Probabilidad II | **Matemáticas**

Curso 2008-2009 | 5 de junio de 2009

Soluciones del Examen Final

---

1. Sean X e Y variables aleatorias independientes idénticamente distribuidas con función de densidad de probabilidad $f(x) = e^{-x}\mathbf{1}_{\{x>0\}}$.

Se consideran las variables aleatorias $S = X + Y$ y $Q = X/Y$. Se pide

(1) Calcular las distribuciones de probabilidad de S y Q, encontrando, si existen, las correspondientes funciones de densidad de probabilidad.

(2) Calcular la distribución de probabilidad conjunta $P_{(S,Q)}$ y explicar si las variables S y Q son independientes.

Solución: • **Primera forma**

(1) – Para todo $A \in \mathcal{B}(\mathbb{R})$,

$$P\{S \in A\} = P_{(X,Y)}\{(x,y) \in \mathbb{R}^2 : x + y \in A\} = \int_0^\infty \int_0^\infty \mathbf{1}_A(x+y)e^{-x}e^{-y}dxdy.$$

Para calcular esta integral hacemos el cambio de variable $x + y = u, x/y = v$, o, lo que es lo mismo, $x = \frac{vu}{1+v}, y = \frac{u}{1+v}$, cuyo jacobiano es, en valor absoluto, $\frac{u}{(1+v)^2}$. Entonces, resulta

$$P\{S \in A\} = \int_0^\infty \mathbf{1}_A(u)e^{-u}udu \int_0^\infty \frac{dv}{(1+v)^2} = \int_0^\infty \mathbf{1}_A(u)e^{-u}udu,$$

de modo que $dP_S(u) = e^{-u}u\mathbf{1}_{\{u>0\}}$. En otras palabras, S es una variable continua con función de densidad $f_S(u) = e^{-u}u\mathbf{1}_{\{u>0\}}$.

– Para todo $B \in \mathcal{B}(\mathbb{R})$,

$$P\{Q \in B\} = P_{(X,Y)}\{(x,y) \in \mathbb{R}^2 : x/y \in B\} = \int_0^\infty \int_0^\infty \mathbf{1}_B(x/y)e^{-x}e^{-y}dxdy.$$

Con el mismo cambio de variable que en el punto anterior, resulta

$$P\{Q \in B\} = \int_0^\infty \mathbf{1}_B(v) \int_0^\infty e^{-u}udu\frac{dv}{(1+v)^2} = \int_0^\infty \mathbf{1}_B(v)\frac{dv}{(1+v)^2},$$

y así vemos que $dP_Q(v) = \frac{1}{(1+v)^2}\mathbf{1}_{\{v>0\}}$. En otras palabras, Q es una variable continua con función de densidad $f_Q(v) = \frac{1}{(1+v)^2}\mathbf{1}_{\{v>0\}}$.

(2) Sean ahora, $A, B \in \mathcal{B}(\mathbb{R})$. Entonces

$$P_{(S,Q)}(A \times B) = P\{S \in A \wedge Q \in B\}$$

$$= P_{(X,Y)}\{(x,y) \in \mathbb{R}^2 : x + y \in A \wedge x/y \in B\} = \int_0^\infty \int_0^\infty \mathbf{1}_A(x+y)\mathbf{1}_B(x/y)e^{-x}e^{-y}dxdy.$$

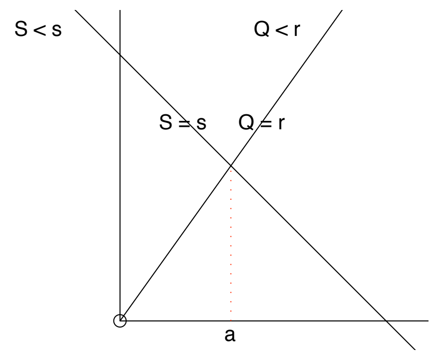
Con el mismo cambio de variable de los dos puntos anteriores obtenemos

$$P_{(S,Q)}(A \times B) = \left(\int_0^\infty \mathbf{1}_A(u)e^{-u}udu\right)\left(\int_0^\infty \mathbf{1}_B(v)\frac{dv}{(1+v)^2}\right) = P_S(A)P_S(B).$$

Queda así probado que S y Q son independientes.

