Vectores Aleatorios: Cambios 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

Parte I

Vectores aleatorios n-dimensionales

Vectores aleatorios n-dimensionales

Las ideas que vimos en la clase pasada se generalizan sin dificultad a vectores aleatorios multidimensionales, pero la notación resulta más complicada. Así pues si $X : \Omega \to \mathbb{R}^n$ es un vector aleatorio n-dimensional, que se distribuye según una densidad conjunta $f(x) = f(x_1, x_2, \ldots, x_n)$ que supongremos por simplicidad continua, tendremos que:

- La esperanza de una función $\varphi(X)$ del vector X, donde $\varphi : \mathbb{R}^n \to \mathbb{R}$ es una función continua, se puede calcular mediante la fórmula:

$$E[\varphi(X)] = \int_{\mathbb{R}^n} \varphi(x) f(x) \, dx$$

- La k-ésima componente X_k del vector X $(1 \leq k \leq n)$ se distribuye según la densidad marginal:

$$f_{X_k}(x) = \int_{\mathbb{R}^{n-1}} f(x_1, x_2, \ldots, x_{k-1}, x, x_{k+1}, \ldots, x_n) \, d\hat{x}_k$$

donde

$$d\hat{x}_k = dx_1 \, dx_2 \, \ldots \, dx_{k-1} \, dx_{k+1} \ldots dx_n$$

Vectores aleatorios n-dimensionales (2)

- Las componentes X_1, X_2, $\ldots X_n$ del vector X se dirán mutuamente independientes si para cualquier rectángulo n-dimensional (producto de intervalos)

$$I = \prod_{k=1}^{n} (a_k, b_n]$$

se verifica que:

$$P\{X \in I\} = \prod_{k=1}^{n} P\{a_k < X_k \leq b_k\}$$

En términos de la función de distribución conjunta, $X_1, X_2, \ldots, X_n$ son mutuamente independientes si y sólo si $f(x)$ se factoriza en la forma:

$$f(x) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

Recordamos de la clase 9:Cambios de variable unidimensionales

Supongamos que X es una variable aleatoria con función de distribución acumuladaF_X, y sea $Y = \varphi(X)$ donde $\varphi : U \to V$ es alguna función creciente siendo U y V intervalos de $\overline{\mathbb{R}}$, biyectiva y con una inversa $\varphi^{-1} : V \to U$ también creciente.

$$F_Y(y) = P\{Y \leq y\} = P\{\varphi(X) \leq y\} = P\{X \leq \varphi^{-1}(y)\} = F_X(\varphi^{-1}(y))$$

En particular si X es una variable continua, derivando vemos que

$$f_Y(y) = f_X(\varphi^{-1}(y)) \cdot (\varphi{-}1)'(y) = \frac{1}{\varphi(x)} f_Y(x) \text{ donde } x = \varphi^{-1}(y)$$

El caso más simple e importante, es el de un cambio de variable lineal dado por $\varphi(x) = ax + b$ con $a > 0$. Entonces $\varphi^{-1}(y) = \frac{y-b}{a}$.
Luego si $Y = aX + b$ entontramos que:

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Un ejemplo de un cambio de variable no monótono

Recordamos también que la situación es bastante más compleja si admitimos cambios de variables que no son monótonos o biyectivos.

Allí consideremos por ejemplo el cambio de variable $Y = X^2$, y encontramos que

$$f_Y(y) = \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right] \quad (y > 0)$$

Cambios de variables n-dimensionales

Proposición

Supongamos que X es una vector que se distribuye según una densidad $f(x)$ con soporte en $\overline{U}$ siendo U un abierto $\mathbb{R}^n$, y que $\varphi : U \to V$ es un difeomorfismo C^1, donde V es otro abierto de $\mathbb{R}^n$ entonces, si consideramos el vector aleatorio $Y = \varphi(X)$, Y se distribuye en V según la densidad

$$f(\varphi^{-1}(y))|det(D\varphi^{-1})(y)|$$

Demostración

Sea $W \subset V$ un abierto cualquiera, entonces

$$P\{Y \in W\} = P\{X \in \varphi^{-1}(W)\} = \int_{\varphi^{-1}(W)} f(x)dx$$

En esta integral, hagamos el cambio de variable $y = \varphi(x)$, $x = \varphi^{-1}(y)$. Entonces, según el teorema de cambio de variable

$$P\{Y \in W\} = \int_W f(\varphi^{-1}(y))|\det D(\varphi^{-1})(y)|dy$$

Como esto vale para todo $W \subset V$, concluimos que Y se distribuye en V según la densidad $f(\varphi^{-1}(y))|\det(D\varphi^{-1})(y))|$.

