Repaso: Más ejermplos sobre variables continuas. Resolvemos algunos ejercicios de la guía 6. Y un ejemplo de cambio de variable.

Pablo L. De Nápoli

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Probabilidades y Estadística para Matemática
Segundo cuatrimestre de 2021

Introducción

- En la clase de hoy repasaremos muchos de los conceptos que vimos hasta ahora a lo largo del curso.
- Al mismo tiempo, veremos algunos ejemplos más, y algunos conceptos nuevos (¡que obviamente no vamos a tomar en el parcial!).

Recordamos: Fórmulas para calcular esperanzas

Si X es una variable aleatoria y $\varphi : \mathbb{R} \to \mathbb{R}$ es una función vimos que

$$E[\varphi(X)] = \sum_k \varphi(x_k) \cdot p_k \text{ donde } p_k = P\{X = x_k\}$$

cuando X es una variable discreta que toma los valores (x_k).
En general para cualquier variable aleatoria

$$E[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x)\, dF(x)$$

cuando φ es continua, siendo $F = F_X$ la función de distribución acumulada de X.
Y en particular,

$$E[\varphi(X)] = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx$$

si X es una variable continua con densidad f.
Estas fórmulas se aplican con $\varphi(x) = x$ para calcular la esperanza $\mu_X = E[X]$ y con $\varphi(x) = (x - \mu_X)^2$ para calcular la varianza.

Recordamos: las distribuciones beta

En la clase 11, introdujimos la función beta de Euler,

$$B(\alpha_1, \alpha_2) = \int_0^1 (1-u)^{\alpha_1 - 1}\, u^{\alpha_2 - 1}\, du \quad \alpha_1, \alpha_2 > 0$$

y las distribuciones beta asociadas a esta función. Diremos que $X \sim \beta(\alpha_1, \alpha_2)$ si se distribuye según la densidad:

$$f_X(x) = \frac{1}{B(\alpha_1, \alpha_2)}\, x^{\alpha_1 - 1}\, (1-x)^{\alpha_2 - 1}\, I_{(0,1)}(x)$$

Allí probamos la fórmula que la relaciona con la función gama:

$$B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \quad \text{donde } \Gamma(\alpha) = \int_0^\infty x^{\alpha - 1}\, e^{-x}\, dx$$

Nos acordamos también de que Γ satisface la ecuación funcional $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$. Pregunta: ¿cuánto valen la esperanza y la varianza de X?

Recordamos: las distribuciones beta (2)

Más generalmente, podemos calcular los momentos de X.

$$\mu_k = E[X^k] = E[\varphi(X)] \ k \in \mathbb{N}$$

donde $\varphi(x) = x^k$. Entonces usando la ecuación funcional para la función gama:

$$\begin{aligned} E[X^k] &= \frac{1}{B(\alpha_1, \alpha_2)} \int_0^1 x^k \, x^{\alpha_1 - 1} \, (1-x)^{\alpha_2 - 1} \, dx \qquad = \frac{B(\alpha_1 + k, \alpha_2)}{B(\alpha_1, \alpha_2)} \\ &= \frac{\Gamma(\alpha_1 + k)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + k + \alpha_2)} \cdot \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \\ &= \frac{\alpha^{(k)}}{(\alpha_1 + \alpha_2)^{(k)}} = \prod_{r=0}^{k} \frac{\alpha_1 + r}{\alpha_1 + \alpha_2 + r} \end{aligned}$$

donde

$$\alpha^{(k)} = \alpha \cdot (\alpha + 1) \cdot \ldots \cdot (\alpha + k - 1)$$

se llama el símbolo de Pochhammer.

Recordamos: las distribuciones beta (3)

En particular, para $k = 1$ vemos que la esperanza vale,

$$E[X] = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

y el momento de segundo orden:

$$E[X^2] = \frac{\alpha_1(\alpha + 1)}{(\alpha_1 + \alpha_2) \cdot (\alpha_1 + \alpha_2 + 1)}$$

Finalmente

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}$$

Función generadora de Momentos

Definición

Si $X : \Omega \to \mathbb{R}$ *es una variable aleatoria, su función generadora de momentos o función generatriz de momentos de define por*

$$M_X(t) = E[e^{tX}]$$

(siempre que esta esperanza sea finita). Notamos que $M_X(0) = 1$.