- **Segunda forma**

(1) La densidad conjunta del vector (X,Y) tiene como soporte el cuadrante $C = \{x > 0,\ y > 0\}$, en el que $f_{(X,Y)}(x,y) = e^{-x-y}$, y las dos funciones S, Q toman sólo valores positivos.
Las funciones de distribución de S y de Q se hallan integrando $f_{(X,Y)}$ sobre la intersección de C con $\{S \le s\}$ o con $\{Q \le r\}$ respectivamente, cuyas líneas frontera podemos ver en el gráfico.



$$\forall s > 0, \quad P(S \le s) = \int_0^s \left\{ \int_0^{s-x} e^{-x-y}\,dy \right\} dx = \int_0^s \left(e^{-x} - e^{-s}\right) dx = 1 - e^{-s} - se^{-s},$$

$$\forall r > 0, \quad P(Q \le r) = \int_0^\infty \left\{ \int_{x/r}^{\infty} e^{-x-y}\,dy \right\} dx = \int_0^\infty e^{-x-x/r}\,dx = \frac{r}{1+r}\ ,$$

que son en efecto las integrales de sus derivadas, puesto que éstas existen y son continuas salvo en 0:

$$\boxed{f_S(s) = se^{-s} \ \text{ para } s > 0\ , \qquad f_Q(r) = 1/(1+r)^2 \ \text{ para } r > 0\ .}$$

(2) Para hallar la función de distribución conjunta basta integrar sobre la intersección de las dos regiones anteriores, cuyo vértice derecho cumple
$x = ry = s - y\ ,\ \ y = s/(r+1)\ ,\ \ x = rs/(r+1)$, y llamando $a = rs/(r+1)$ para s, r dados,

$$P(S \le s,\ Q \le r) = \int_0^a \left\{ \int_{x/r}^{s-x} e^{-x-y}\,dy \right\} dx = \int_0^a \left(e^{-x-x/r} - e^{-s}\right) dx$$

$$= \frac{r}{1+r}(1 - e^{-a-a/r}) - ae^{-s} = \frac{r}{1+r}(1 - e^{-s} - se^{-s}) = P(Q \le r)\ P(S \le s)$$

lo que prueba que S, Q son independientes.

---

2. Sean $\{X_j\}_{j\in\mathbb{N}}$ variables aleatorias independientes idénticamente distribuidas, con $E[X_j] = 0$ y $\mathrm{Var}[X_j] = \sigma^2$. Se definen, $\forall n \in \mathbb{N},\ S_n = \sum_{j=1}^{n} X_j$. Se pide

(1) Explicar, usando teoremas apropiados claramente enunciados, por qué existe el límite
$\lim_{n\to\infty} P\{|S_n|^2 < n\}$ y expresarlo de forma precisa.

(2) Explicar cómo tiene que ser σ para que el límite de (1) sea $> 1/2$.

Solución: (1) Por el **Teorema central del límite**,

$$\frac{S_n}{\sqrt{n}\sigma} \underset{n\to\infty}{\overset{d}{\to}} N(0,1),$$

es decir que, para todo $a > 0$,

$$P\{S_n^2 < n\sigma^2 a^2\} = P\left\{\frac{S_n^2}{n\sigma^2} < a^2\right\} = P\left\{\left|\frac{S_n}{\sqrt{n}\sigma}\right| < a\right\} \underset{n\to\infty}{\to} \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.$$

Tomando $a = \frac{1}{\sigma}$ resulta que

$$\lim_{n\to\infty} P\{S_n^2 < n\} = \int_{-\frac{1}{\sigma}}^{\frac{1}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1 - 2P\{Z > \frac{1}{\sigma}\},$$

siendo Z normal estándar.

(2)

$$1 - 2P\left\{Z > \frac{1}{\sigma}\right\} > \frac{1}{2} \Leftrightarrow P\left\{Z > \frac{1}{\sigma}\right\} < \frac{1}{4} \Leftarrow \frac{1}{\sigma} > 0.68 \Leftarrow \sigma < 1.47.$$

Por tanto, para que el límite de (1) sea $> 1/2$, es suficiente que sea $\sigma < 1.47$. Por otro lado $P\left\{Z > \frac{1}{\sigma}\right\} < \frac{1}{4} \Rightarrow \frac{1}{\sigma} > 0.67 \Rightarrow \sigma < 1.49$. Así que, para que el límite de (1) sea $> 1/2$, es necesario que $\sigma < 1.49$.

---

3. Sean, $\forall j \in \mathbb{N},\ X_j \geq 0$, variables aleatorias independientes, idénticamente distribuidas, (absolutamente) continuas con función de densidad de probabilidad f tal que
$\lim_{x\downarrow 0} f(x) = a > 0$. Consideramos $\forall n \in \mathbb{N},\ Y_n = \min\{X_1, X_2, \cdots, X_n\}$. Se pide

(1) Demostrar que $Y_n \to 0$ para $n \to \infty$, en probabilidad y también casi seguramente.

(2) Demostrar que $\forall x > 0,\ P\{nY_n > x\} \to e^{-ax}$ cuando $n \to \infty$.