Parte II

Suma de variables aleatorias independientes

Suma de variables aleatorias independientes

Definición

Sean $f, g : \mathbb{R} \to \mathbb{R}$ *funciones integrables. Definimos su* **convolución** $f * g$ *de la siguiente manera:*

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

Como ejemplo de la aplicación del teorema de cambio de variable, demostramos la siguiente afirmación:

Proposición

Supongamos que X *e* Y *son variables aleatorias continuas independientes, que se distribuyen en* $\mathbb{R}$ *según las densidades* $f(x)$ *y* $g(x)$ *respectivamente, entonces* $X + Y$ *se distribuye según la densidad* $f * g(x)$.

Demostración

Como X e Y son independientes,

$$(X, Y) \sim f(x)g(y)$$

Hacemos el cambio de variable lineal $(U, V) = \varphi(X, Y) = (X + Y, Y)$. Entonces $(X, Y) = \varphi^{-1}(U, V) = (U - V, V)$. Como φ es una transformación lineal, su diferencial coincide con ella misma. Para calcular el determinante de φ observamos que su matriz en la base canónica de $\mathbb{R}^2$ es:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

En consecuencia, el determinante de φ es 1. Por el teorema anterior, tenemos que (U, V) que:

$$(U, V) \sim f(u - v)g(v) \text{ (densidad conjunta)}$$

Para recuperar la densidad de U (densidad marginal) debemos integrar en la variable v:

$$U \sim \int_{-\infty}^{\infty} f(u - v)g(v)\, dv$$

Algunas Observaciones sobre la convolución

1. La convolución es conmutativa:

$$f * g = g * f$$

   También es posible probar que es asociativa:

$$(f * g) * h = f * (g * h)$$

2. Si f y g son densidades de probabilidad, entonces $f * g$ también lo es.
3. Si f y g están soportadas en la semirrecta $[0, +\infty)$ (es decir: $f(t) = g(t) = 0$ si $t < 0$), entonces:

$$(f * g)(x) = \int_0^x f(t)\, g(x - t)\, dt$$

Suma de variables normales independientes

Proposición

Si $X \sim N(0, \sigma_1^2)$ *e* $Y \sim N(0, \sigma_2^2)$ *son variables aleatorias independientes, entonces* $X + Y \sim N(0, \sigma_1^2 + \sigma_2^2)$

Demostracion (ver detalles en mi apunte)

Aplicamos la proposción anterior con

$$f(x) = \frac{1}{\sigma_1\sqrt{2\pi}}\ e^{-x^2/(2\sigma_1^2)}, \quad g(x) = \frac{1}{\sigma_2\sqrt{2\pi}}\ e^{-x^2/(2\sigma_2^2)}$$

Entonces $X \sim f * g$, donde

$$\begin{aligned}(f * g)(x) &= \int_{-\infty}^{\infty} \frac{1}{\sigma_1\sqrt{2\pi}}\ e^{-t^2/(2\sigma_1^2)} \frac{1}{\sigma_2\sqrt{2\pi}}\ e^{-(x-t)^2/(2\sigma_2^2)}\ dt \\ &= \frac{1}{\sigma_1\sigma_2 2\pi} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}A(x,t)\right\}\ dt\end{aligned}$$

donde

$$A(x,t) := \frac{t^2}{\sigma_1^2} + \frac{(x-t)^2}{\sigma_2^2}$$

Demostracion (2)

