Transformaciones de variables aleatorias múltiples

Fabián Abarca Calderón

IE0405 – Modelos Probabilísticos de Señales y Sistemas

11

Tema III

UNIVERSIDAD DE COSTA RICA

EIE Escuela de Ingeniería Eléctrica

Similar al caso de las variables aleatorias marginales, es posible “transformar” dos o más variables aleatorias conjuntas mediante funciones, y el objetivo es encontrar la función de densidad de las nuevas variables aleatorias conjuntas.

Transformaciones de variables aleatorias múltiples

Considérese el caso de hallar la densidad conjunta para un conjunto de nuevas variables aleatorias Y_i:

$$Y_i = T_i(X_1, X_2, \ldots, X_N) \qquad i = 1, 2, \ldots, N \tag{1}$$

definidas por transformaciones funcionales T_i.

Nota: X_i puede ser continua, discreta o mixta, mientras las funciones T_i pueden ser lineales o no, continuas, segmentadas, etcétera.

Transformaciones de variables aleatorias múltiples II

- Las nuevas variables aleatorias Y_i son producidas por funciones **univaluadas** continuas T_i con derivadas parciales continuas en todas partes.
- Existe un conjunto de funciones inversas continuas T_j^{-1} tal que las variables originales puedan expresarse como funciones continuas univaluadas de las variables nuevas:

$$X_j = T_j^{-1}(Y_1, Y_2, \ldots, Y_N) \qquad j = 1, 2, \ldots, N \tag{2}$$

Estas suposiciones implican que un punto en el espacio de muestras conjunto de las X_i mapea en un solo punto en el espacio de las nuevas variables Y_j.

Premisa de la relación entre X_i e Y_i I

- Sea $\mathcal{R}_X$ una región cerrada de puntos en el espacio de las X_i, y $\mathcal{R}_Y$ sea la región correspondiente de puntos mapeados en el espacio de las Y_j.
- Entonces, **la probabilidad de que un punto caiga en $\mathcal{R}_X$ iguala a la probabilidad de que su punto mapeado caiga en $\mathcal{R}_Y$.**

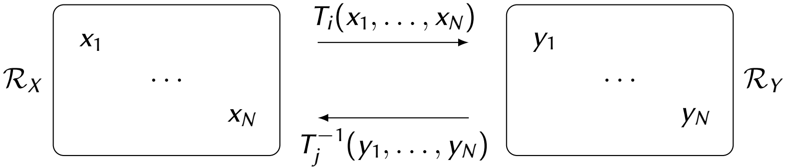
Estas probabilidades, en términos de densidades conjuntas, son:

$$\begin{aligned}
&\underbrace{\int \cdots \int}_{\text{sobre } \mathcal{R}_X} f_{X_1,\ldots,X_N}(x_1,\ldots,x_N)\,\mathrm{d}x_1 \cdots \,\mathrm{d}x_N \\
&\qquad = \underbrace{\int \cdots \int}_{\text{sobre } \mathcal{R}_Y} f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N)\,\mathrm{d}y_1 \ldots \,\mathrm{d}y_N
\end{aligned} \tag{3}$$

Resolver para $f_{Y_1,\ldots,Y_N}(y_1, \ldots, y_N)$ I

La ecuación (3) puede resolverse para $f_{Y_1,\ldots,Y_N}(y_1, \ldots, y_N)$ tratándola como una integral múltiple donde se hace un cambio de variables.

- Las variables x_i se cambian a nuevas variables y_i por medio de la transformación, de forma que el integrando se cambia por sustitución funcional directa.
- Los límites cambian de la región $\mathcal{R}_X$ a la región $\mathcal{R}_Y$.
- Finalmente, el diferencial de volumen $\,\mathrm{d}x_1 \ldots \,\mathrm{d}x_N$ cambiará al valor $|J| \,\mathrm{d}y_1 \ldots \,\mathrm{d}y_N$, donde $|J|$ es la magnitud del jacobiano J de las transformaciones.



El lado izquierdo de la ecuación (3) se convierte en:

$$\begin{aligned}\underbrace{\int\cdots\int}_{\text{sobre } \mathcal{R}_X} f_{X_1,\ldots,X_N}(x_1,\ldots,x_N)\,\mathrm{d}x_1\cdots\,\mathrm{d}x_N = \\ \underbrace{\int\cdots\int}_{\text{sobre } \mathcal{R}_Y} f_{X_1,\ldots,X_N}(x_1 = T_1^{-1},\ldots,x_N = T_N^{-1})|J|\,\mathrm{d}y_1\cdots\,\mathrm{d}y_N\end{aligned} \tag{4}$$

Dado que este resultado debe igualar el lado derecho de la ecuación 3, entonces:

$$f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N) = f_{X_1,\ldots,X_N}(x_1 = T_1^{-1},\ldots,x_N = T_N^{-1})|J| \tag{5}$$