Observación

Si $X : \Omega \to \mathbb{N}_0$ *es una variable aleatoria discreta con valores enteros no negativos,*

$$M_X(t) = \sum_{k=0}^{\infty} e^{kt} \cdot p_k = \sum_{k=0}^{\infty} (e^t)^k \cdot p_k = g_X(e^t)$$

donde $p_k = P\{X = x_k\}$ *y* g_X *es la función generatriz que definimos en la clase 6.*

Algunos ejemplos fáciles

- Supongamos que $X \sim \mathcal{U}(a, b)$ entonces:

$$M_X(t) = E[e^{tX}] = \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{e^{tb} - e^{ta}}{t(b-a)} \text{ para } t \neq 0$$

mientras que $M_X(0) = 1$.

- Supongamos que $X \sim \text{Exp}(\lambda)$ entonces:

$$M_X(t) = E[e^{tX}] = \int_0^\infty \lambda \, e^{tx} e^{-\lambda x} \, dx = \int_0^\infty \lambda \, e^{(t-\lambda)x} \, dx = \frac{\lambda}{\lambda - t}$$

siempre que $t < \lambda$.

Función generadora de Momentos (2)

Definición

Si $f : [0, \infty) \to \mathbb{R}$ *es una función, su transformada de Laplace se define por*

$$\mathcal{L}(f)(s) = \int_0^\infty f(x)\, e^{-sx}\, dx$$

$$Si \int_{-\infty}^{\infty} |f(x)|\, e^{-s_0 x}\, dx < +\infty$$

para algún $s_0 \in \mathbb{R}$ *esta integral existe para* $s > s_0$ *(o incluso para valores complejos de* s *con* $Re(s) > s_0$*.)*

Consideremos ahora el caso de una variable aleatoria X continua no negativa (o sea soportada en la semirrecta $[0, \infty)$), y con densidad de probabilidad $f(x)$. Entonces

$$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx}\, f(x)\, dx = \mathcal{L}(f)(-t) \text{ para } t < -s_0$$

¿Porqué se llama así?

Podemos calcular las derivadas de M_X derivando formalmente bajo el signo de esperanza:

$$M_X^{(k)}(t) = E[X^k e^{tX}]$$

con lo que los momentos de X se pueden calcular evaluando las derivadas de M_X en el origen,

$$\mu_k = E[X^k] = M_X^{(k)}(0)$$

Esta cuenta se puede justificar rigurosamente para $t > t_0$ usando teoremas de análisis real siempre que

$$E[|X|^k e^{t_0 X}] < +\infty.$$

para algún $t_0 = t_0(X) > 0$ (abscisa de convergencia). Bajo esta hipótesis, la función generatriz de momentos va a existir y ser derivable para $t < t_0$.

Un ejemplo

Supongamos que $X \sim \text{Exp}(\lambda)$. Vimos que

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

Las derivadas son

$$M_X^{(k)}(t) = k! \frac{\lambda}{(\lambda - t)^{k+1}}$$

Luego los momentos de la distribución exponencial son

$$\mu_k = E[X^k] = M_X^{(k)}(0) = \frac{k!}{\lambda^k}$$

En particular X tiene esperanza $1/\lambda$ y varianza

$$\text{Var}(X) = E[X^2] - E[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

como vimos antes.

Otro ejemplo

Consideremos una vez más las distribuciones gama que introdujimos en la clase 8. Si $X \sim \Gamma(\alpha, \lambda)$

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \, x^{\alpha-1} \, e^{-\lambda x} I_{(0,+\infty)}(x)$$

[Aqui $\alpha, \lambda > 0$]. Su función generatriz de momentos es

$$\begin{aligned} M_X(t) = E[e^{tX}] &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{tx} \, x^{\alpha-1} \, e^{-\lambda x} \, dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{(t-\lambda)x} \, x^{\alpha-1} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(t-\lambda)^\alpha} = \left(\frac{\lambda}{\lambda - t} \right)^\alpha \quad \text{siempre que } t < t_0 = \lambda \end{aligned}$$

Como $\Gamma(1, \lambda) = \text{Exp}(\lambda)$, este ejemplo generaliza el de la distribución exponencial que vimos antes.

Funciones generatrices de momentos e independencia

Proposición

Sean X e Y variables aleatorias independientes, entonces

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

si $t < \min(t_0(X), t_0(Y))$

Demostración.