Solución: (1)

$$P\{Y_n \geq \varepsilon\} = P\{X_1 \geq \varepsilon \wedge X_2 \geq \varepsilon \wedge \cdots \wedge X_n \geq \varepsilon\} = \prod_{j=1}^{n} P\{X_j \geq \varepsilon\} = \left(1 - \int_0^{\varepsilon} f(x)dx\right)^n$$
$$\leq (\text{para } \varepsilon \text{ pequeño}) \left(1 - \frac{a}{2}\varepsilon\right)^n \to 0, \text{ cuando } n \to \infty.$$

Así queda probado que $Y_n \xrightarrow{P} 0$ cuando $n \to \infty$.
Además, para ε como arriba,

$$\sum_{j=1}^{\infty} P\{Y_j \geq \varepsilon\} \leq \sum_{j=1}^{\infty} \left(1 - \frac{a}{2}\varepsilon\right)^j = \frac{1 - (a/2)\varepsilon}{(a/2)\varepsilon}.$$

Por el lema de Borel-Cantelli (¡el primero, que no necesita independencia!), para j grande $Y_j \leq \varepsilon$ c. s.. Como la unión de una colección numerable de conjuntos nulos es nulo, se sigue que $Y_n \to 0$ casi seguramente para $n \to \infty$.

(2)

$$P\{nY_n > x\} = \left(1 - \int_0^{\frac{x}{n}} f(y)dy\right)^n = e^{n\log\left(1-\int_0^{\frac{x}{n}} f(y)dy\right)}$$

$$\sim e^{-n\int_0^{\frac{x}{n}} f(y)dy} = e^{-x\frac{1}{x/n}\int_0^{x/n} f(y)dy} \to e^{-ax} \text{ cuando } n \to \infty.$$

---

4. Sea $\{S_n\}_{n\in\mathbb{N}}$ el paseo aleatorio definido por $P\{S_{n+1} - S_n = 1\} = p = 1 - P\{S_{n+1} - S_n = -1\}$, con $S_0 = 0$. Se pide

(1) Determinar para qué valores de p existe alguna constante $c > 1$ tal que las variables $X_n = c^{S_n}$ forman una martingala, hallando también c.

(2) Sea $m_n = \min\{x, P\{X_n \leq x\} \geq 1/2\}$ (la **mediana** de X_n). Dar una aproximación de m_n para n grande. *Indicación: usar el teorema central del límite para* S_n.

Solución: (1)

$$E[X_n|\mathcal{F}_{n-1}] = E[c^{S_n}|\mathcal{F}_{n-1}] = E[c^{S_{n-1}+S_n-S_{n-1}}|\mathcal{F}_{n-1}] = E[c^{S_{n-1}}c^{S_n-S_{n-1}}|\mathcal{F}_{n-1}]$$
$$= c^{S_{n-1}}E[c^{S_n-S_{n-1}}] = X_{n-1}E[c^{S_n-S_{n-1}}],$$

ya que $c^{S_{n-1}}$ es $\mathcal{F}_{n-1}-$medible y $c^{S_n-S_{n-1}}$ es independiente de $\mathcal{F}_{n-1}$.
Así pues, X es martingala si y sólo si $E[c^{S_n-S_{n-1}}] = 1$. Pero $E[c^{S_n-S_{n-1}}] = cp + c^{-1}(1-p)$, de modo que $E[c^{S_n-S_{n-1}}] = 1 \Leftrightarrow \left(c - \frac{1}{c}\right)p = 1 - \frac{1}{c} \Leftrightarrow (c^2-1)p = c - 1 \Leftrightarrow c + 1 = 1/p \Leftrightarrow c = \frac{1-p}{p}$.
Para que sea $c > 1$, habrá de ser $1 - p > p$, es decir $1 > 2p$ o $p < 1/2$.

(2) Como $X_n = c^{S_n}$ es función creciente de S_n , el suceso $\{X_n \leq x\}$ coincide con el $\{S_n \leq s\}$ si $x = c^s$. Por lo tanto basta responder la pregunta para la mediana a_n de S_n, ya que se tendrá $m_n = c^{a_n}$. La ventaja de usar S_n es el TCL: como S_n es una suma de v.a.i.i.d. con $E[S_n] = n\mu$, $\mu = 2p - 1 < 0$, $\mathrm{Var}S_n = n\sigma^2$, $\sigma^2 = 4p(1-p)$, se tiene que

$$\boxed{Z_n = (S_n - n\mu)/\sigma\sqrt{n} \ \ \textit{tienden en distribución a la } \text{Normal(0,1)} \ ,}$$

y en particular $\forall\varepsilon > 0$ hay un $h_\varepsilon > 0$ tal que $P\{S_n \leq n\mu \pm \varepsilon\sqrt{n}\} \to \frac{1}{2} \pm h_\varepsilon$; eso implica que $a_n - n\mu = o(\sqrt{n})$ y por lo tanto que

$\boxed{\log(m_n) \approx n(2p-1)\log(c) \to -\infty}$, con *errores relativos* que son $o(1/\sqrt{n})$.