Reescribimos esta expresión, completando el cuadrado:

$$A(x,t) = \frac{\sigma^2}{\sigma_1^2\sigma_2^2}\left(t - x\frac{\sigma_1^2}{\sigma^2}\right)^2 + \frac{1}{\sigma^2}x^2$$

donde $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Sustituyendo y usando las propiedades de la exponencial nos queda

$$(f * g)(x) = \frac{1}{\sigma_1\sigma_2 2\pi}\exp\left(-\frac{x^2}{2\sigma^2}\right)\int_{-\infty}^{\infty}\exp\left\{-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2}\left(t - x\frac{\sigma_1^2}{\sigma^2}\right)^2\right\}\,dt$$

Demostracion (3)

Sólo nos falta pues calcular la integral,

$$I(x) = \int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2}\left(t - x\frac{\sigma_1^2}{\sigma^2}\right)^2\right\} dt$$

pero haciendo el cambio de variable

$$u = t - x\frac{\sigma_1^2}{\sigma^2}$$

vemos que no depende en realidad de x, y es

$$I(x) = \int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2}u^2\right\} du$$

Demostracion (5)

Y haciendo un último cambio de variable

$$v = \frac{\sigma}{\sigma_1 \sigma 2} u$$

nos queda que

$$I(x) = \frac{\sigma_1 \sigma_2}{\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{v^2}{2}\right\} dv = \sqrt{2\pi} \frac{\sigma_1 \sigma_2}{\sigma}$$

Reemplazando nos queda que

$$X + Y \sim (f * g)(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

Es decir, que $X + Y \sim N(0, \sigma^2)$.

Otro Ejemplo: Suma de variables independientes con distribución gama

Recordamos las densidades gama $\Gamma(\alpha, \lambda)$

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

Proposición

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

Demostración

$X + Y \sim f_{\alpha_1,\lambda} * f_{\alpha_2,\lambda}$. Hemos de calcular esta convolución:

$$
\begin{aligned}
(f_{\alpha_1,\lambda} * f_{\alpha_2,\lambda})(x) &= \int_0^x \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} \, (x-t)^{\alpha_1 - 1} \, e^{-\lambda(x-t)} \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} \, t^{\alpha_2 - 1} \, e^{-\lambda t} \, dt \\
&= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \left(\int_0^x (x-t)^{\alpha_1-1} t^{\alpha_2-1} dt \right) e^{-\lambda x}
\end{aligned}
$$

Demostración (2)

En esta integral hacemos el cambio de variable $u = t/x$ $(0 \leq u \leq 1)$. Entonces:

$$(f_{\alpha_1,\lambda} * f_{\alpha_2,\lambda})(x) = \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1))\Gamma(\alpha_2)} \left(\int_0^1 (x - xu)^{\alpha_1 - 1} \, (xu)^{\alpha_2 - 1} \, x \, du \right) \, e^{-\lambda x}$$

$$= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1+\alpha_2-1} \left(\int_0^1 (1 - u)^{\alpha_1 - 1} \, u^{\alpha_2 - 1} \, du \right) e^{-\lambda x}$$

$$= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \, B(\alpha_1, \alpha_2) \, x^{\alpha_1+\alpha_2-1} \, e^{-\lambda x}$$

donde

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

se llama la función Beta de Euler.

Demostración (3)

Notamos que esta es salvo la constante, la densidad gama $f_{\alpha_1+\alpha_2,\lambda}$, pero como la convolución de dos densidades de probabilidad es una densidad de probabilidad, y hay una única constante que hace que la integral sobre $(0, +\infty)$ dé 1 deducimos que:

$$f_{\alpha_1,\lambda} * f_{\alpha_2,\lambda} = f_{\alpha_1,\alpha_2,\lambda}$$

Como subproducto de la demostración obtenemos que:

$$\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} B(\alpha_1, \alpha_2) = \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}$$