Como X e Y son idependientes, e^{tX} y e^{tY} también lo serán, luego:

$$\begin{aligned} M_{X+Y}(t) &= E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}] \\ &= E[e^{tX}] \cdot E[e^{tY}] = M_X(t) \cdot M_Y(t) \end{aligned}$$

$\square$

Esto corresponde a la siguiente propiedad de la transformada de Laplace: si $f, g : [0, \infty) \to \mathbb{R}$

$$\mathcal{L}(f * g)(s) = \mathcal{L}(f)(s) \cdot \mathcal{L}(g)(s)$$

El teorema de Lerch

Para el siguiente ejemplo, necesitamos un teorema sobre la transformada de Laplace:

Teorema

La transformada de Laplace es inyectiva. Con más precisión si $f : [0, \infty) \to \mathbb{R}$ *es una función continua, tal que*

$$\int_{-\infty}^{\infty} |f(x)| \, e^{-s_0 x} \, dx < +\infty$$

para algún s_0*,*

$$\mathcal{L}(f)(s) = 0$$

para $s > s_0$ *entonces* $f \equiv 0$*.*

Corolario

La función generatriz de momentos de una variable continua soportada en $[0, \infty)$ *determina su distribución.*

Un ejemplo

En la clase 11 demostramos (a mano) el siguiente resultado:

Proposición

Si $X \sim \Gamma(\alpha_1, \lambda)$, $Y \sim \Gamma(\alpha_2, \lambda)$ y son independientes, entonces $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$.

Ahora veremos otra forma de demostrarlo usando la función generatriz de momentos. Como X e Y son independientes, si $< \lambda$:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = \left(\frac{t}{\lambda - t}\right)^{\alpha_1} \cdot \left(\frac{t}{\lambda - t}\right)^{\alpha_2}$$
$$= \left(\frac{t}{\lambda - t}\right)^{\alpha_1+\alpha_2}$$

Pero entoces como esta es la función generatriz de momentos de la distribución $\Gamma(\alpha_1 + \alpha_2, \lambda)$, por el corolario del teorema de Lerch $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$.

Estadísticos de orden

Ejercicio 8 de la práctica 6, item a)

Dadas $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con función de distribución acumulada F, se definen sus estadísticos de orden $X^{(1)}, \ldots, X^{(n)}$ como aquellas variables aleatorias que se obtienen ordenando las X_i de manera creciente. En particular, tenemos que

$$X^{(1)} = \min_{1 \leq i \leq n} X_i$$

$$X^{(n)} = \max_{1 \leq i \leq n} X_i.$$

Hallar para cada $k = 1, \ldots, n$ la función de distribución acumulada de $X^{(k)}$ en términos de F.

En estadística, cuando $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas con función de distribución acumulada F, decimos que tenemos una muestra aleatoria de tamaño n de la distribución F (con reposición).

Distribución del máximo

Empezemos mirando el máximo $X^{(n)}$. Dado $x \in \mathbb{R}$, será $X^{(n)} \leq x$ si $X_i \leq x$ para todo i. De modo que

$$P\{X^{(n)} \leq x\} = \prod_{i=1}^{n} P\{X_i \leq x\} \text{ por independencia}$$

o sea

$$F_{X^{(n)}}(x) = \prod_{i=1}^{n} F_{X_i}(x) = F^n(x)$$

al ser las X_i idénticamente distribuidas.

Distribución del mínimo

Similarmente miremos el máximo $X^{(1)}$. Dado $x \in \mathbb{R}$, será $X^{(1)} \leq x$ si y $X_i \leq x$ para algún i.
Queremos hallar $F_{X^{(1)}}(x) = P\{X^{(1)} \leq x\}$. Es más fácil mirar la probabilidad complementaria: Nuevamente como las variables son independientes,

$$
\begin{aligned}
F_{X^{(1)}}(x) &= 1 - P\{X^{(1)} > x\} \\
&= 1 - P\{X_i > x \text{ para todo } i\} \\
&= 1 - \prod_{i=1}^{n} P\{X_i > x\} \quad \text{por independencia} \\
&= 1 - \prod_{i=1}^{n} [1 - F_{X_i}(x)] \\
&= 1 - [1 - F(x)]^n
\end{aligned}
$$

Distribución de los estadísticos de orden

Consideremos ahora uno cualquiera de los estadísticos de orden $X^{(k)}$ y, dado un $x \in \mathbb{R}$, preguntémosnos cuando $X^{(k)}$ eso significa que tenemos k observaciones que son menores o iguales que x.
Definimos las variables

$$Z_i = \begin{cases} 1 & \text{si} \quad X_i \leq x \\ 0 & \text{si} \quad X_i > x \end{cases}$$

