leq c,$$

y que las funciones g y f son ambas conocidas y sus valores calculables en todo punto $x \in \mathbb{R}^d$.

El objetivo de este problema es construir, bajo las hipótesis anteriores, un método que permite simular una v.a. Z cuya ley tiene densidad f.

Para esto, considere una sucesión $(X_n)_{n\geq 1}$ de v.a. i.i.d. en $\mathbb{R}^d$ con densidad g, y una sucesión $(U_n)_{n\geq 1}$ de v.a. i.i.d. de ley $U[0,1]$ independientes de las v.a. $(X_n)_{n\geq 1}$.

a. Verifique que $\mathbb{P}(g(X_n) = 0) = 0$ para todo $n \geq 1$.

b. Pruebe que

$$\mathbb{P}\left(U_n \geq \frac{f(X_n)}{cg(X_n)}\right) = 1 - p$$

donde $p := \frac{1}{c}$. (Ind.: puede calcular primero $\mathbb{P}(U_n \geq \frac{f(X_n)}{cg(X_n)} | X_n = x)$ o si lo prefiere, hacer el cálculo directamente integrando en un dominio de $\mathbb{R}^{d+1}$ apropiado.)

c. Se define la variable aleatoria

$$T(\omega) := \inf \left\{ n \in \mathbb{N}\backslash\{0\} : U_n(\omega) < \frac{f(X_n(\omega))}{cg(X_n(\omega))} \right\},$$

(es decir, suponiendo que se simulan sucesivamente $(U_1, X_1), (U_2, X_2), ...$ entonces T es el n tal que $U_n < \frac{f(X_n)}{cg(X_n)}$ ocurre por primera vez). Calcule $\mathbb{P}(T > k)$ y deduzca que T tiene ley geométrica de parámetro p.

d. Sea ahora $A \subseteq \mathbb{R}^d$ un Boreliano. Pruebe que

$$\mathbb{P}\left(U_n < \frac{f(X_n)}{cg(X_n)}, X_n \in A\right) = \frac{1}{c} \int_A f(z)dz.$$

(Ind.: tiene alternativas similares a las de la parte b. para proceder.)

Se define ahora la v.a. Z como $Z(\omega) := X_{T(\omega)}(\omega)$ o, de manera equivalente,

$$Z = X_m \text{ en el evento } \{T = m\},$$

(es decir, el valor que tomó X_n la primera vez n que ocurrió que $U_n < \frac{f(X_n)}{cg(X_n)}$).

e. Para $A \subseteq \mathbb{R}^d$ un Boreliano y $m \in \mathbb{N}\backslash\{0\}$, exprese el evento $\{Z \in A, T = m\}$ en términos sólo de las v.a. $(X_n)_{n\geq 1}$ y $(U_n)_{n\geq 1}$, y deduzca de esto y todo lo anterior que

$$\mathbb{P}(Z \in A) = \int_A f(z)dz.$$

f. En base a las partes anteriores describa un algoritmo para simular v.a. $Z_1, \ldots, Z_k$ i.i.d. con densidad f.

7. Se sabe que un átomo individual del isótopo radioactivo Carbono 14 (C14), después de haberse constituido o sintetizado, se degrada tras un "tiempo de vida" aleatorio, que tiene ley exponencial de parámetro λ. Al hacerlo, se transforma en un átomo estable de Carbono "usual" (C12) . Además, la desintegración de un átomo no tiene ninguna influencia sobre átomos próximos.

   Por otro lado, es sabido que cada organismo vivo absorbe constantemente (al respirar) átomos de C14, de forma tal que la razón entre la cantidad de átomos de este isótopo y la cantidad total de átomos de Carbono que porta (es decir, C12 más C14) es constante durante su vida e igual a un número $\alpha \in (0,1)$ conocido (que corresponde a la proporción de C14 en la atmósfera). Sin embargo, cuando muere, ya no absorbe más C14, y la proporción de C14 comienza a disminuir debido a la desintegración natural de los átomos de isótopo.

   a. Se estudia el tejido de un organismo muerto hace mucho tiempo. Explique por qué es razonable suponer, en base a todo lo anterior y propiedades que conoce de la ley exponencial, que todos los átomos de C14 hayados en el presente en este tejido (es decir, aquellos que no se han desintegrado todavía) fueron sintetizados en el momento de la muerte del organismo.

   b. En base a lo anterior, y suponiendo que la cantidad de átomos de C14 en el tejido es muy grande, muestre que la proporción $\beta \in (0,1)$ de átomos de C14 que no se han desintegrado (entre el total del átomos de C14 presentes en el momento de la muerte) es, pasado un tiempo t desde la muerte del organismo, aproximadamente igual a $e^{-\lambda t}$.