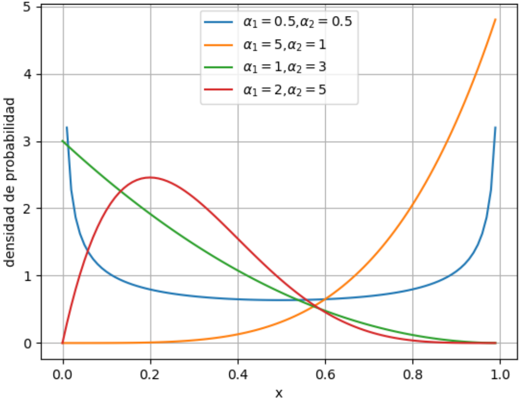
o sea

$$B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$$

Densidades beta

La función beta también puede usarse para definir una familia de distribuciones: las distribuciones beta. Diremos que $X \sim \beta(\alpha_1, \alpha_2)$ si se distribuye según la densidad:

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



Parte III

Algunas densidades útiles en estadística

Las densidades χ^2

Recordamos de la clase 9 que si $X \sim N(0,1)$ una variable aleatoria con distribución normal estándar $Y = X^2$ se distribuye según la densidad $\chi_1^2 = \Gamma(1/2, 1/2)$

Sean ahora $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con distribución normal estándar, y consideremos la variable aleatoria

$$Y_n = X_1^2 + X_2^2 + \ldots + X_n^2$$

¿cuál es la distribución de Y_n ? Por lo anterior, la densidad de Y_n será (por la independencia) la convolución de la densidad $\Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$ n veces con sigo misma, que por el resultado anterior es $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$. Es decir, que la densidad de Z_n será

Densidad χ_n^2

$$f_{Y_n}(y) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} y^{n/2-1} e^{-y/2} \quad (y > 0)$$

Esta densidad se conoce como densidad χ^2 con n grados de libertad. Las fórmulas para la esperanza y variancia de las distribuciones gama nos dan que

$$E[Y_n] = n, \quad \mathrm{Var}[Y_n] = 2n$$

Gráfico de la densidad χ^2_n

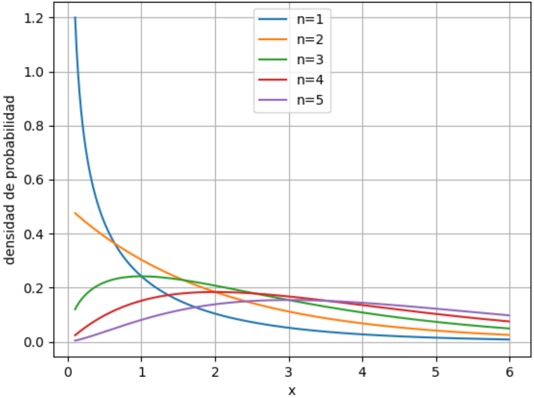
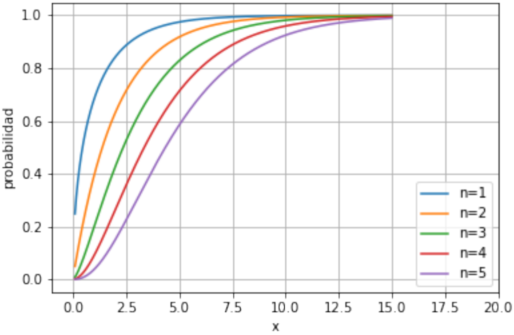


Gráfico de la distribución acumulada de una χ^2_n



Las densidades χ_n

Si consideramos el vector aleatorio $X = (X_1, X_2, \ldots, X_n)$ con las $X_i \sim N(0,1)$ independientes,

$$Z_n = \|X\| = \sqrt{Y_n}$$

entonces

$$\begin{aligned} f_{Z_n}(z) &= 2z \frac{(1/2)^{n/2}}{\Gamma(n/2)} \, (z^2)^{n/2-1} \cdot e^{-z^2/2} \\ &= \frac{2(1/2)^{n/2}}{\Gamma(n/2)} z^{n-1} \cdot e^{-z^2/2} \quad (z > 0) \end{aligned}$$

Esta distribución se llama χ_n. Con $n = 3$ esta distribución aparece en física, como la distribución de Maxwell-Boltzmann, que es la distribución de probabilidad de las velocidades de un gas asociada a la estadística de Maxwell-Boltzmann para dicho sistema.

$$f(v) = 4\pi \left(\frac{m}{2\pi k T} \right)^{\frac{3}{2}} v^2 e^{\frac{-mv^2}{2kT}}$$

donde m es la masa de la partícula, T es la temperatura absoluta y k es una constante (constante de Boltzmann).