Vemos que son variables de Bernoulli con probabilidad de éxito $p = F(x)$. Son independientes porque las X_i lo eran.
La variable aleatoria

$$N = \sum_{i=1}^{n} Z_i$$

representa el número total de observaciones X_i que son menores o iguales que x.
Notamos que

$$N \sim \text{Bi}(n, p)$$

Distribución de los estadísticos de orden(2)

Entonces

$$P\{X^{(k)} \leq x\} = P\{N \geq k\} = \sum_{j=k}^{n} b(j,n,p)$$

donde

$$b(j,n,p) = \binom{n}{j} p^j q^{n-j} \quad q = 1 - p$$

O sea:

Solución

$$F_{X^{(k)}}(x) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j [1 - F(x)]^{n-j}$$

Un ejemplo

Ejercicio 9, práctica 6

Sean $X_1, \ldots, X_n$ variables aleatorias independientes con distribución exponencial de parámetros $\alpha_1, \ldots, \alpha_n$ respectivamente. Mostrar que la distribución de $X^{(1)}$ es exponencial. ¿De qué parámetro?

Solución: Recordamos que para una distribución exponencial $\mathrm{Exp}(\alpha)$

$$F(x) = \int_0^x \alpha e^{-\alpha x}\, dx = 1 - e^{-\alpha x}$$

Entonces

$$\begin{aligned} F_{X^{(1)}} = 1 - P\{X^{(1)} > x\} &= 1 - P\{X_i > x \text{ para todo } i\} \\ &= 1 - \prod_{i=1}^{n} P\{X_i > x\} = 1 - \prod_{i=1}^{n} [1 - F_{X_i}(x)] \\ &= 1 - \prod_{i=1}^{n} e^{-\alpha_i x} = 1 - e^{-sx} \end{aligned}$$

donde $s = \alpha_1 + \alpha_2 + \ldots + \alpha_n$. Luego $X^{(1)} \sim \mathrm{Exp}(s)$.

Densidad de los estadísticos de orden

Teorema

Si $X_1, X_2, \ldots, X_n$ *son variables continuas independientes identicamente distribuidas con densidad* f *y función de distribución acumulada*

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

entonces los estadísticos de orden $X^{(k)}$ *también son variables continuas con la densidad*

$$f_{X^{(k)}}(x) = c_k\, [F(k)]^{k-1} (1 - F(x))^{n-k} f(x)$$

donde

$$c_k = \frac{m!}{(k-1)!(n-k)!} = n \binom{n-1}{k-1} = k \binom{n}{k}$$

Idea de la Demostración

Antes vimos que

$$F_{X^{(k)}}(x) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j [1 - F(x)]^{n-j}$$

Derivando

$$f_{X^{(k)}} = \sum_{j=k}^{n} \binom{n}{j} \left[jF(x)^{j-1}[1 - F(x)]^{n-j} - (n-j)F(x)^j[1 - F(x)]^{n-j-1} \right] f(x)$$

Pero

$$j\binom{n}{j} = j \cdot \frac{n!}{j!(n-j)!} = n \cdot \frac{(n-1)!}{(j-1)!(n-j)!} = n\binom{n-1}{j-1} = (n-j)\binom{n-1}{j}$$

¡Entonces la suma es telescópica y sólo sobrevive un término! (les dejo terminar la cuenta como ejercicio)

Las densidades beta como estadísticos de orden de la uniforme

Ejercicio 8, item d)

Probar que si las X_i tienen distribución uniforme en el intervalo $[0,1]$ entonces para cada $k = 1, \ldots, n$ la variable aleatoria $X^{(k)}$ tiene distribución $\beta(k, n-k+1)$.

Solución: Antes vimos que

$$f_{X^{(k)}}(x) = c_k \, [F(k)]^{k-1} \, (1 - F(x))^{n-k} \, f(x)$$

Para la distribución uniforme si $x \in (0,1)$, $f(x) = 1$, $F(x) = x$, entonces

$$f_{X^{(k)}}(x) = c_k \; x^{k-1} \, (1-x)^{n-k}$$

Por lo que vemos que $X^{(k)} \sim \beta(k, n-k+1)$.

Otro ejercicio sobre estadísticos de orden, para comparar

Ejercicio 15 de la práctica 6, item a)

Sean $X_1, \ldots, X_n$ variables aleatorias absolutamente continuas, independientes e idénticamente distribuidas con función de densidad f y consideremos el vector aleatorio $\overline{X} = (X^{(1)}, \ldots, X^{(n)})$ conformado por sus estadísticos de orden. Mostrar que $\overline{X}$ es absolutamente continuo y que su función de densidad viene dada por

$$f_{\overline{X}}(x) = n! \prod_{i=1}^{n} f(x_i) I_{\{x : x_1 < \cdots < x_n\}}(x).$$

Este ejercicio no lo voy a resolver porque lo resolvieron en la última clase práctica.