      Ind: Suponga que hay infinitos átomos de C14 presentes en el momento de la muerte del organismo, y "enumerelos" $i = 1, \ldots.$ Sean $(X_i)_{i \geq 1}$ los tiempos de "vida" de cada uno de ellos. Note entonces que si se consideran "los n primeros átomos de C14" de la lista, la proporción de ellos que no se han desintegrado transcurrido un tiempo t es

      $$\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i > t\}}.$$

      Luego, vea que pasa cuando n es grande.

   c. Se calcula la masa total de Carbono del tejido, resultando esta ser igual a m. Por otro lado, se determina que la masa de C14 en el mismo tejido es m'. Hace cuánto tiempo murió el organismo estudiado? (Asuma que la diferencia de masa entre átomos de C12 y C14 es despreciable) R: $\frac{\ln(\frac{m'}{\alpha m})}{-\lambda}$)

      Nota: el tiempo tal que $\beta = \frac{1}{2}$ se llama "vida media" del isótopo.

8. "La aguja de Buffon"

   Considere un piso de "parquet" consistente en bandas paralelas de madera muy largas y de ancho $a > 0$.

   a. Se lanza una aguja de largo $l \in (0, a)$ sobre el piso. Cuál es la probabilidad de que caiga sobre dos bandas de madera.

Ind.: Suponga que la distancia X desde el centro de la aguja al momento de caer, hasta el borde de banda más cercano, se distribuye uniformemente en $[0, a/2]$. Suponga además que al caer, el ángulo θ que forma la aguja con el eje perpendicular a la dirección de las bandas se distribuye uniformemente en $[-\frac{\pi}{2}, \frac{\pi}{2}]$, y que X y θ son independientes.

Escriba el evento estudiado en términos de ambas variables aleatorias y deduzca que la probabilidad de que ocurra es $\frac{2l}{\pi a}$.

(Nota: este problema fue planteado y resuelto por Georges Louis Leclerc Conde de Buffon, naturalista francés, 1707 - 1788)

b. Ud. dispone de una aguja y un parquet como en la parte a., además de mucho tiempo libre y paciencia. Diga como calcular π con estos elementos.

9. a. Sean $X_1, X_2, ...$ la cantidad de público que llega a un cine nuevo con 5 salas en los días $1, 2, ...$ Se considera que X_i se modelan bien por v.a. de Poisson independientes de parámetro λ.

   a.i Dos meses después de la inauguración, han venido 30.000 personas. Dé una estimación de λ.

   a.ii El cine deberé cerrar si el público total durante el próximo año es menor a 100000 personas. Usando TCL, estime la probabilidad de que esto ocurra.

   b. Cuántas veces hay que lanzar una moneda equilibrada para que la proporción de caras obtenidas esté con una probabilidad superior o igual a 0.96 en el intervalo $[1/2 - 0.01, 1/2 + 0.01]$? (Ind: use TCL)

10. (TCL) Una linea aerea debe optimizar sus vuelos de 200 pasajeros, teniendo en cuenta la siguiente información:

    • el peso en kilos de cada pasajero es una v.a. de esperanza 65 y desviación estandar (es decir, raíz de la varianza) igual a 15;
    • cuando el peso máximo de equipaje autorizado por pasajero es Mkg, el peso del equipaje de cada pasajero es una v.a. de esperanza $0.7M$ y desviación estandar $0.2M$;
    • las 400 v.a. correspondientes a los pesos de los pasajeros y el de sus equipajes son independientes.

    Cúal es el valor máximo de M que permite que en el 95% de los vuelos el peso total de los 200 pasajeros y su equipaje no sobrepase el valor máximo autorizado de 18 toneladas?

11. (TCL) Un restaurante puede servir 75 comidas por noche, y sólo acepta clientes con reservación. La experiencia muestra que 20% de los clientes que reservan no vienen. Cuántas reservaciones debe aceptar el restaurant para que la probabilidad de poder servir a todos los clientes que vendrán sea mayor o igual a 0.9? (Ind: utilise TCL)

12. Una compañia de seguros se propone asegurar 100000 clientes contra robo. Los montos en miles de pesos $X_1, \ldots, X_{100000}$ que serán cobrados cada año a la compañia (la gran mayoría de ellos 0) son v.a. independientes de esperanza 75 y desviación estandar 750. Qué prima anual debe cobrar la companía a cada asegurado, para que sus gastos de operación, evaluados en 1500 millones de pesos, sean cubiertos con probabilidad mayor o igual a 0.999? (Ind: utilise TCL)

13. Una central telefónica se contruye para 5000 lineas. Cada día, cada usuario llama en hora de punta con probabilidad $p = 0.02$.

    Qué capacidad de llamadas simultáneas debe tener la central para que la probabilidad de saturarse sea menor que 0.1? Haga el cálculo con dos aproximaciones distintas de la ley binomial. Compare los resultados y diga cuál aproximación es más "razonable" (Ind: note que p es bastante pequeño.)