Densidad del cociente de dos variables aleatorias independientes

Supongamos que X e Y son variables aleatorias continuas independientes, con densidades f_X y f_Y respectivamente. Supongamos además que Y está concentrada en la semirrecta positiva $(0, +\infty)$. Quremos calcular la densidad del cociente $T = X/Y$.

La densidad conjunta del vector aleatorio (X, Y) será $f_X(x)f_Y(y)$ como consecuencia de independencia de las variables X e Y.

Consideramos ahora el cambio de variable $(T, V) = \varphi(X, Y)$ donde donde

$$(u, v) = \varphi(x, y) = (x/y, y)$$

entonces la función inversa será

$$(x, y) = \varphi^{-1}(t, v) = (tv, v)$$

Y la diferencial de φ^{-1} es

$$D\varphi^{-1}(t, v) = \begin{pmatrix} v & t \\ 0 & 1 \end{pmatrix}$$

de modo que el Jacobiano es v.

Densidad del cociente de dos variables aleatorias independientes (2)

De acuerdo al teorema de cambio de variable, encontramos que el vector (T, V) se distribuye según la densidad conjunta

$$f_X(tv)f_Y(v)v$$

e integrando respecto la variable v podemos recuperar la densidad (marginal) de T que resulta ser:

$$T \simeq \int_0^\infty f_X(tv)f_Y(v)v \, dv \tag{1}$$

La densidad t de Student

Sea X una variable aleatoria con distribución χ^2 con n grados de libertad, Y una variable aleatoria con distribución normal estándar y supongamos que X e Y son independientes. Queremos calcular la densidad de la variable aleatoria

$$T = \frac{\sqrt{\frac{X}{n}}}{Y}$$

[El porqué esta variable aleatoria es interesante, lo veremos más adelante al desarrollar conceptos de estadística]

Ya vimos que la densidad de X es la χ^2_n. Consideramos $\varphi : (0, +\infty) \to (0, +\infty)$ dada por

$$\varphi(x) = \sqrt{\frac{x}{n}}$$

es un difeomorfismo cuya inversa es $\varphi^{-1}(y) = ny^2$.

La densidad t de Student (2)

Aplicando la fórmula de cambio de variables, encontramos que la densidad de $U = \sqrt{\frac{X}{n}}$ es

$$\begin{aligned} f_Y(y) &= \frac{(1/2)^{n/2}}{\Gamma(n/2)} \, (ny^2)^{n/2-1} e^{-ny^2/2} 2ny \; I_{(0,+\infty)}(y) \\ &= \frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)} \, y^{n-1} \; e^{-ny^2/2} \; I_{(0,+\infty)}(y) \end{aligned}$$

La densidad t de Student (3)

Utilizando la fórmula para la densidad del cociente de dos variables aleatorias independientes, vemos que T se distribuye según la densidad

$$\begin{aligned} f_T(t) &= \int_0^\infty f_X(tv) f_Y(v) v \, dv \\ &= \frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)\sqrt{2\pi}} \int_0^\infty e^{-t^2v^2/2} \, v^{n-1} \, e^{-nv^2/2} \, v \, dv \\ &= \frac{2^{(1-n)/2} n^{n/2}}{\Gamma(n/2)\sqrt{\pi}} \int_0^\infty e^{-(t^2+n)v^2/2} \, v^n \, dv \quad (t > 0) \end{aligned}$$

La densidad t de Student (4)

Hacemos el cambio de variable $x = \frac{v^2}{2}(t^2+n)$, entonces esta integral se transforma en