Nota sobre el jacobiano

El jacobiano es el *determinante* de una matriz de derivadas parciales, definido por:

$$J = \begin{vmatrix} \frac{\partial T_1^{-1}}{\partial Y_1} & \cdots & \frac{\partial T_1^{-1}}{\partial Y_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial T_N^{-1}}{\partial Y_1} & \cdots & \frac{\partial T_N^{-1}}{\partial Y_N} \end{vmatrix} = \det \begin{pmatrix} \frac{\partial T_1^{-1}}{\partial Y_1} & \cdots & \frac{\partial T_1^{-1}}{\partial Y_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial T_N^{-1}}{\partial Y_1} & \cdots & \frac{\partial T_N^{-1}}{\partial Y_N} \end{pmatrix} \tag{6}$$

Dada una densidad conjunta $f_{X,Y}(x, y)$ de dos variables aleatorias, y dos funciones[1]

$$Z = g(X, Y) \quad \text{y} \quad W = h(X, Y) \tag{7}$$

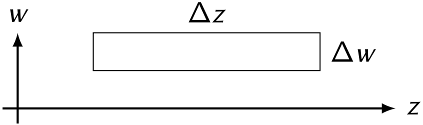
queremos encontrar la densidad conjunta $f_{Z,W}(z, w)$.
Igual que antes:

- Si $g(X, Y)$ y $h(X, Y)$ son continuas y diferenciables, es posible desarrollar una forma de obtener el pdf directamente.
- Para derivar el resultado anterior, se asume primero que $g(X, Y)$ y $h(X, Y)$ son transformaciones uno a uno, de forma que su inversa existe, aquí definidas como[2]

$$X \triangleq P(Z, W) \quad \text{y} \quad Y \triangleq Q(Z, W) \tag{8}$$

[1] Con la notación ya descrita, g y h bien podrían llamarse T_1 y T_2.
[2] Con la notación ya descrita, P y Q bien podrían llamarse T_1^{-1} y T_2^{-1}.

En el plano zw la región de interés está definida por un rectángulo de lados Δz y Δw.



- Si Δz y Δw son suficientemente pequeños, se puede aproximar la siguiente probabilidad $P(\cdot)$:

$$P\left[(Z, W) \in \Delta z \Delta w\right] \approx f_{Z,W}(z, w) \cdot \Delta z \Delta w \tag{9}$$

Nota: recordar la interpretación de la probabilidad conjunta como el volumen debajo de la curva bidimensional de $f_{X,Y}(x, y)$. En este caso sería "base por altura".

Esta área se mapea mediante las transformaciones $x = P(z, w)$ y $y = Q(z, w)$ a una nueva área Δ en el plano xy. Debido al mapeo entre (X, Y) y (Z, W), lo anterior debe ser igual a

$$P\left((X, Y) \in \Delta\right) \approx f_{X,Y}(x = P(z, w), y = Q(z, w)) \cdot \Delta \tag{10}$$

- ¿Cómo relacionar entonces las áreas $\Delta z \Delta w$ y Δ? Igualando el área de un rectángulo con el de un paralelogramo, se puede demostrar (https://youtu.be/rnnrUDdKzm0) que

$$\Delta = |J| \Delta z \Delta w \tag{11}$$

donde

$$J(z,w) \triangleq \frac{\partial(P,Q)}{\partial(z,w)} = \det \begin{pmatrix} \frac{\partial P}{\partial z} & \frac{\partial P}{\partial w} \\ \frac{\partial Q}{\partial z} & \frac{\partial Q}{\partial w} \end{pmatrix} \tag{12}$$

y se le llama “el jacobiano de la transformación (P, Q)”.

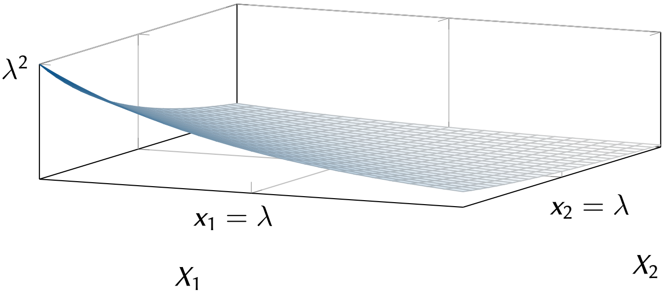
Así entonces,

$$f_{Z,W}(z,w) = f_{X,Y}\left[x = P(z,w), y = Q(z,w)\right] |J| \tag{13}$$

Comparar con $f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N) = f_{X_1,\ldots,X_N}(x_1 = T_1^{-1},\ldots,x_N = T_N^{-1})|J|$.

Sea X_1 una VA para el tiempo de vida de un componente. El componente puede sustituirse una única vez, y X_2 representa el tiempo de vida del repuesto. X_1 y X_2 son **iid** con distribución exponencial de parámetro λ. Por lo anterior, la PDF conjunta es

$$f_{X_1,X_2}(x_1, x_2) = \lambda^2 e^{-\lambda(x_1+x_2)} \quad \text{para} \quad x_1 > 0, x_2 > 0$$



Podríamos estar interesados en dos cantidades (transformaciones arbitrarias u_1 y u_2):

- La vida total del dispositivo

$$Y_1 = u_1(X_1, X_2) = X_1 + X_2$$

- La proporción de la vida del primer componente a la vida total del dispositivo

$$Y_2 = u_2(X_1, X_2) = \frac{X_1}{X_1 + X_2}$$

¿Cuál es la distribución conjunta de Y_1 y Y_2?