Un ejercicio de cambio de variable

Ejercicio

Se tienen dos variables aleatorias independientes $U, V \sim \mathcal{U}(0,1)$. A partir de ellas se definen las variables aleatorias R y W :

$$R = \sqrt{-2\log U}, \quad W = 2\pi V$$

y

$$X = R \cdot \cos W, \quad Y = R \cdot \operatorname{sen} W$$

Caracterizar la distribución del vector (X, Y).

Solución

Notamos que R toma valores en $(0,+\infty)$ y W en $(0,2\pi)$
Para la primera parte consideramos el cambio de variable

$$(R,W)=\varphi_1(U,V) \text{ donde } \varphi_1:\Omega_1=(0,1)\times(0,1)\to\Omega_2=(0,+\infty)\times(0,2\pi)$$

dado por $\varphi_1(u,v)=(\sqrt{-2\log r},2\pi v)$. Este cambio de variable es biyectivo y su inversa $\varphi_1^{-1}:\Omega_2\to\Omega_1$ es

$$\varphi_1^{-1}(r,w)=\left(e^{-r^2/2},\frac{w}{2\pi}\right)$$

Para encontrarla, obervé que:

$$r=\sqrt{-2\log u}\Leftrightarrow r^2=-2\ \log r\Leftrightarrow -\frac{r^2}{2}=\log u\Leftrightarrow u=e^{-r^2/2}$$

$$w=2\pi v\Leftrightarrow v=\frac{w}{2\pi}$$

Además observamos que

$$r\in(0,1)\Leftrightarrow w\in(0,\infty), v\in(0,1)\Leftrightarrow w\in(0,2\pi)$$

¡Esta cuenta es fácil porque las varialbes no se mezclan!

Solución (2)

Entonces según el teorema de cambio de variable

$$f_{(R,W)}(r,w) = f_{(U,V)}(\varphi^{-1}(r,w)) \cdot |\det D(\varphi_1^{-1})(r,w)|$$

Pero

$$f_{(U,V)}(\varphi_1^{-1}(r,w)) = I_{\Omega_2}(z,w) = I_{(0,\infty)}(r) \cdot I_{(0,2\pi)}(w)$$

El jacobiano es:

$$|\det D(\varphi^{-1})(z,w)| = \left|\det \begin{pmatrix} re^{-r^2} & 0 \\ 0 & \frac{1}{2\pi} \end{pmatrix}\right| = \frac{1}{2\pi} z e^{-r^2/2}$$

Luego

$$f_{(R,W)}(r,w) = \frac{1}{2\pi} r e^{-r^2/2} I_{(0,\infty)}(r) \cdot I_{(0,2\pi)}(w)$$

Notamos que R y W son independientes. $W \sim \mathcal{U}(0,2\pi)$ mientras que $R \sim \chi_2$ (una de las distribuciones que introdujimos en la clase 11).

Solución (3)

Ahora hacemos un nuevo cambio de variable $\varphi_2 : \Omega_2 \to \Omega_3 = \mathbb{R}^2$ dado por

$$(x, y) = \varphi_2(r, w) = (r \cos w, r \operatorname{sen} w)$$

Este cambio de variable lo conocemos bien: es el cambio de variables polares. Sabemos que su jacobiano es r, y que podemos hacerlo biyectivo quitando un conjunto de área cero. Entonces, el teorema de cambio de variable se aplica también. Además

$$r^2 = x^2 + y^2$$

Y como $\det(D\varphi)(r, w) = r \Rightarrow \det(D\varphi^{-1}) = \frac{1}{r}$. Encontramos que:

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi} r e^{-r^2/2} \cdot \frac{1}{r} = e^{-r^2/2} = \frac{1}{2\pi} e^{(x^2+y^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Como ya vimos, esto significa que X e Y con variables con distribución normal estándar independientes.

Interés de este ejercicio

- Este ejercicio proporciona un método para simular en la computadora la distribución normal, a partir de un generador de números pseudo-aleatorios que simula la distribución uniforme.
- La cuenta del ejercicio es la misma que la que se hace en análisis 2 para calcular el área bajo la curva normal.