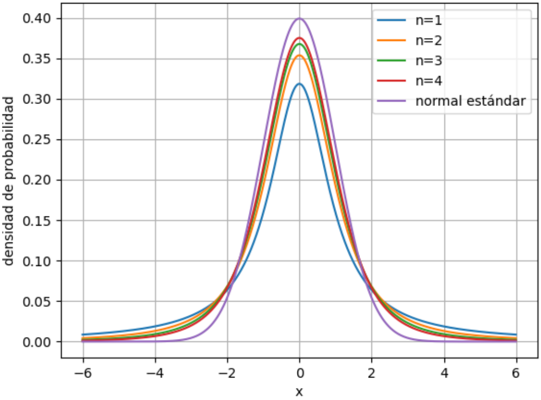
$$
\begin{aligned}
f_T(t) &= \frac{2^{(1-n)/2} n^{n/2}}{\Gamma(n/2)\sqrt{\pi}} \frac{1}{n+t^2} \int_0^\infty e^{-x} \left(\frac{2x}{n+t^2}\right)^{(n-1)/2} dx \\
&= \frac{n^{n/2}}{\Gamma(n/2)\sqrt{\pi}} \frac{1}{(n+t^2)^{(n+1)/2}} \int_0^\infty e^{-x} \, x^{(n-1)/2} \, dx \\
&= \frac{n^{n/2}}{\Gamma(n/2)\sqrt{\pi}} \Gamma\left(\frac{n+1}{2}\right) \frac{1}{(n+t^2)^{(n+1)/2}} \\
&= \frac{1}{\Gamma(n/2)\sqrt{n\pi}} \Gamma\left(\frac{n+1}{2}\right) \frac{n^{(n+1)/2}}{(n+t^2)^{(n+1)/2}}
\end{aligned}
$$

Finalmente obtenemos

$$
f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(n/2)\sqrt{n\pi}} \left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \qquad (t > 0) \tag{2}
$$

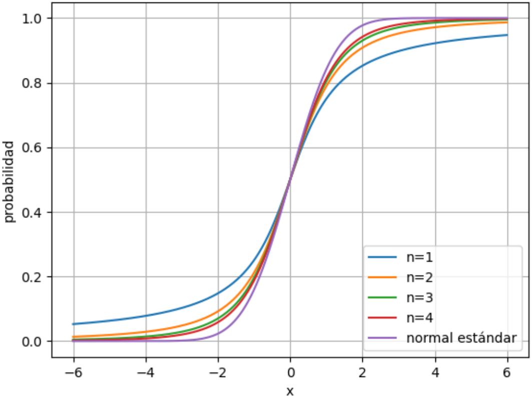
Esta distribución se conoce como distribución t de Student con n grados de libertad.

Gráfico de la densidad t de Student



Cuando $n \to +\infty$, estas curvas convergen a la densidad normal estándar [¡ejercicio fácil de límites!].

Gráfico de la distribución acumulada de una t de Studient



Cuando $n \to +\infty$, estas curvas convergen a la distribución acumulada de una normal estándar.

Dato curioso

Como veremos más adelante, esta distribución de probabilidad surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño y la desviación estándar poblacional es desconocida.

La distribución de Student fue descripta en el año 1908 por William Sealy Gosset. Gosset trabajaba en una fábrica de cerveza, Guinness, que prohibía a sus empleados la publicación de artículos científicos debido a una difusión previa de secretos industriales. De ahí que Gosset publicase sus resultados bajo el pseudónimo de "Student".

Parte IV

La distribución exponencial y los Procesos de Poisson

La Distribución Exponencial y la propiedad de Falta de Memoria

Recordamos que una variable aleatoria X tiene distribución exponencial $\text{Exp}(\lambda)$ (donde $\lambda > 0$) cuando su densidad de probabilidad es

$$f_X(x) = \lambda e^{-\lambda x} I_{(0,+\infty)}(x)$$

La distribución exponencial es un modelo muy útil para distintos procesos: llamadas que llegan a una central telefónica o peticiones a un servidor informático, tiempo de duración de una lámpara, desintegración radiactiva, etc.