Dicho de manera más clara aunque menos precisa:

m_n es aproximadamente $c^{E[S_n]} = b^n$, *donde* $b = 1/c^{1-2p} < 1 \Rightarrow b^n \to 0$.

% *No cabe esperar que sea $m_n/b^n \to 1$, porque la mediana a_n será siempre un entero, y en el mejor de los casos mantendrá errores absolutos < 1 respecto de $E[S_n]$, que suponen para m_n un factor*

$\in (1, c)$ arriba o abajo. Para poder probar si efectivamente se llega a ese caso óptimo no basta con aplicar el TCL, hay que apretar los tornillos de su prueba ...

---

5. Sea $\mathtt{a} = (\mathtt{a}_1, \mathtt{a}_2 \cdots, \mathtt{a}_L)$ una "palabra"de longitud L formada por caracteres de un alfabeto finito Σ y sean $\{X_j\}_{j\in\mathbb{N}}$ variables aleatorias independientes idénticamente distribuidas que toman valores en Σ con una distribución en la que cada elemento de Σ tiene probabilidad positiva. Se pide

(1) Si llamamos A_n al suceso de que se tenga $X_{n+i} = \mathtt{a}_i,\ \forall i \in \{1, \cdots, L\}$, probar que, casi seguramente, ocurren infinitos de los sucesos A_n, es decir, con probabilidad 1, la palabra $\mathtt{a}$ se escribe infinitas veces. *Idea: usar sólo los $n = kL,\ k = 0, 1, \ldots$.*

(2) Si la distribución de cada X_i es uniforme en Σ, es decir si hay N caracteres y cada uno tiene probabilidad $1/N$, hallar $p = P(A_n | X_{n+1} = \mathtt{a}_1)$, y deducir que el número esperado de caracteres hasta que llegue a completarse la palabra por primera vez es $> N^L$. *Idea: cada vez que se comienza la palabra es un "intento"y el "éxito"es completarla.*

Solución: (1) Para diferentes $k = 0, 1, \ldots$, los sucesos A_{kL} se definen en función de los valores de 'bloques disjuntos' de las variables independientes X_i, luego son sucesos independientes; por otro lado, cada A_n tiene $P(A_n) > 0$ (por tener cada $\mathtt{a}_i \in \Sigma$ probabilidad > 0), y todos ellos la misma (por ser las X_i igualmente distribuidas), luego

$$\sum_k P(A_{kL}) = \infty$$

y Borel-Cantelli implica que casi-seguramente ocurren infinitos de los A_{kL} (además de las veces que pueda aparecer *la palabra* en otras posiciones).

(2) Si ya se tiene $X_{n+1} = \mathtt{a}_1$, el *completar la palabra* con los siguientes $L-1$ caracteres tiene probabilidad $p = (1/N)^{L-1}$ si cada $\mathtt{a}_i$ tiene probabilidad $1/N$ de aparecer. Si vemos cada *comienzo de la palabra* como un 'intento', el número I de ellos que esperamos hacer hasta tener un 'éxito' es una variable Geométrica(p), con valor esperado $1/p = N^{L-1}$.

En la secuencia de caracteres hasta *completar la palabra* aparecerá el carácter $\mathtt{a}_1$ un número de veces $M \geq I$ y el número total de caracteres (sin contar los necesarios para *completar por fin la palabra*) será la suma

$$S = \sum_{j=1}^{M} Y_j$$

de variables i.i.d. Y_j que cuentan el número de caracteres hasta la siguiente aparición de $\mathtt{a}_1$, y que son por lo tanto Geométricas($1/N$), con valor esperado N. Luego $E[S] = E[M]N \geq N^L$.

En el caso, al que naturalmente NO nos referimos, de que fuese $L = 1$, todo lo dicho será bastante inútil y la desigualdad es una igualdad: *número esperado* $= N$. De lo contrario, los caracteres que *completan al fin la palabra* dan la desigualdad estricta.

Otra forma de llegar a esa desigualdad será condicionar la esperanza del número total T de caracteres al suceso A_0, aprovechando que A_0^c implica "malgastar al menos uno": $E[T] = E[T\,|A_0]\cdot 1/N^L + E[T\,|A_0^c]\cdot(1-1/N^L) \geq L/N^L + (1+E[T])\cdot(1-1/N^L)$.