Por ejemplo, para fijar ideas, consideremos la desintegración radiactiva de un átomo. La hipótesis fundamental que haremos para describir este fenómeno, es la propiedad de "falta de memoria" que establece que la probabilidad de que un átomo se desintegre en un intervalo de tiempo de longitud Δt sólo depende de la longitud del intervalo y es independiente de la historia anterior del material.

La Distribución Exponencial y la propiedad de Falta de Memoria (2)

Podemos describir con más precisión esta propiedad de la siguiente manera: Si llamamos T al tiempo en el que el átomo se desintegra, T es una variable aleatoria. La probabilidad condicional de que el átomo se desintegre en el intervalo $(t_0, t_0 + \Delta t]$ sabiendo que no se ha desintegrado aún en tiempo $t = t_0$, es igual a la probabilidad de que se desintegre en el intervalo $(0, \Delta t]$:

$$P\{T > t_0 + \Delta t / T > t_0\} = P\{T > \Delta t\}$$

Por definición de probabilidad condicional, esto significa que:

$$\frac{P\{t < T \leq t + \Delta t\}}{P\{T > t\}} = P\{T > \Delta t\}$$

Llammemos F a la función de distribución de T, y sea $G(t) = 1 - F(t)$. Entonces, esta igualdad establece que:

$$G(t + \Delta t) = G(t)G(\Delta t)$$

La Distribución Exponencial y la propiedad de Falta de Memoria (3)

Necesitaremos el siguiente lema:

Lema

Sea $G : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *una función continua que satisface que:*

$$G(t+s) = G(t)G(s)$$

Entonces: $G(t) = G(0)a^t$, *siendo* $a = G(1)$.

La Distribución Exponencial y la propiedad de Falta de Memoria (4)

Volviendo a nuestro problema de la desintegración radiactiva, si ponemos $G(1) = e^{-\lambda}$ (suponiendo $G(0) \neq 0$), y observamos que $G(0) = 1$ pues $T > 0$ (El átomo no se desintegró aún en $t = 0$), obtenemos que:

$$G(t) = e^{-\lambda t}$$

Por consiguiente la función de distribución de T es:

$$F(t) = 1 - e^{-\lambda t}$$

y derivando vemos que su densidad es

$$f(t) = \lambda \, e^{-\lambda t} \ (t > 0)$$

La ley de Rutherford-Soddy

Supongamos ahora que tenemos un material radiactivo formado inicialmente por un grán número de átomos N_0, y llamemos $N(t)$ a la cantidad de átomos no desintegrados hasta el instante t. Hagamos la hipótesis de que las desintegraciones de los distintos átomos son independientes. Podemos pensar que son ensayos de Bernoullí, entonces por la ley de los grandes números

$$\frac{N(t)}{N_0} \approx P\{T > t_0\}$$

y deducimos que:

$$N(t) = N_0 \, e^{-\lambda t}$$

Esta expresión se conoce como la ley de desintegración radiactiva de Rutherford-Soddy (1902). El valor de la constante λ depende de la sustancia.

La ley de Rutherford-Soddy (2)

Se define semivida o período de semi-desintegración $T_{1/2}$ el tiempo en que una muestra de material radiactivo tarda en reducirse a la mitad. De la fórmula anterior, se deduce que

$$T_{1/2} = \frac{\log 2}{\lambda}$$

La siguiente tabla muestra por ejemplo los períodos de semi-desintegración de algunos isótopos radiactivos:

Isótopo	$T_{1/2}$
Berilio-8	$10^{-16} s$
Polonio-213	$4x10^{-6} s$
Aluminio-28	2.25 min
Yodo-131	8 días
Estroncio-90	28 años
Radio-226	1600 años
Carbono-14	5730 años
Rubidio-87	$5{,}7 \times 10^{10}$ años

Tiempos de espera y procesos de Poisson

Llamemos T_i al tiempo en que ocurre la iésima densintegración radiactiva, de modo que:

$$T_1 < T_2 < \ldots < T_n$$

(Podemos suponer para simplificar que no hay dos desintegraciones simultáneas, ya que la probabilidad de que ello ocurra es despreciable). Notemos que:

$$T_n = T_1 + (T_2 - T_1) + (T_3 - T_2) + \ldots + (T_n - T_{n-1})$$

Las variables $T_k - T_{k-1}$ representan el tiempo entre la $(k-1)$-ésima desintegración y la k-ésima desintegración. Por la discusión anterior (y la propiedad de falta de memoria), $T_k - T_{k-1}$ tiene distribución exponencial de parámetro $\lambda > 0$ (donde $\lambda > 0$ es una constante que depende del material que estamos considerando).
Por otra parte, si suponemos que el tiempo que un átomo tarda en desintegrarse es independiente de lo que tardan los demás, las $T_{k+1} - T_k$ serán variables aleatorias independientes. Entonces la variable T_n será dada por una suma de n variables aleatorias independientes, todas con distribución exponencial de parámetro λ.

Tiempos de espera y procesos de Poisson(2)

Como $\text{Exp}(\lambda) = \Gamma(1, \lambda)$, deducimos que T_n tiene distribución $\Gamma(n, \lambda)$, es decir que se distribuye según la densidad $g_n(t)$ dada por:

$$g_n(t) = \begin{cases} \frac{\lambda^n}{(n-1)!}\ t^{n-1}\ e^{-\lambda t} & \text{si} \quad t > 0 \\ 0 & \text{si} \quad t \leq 0 \end{cases}$$

Llamemos $D(t)$ al número de desintegraciones en el intervalo $[0, t]$. Entonces

$$D(t_0) = n \text{ si y sólo si } T_n \leq t_0 < T_{n+1}$$

Deducimos que:

$$\{D(t_0) = n\} = \{T_n \leq t_0\} - \{T_{n+1} \leq t_0\}$$

En consecuencia,

$$P\{D(t_0) = n\} = P\{T_n \leq t_0\} - P\{T_{n+1} \leq t_0\} = \int_0^{t_0} g_n(t)\ dt - \int_0^{t_0} g_{n+1}(t)\ dt$$

Tiempos de espera y procesos de Poisson(3)

Integrando por partes, tenemos que:

$$\begin{aligned}
\int_0^{t_0} g_{n+1}(t)\, dt &= \int_0^{t_0} \frac{\lambda^{n+1}}{n!}\ t^n e^{-\lambda t}\ dt \\
&= \frac{\lambda^{n+1}}{n!} \left[t^n\ \frac{e^{-\lambda t}}{(-\lambda)} \Big|_0^{t_0} - \int_0^{t_0} n\ t^{n-1}\ \frac{e^{-\lambda t}}{(-\lambda)}\ dt \right] \\
&= \frac{\lambda^{n+1}}{n!}\ t_0^n \frac{e^{-\lambda t_0}}{(-\lambda)} - 0 - \int_0^{t_0} \frac{\lambda^{n+1}}{n!}\ n\ t^{n-1}\ \frac{e^{-\lambda t}}{(-\lambda)}\ dt \\
&= -\frac{\lambda^n}{n!}\ t_0^n\ e^{-\lambda t_0} + \int_0^{t_0} \frac{\lambda^n}{(n-1)!}\ t^{n-1}\ e^{-\lambda t}\ dt \\
&= -\frac{\lambda^n}{n!}\ t_0^n\ e^{-\lambda t_0} + \int_0^{t_0} g_n(t)\ dt
\end{aligned}$$

Tiempos de espera y procesos de Poisson(4)

En definitiva concluimos que la distribución del número de desintegraciones viene dada por una distribución de Poisson (proceso de Poisson):

$$P\{D(t_0) = n\} = \frac{(\lambda t_0)^n}{n!}\, e^{-\lambda t_0}$$

Como dijimos al comienzo de la sección, aunque hemos presentado la distribución exponencial y este cálculo de los tiempos de espera como modelo de la desintegración radiactiva, este mismo modelo se puede aplicar a otros procesos donde la hipótesis de falta de memoria resulte razonable como por ejemplo la llegada de eventos a un servidor informático, o los siniestros en una companía de seguros. Esto explica la utilización de las distribuciones exponencial y de Poisson en muchas aplicaciones de las probabilides.