Por el teorema de transformación, escrito como

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}[x_1 = v_1(y_1, y_2), x_2 = v_2(y_1, y_2)] \, |\det(\mathbf{M})|$$

Y dadas las funciones inversas

- $x_1 = v_1(y_1, y_2) = y_1 y_2$
- $x_2 = v_2(y_1, y_2) = y_1(1 - y_2)$

de donde se obtiene también la región de la "imagen de la transformación", en $y_1 > 0, 0 < y_2 < 1$.

En una transformación bivariada se cumple que

$$J(y_1, y_2) \triangleq \det(\mathbf{M}) \triangleq \frac{\partial(v_1, v_2)}{\partial(y_1, y_2)} \triangleq \det\begin{pmatrix} \frac{\partial v_1}{\partial y_1} & \frac{\partial v_1}{\partial y_2} \\ \frac{\partial v_2}{\partial y_1} & \frac{\partial v_2}{\partial y_2} \end{pmatrix}$$

de donde, dados $v_1(y_1, y_2)$ y $v_2(y_1, y_2)$,

- $\frac{\partial v_1}{\partial y_1} = y_2$
- $\frac{\partial v_1}{\partial y_2} = y_1$
- $\frac{\partial v_2}{\partial y_1} = (1 - y_2)$
- $\frac{\partial v_2}{\partial y_2} = -y_1$

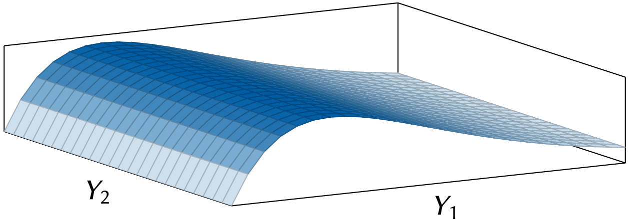
y, por tanto,

$$\det(\mathbf{M}) = \frac{\partial v_1}{\partial y_1} \cdot \frac{\partial v_2}{\partial y_2} - \frac{\partial v_1}{\partial y_2} \cdot \frac{\partial v_2}{\partial y_1} = -y_1 y_2 - y_1(1 - y_2) = -y_1$$

Finalmente, la nueva función de densidad viene de

$$
\begin{aligned}
f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}[x_1 = v_1(y_1,y_2), x_2 = v_2(y_1,y_2)] \left| \frac{\partial(v_1,v_2)}{\partial(y_1,y_2)} \right| \\
&= \lambda^2 e^{-\lambda(y_1 y_2 + y_1(1-y_2))} |-y_1| = \lambda^2 y_1 e^{-\lambda(y_1)} \\
&= \lambda^2 y_1 e^{-\lambda y_1} \cdot 1 \quad \text{para} \quad y_1 > 0, 0 < y_2 < 1
\end{aligned}
$$

donde se observa que no depende de y_2, y 1 está distribuido en $0 < y_2 < 1$.



Se observa además que

- La distribución del tiempo de vida $X_1 + X_2$ se conoce como *distribución gama*, con parámetros $\alpha = 2$ y $\beta = 2$.
- La proporción del tiempo de vida del componente original es uniforme entre 0 y 1.
- Y_1 y Y_2 son independientes entre sí.

Sea $Z = X + Y$, y $W = X - Y$. Entonces

$$X = \frac{Z+W}{2} \triangleq P(Z, W)$$
$$Y = \frac{Z-W}{2} \triangleq Q(Z, W)$$

Por tanto

$$|J(z,w)| = \left|\det\begin{pmatrix} \frac{\partial P}{\partial z} & \frac{\partial P}{\partial w} \\ \frac{\partial Q}{\partial z} & \frac{\partial Q}{\partial w} \end{pmatrix}\right| = \left|\det\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}\right| = \frac{1}{2}$$

y así

$$\boxed{f_{Z,W}(z,w) = \frac{1}{2} f_{X,Y}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)}$$

Sea, para un ángulo fijo θ_0,

$$Z = X\cos\theta_0 + Y\sin\theta_0 \qquad W = X\sin\theta_0 - Y\cos\theta_0$$

y así

$$\begin{aligned} X &= Z\cos\theta_0 + W\sin\theta_0 \triangleq P(Z, W) \\ Y &= Z\sin\theta_0 - W\cos\theta_0 \triangleq Q(Z, W) \end{aligned}$$

por tanto

$$|J(z, w)| = \left|\det\begin{pmatrix} \frac{\partial P}{\partial z} & \frac{\partial P}{\partial w} \\ \frac{\partial Q}{\partial z} & \frac{\partial Q}{\partial w} \end{pmatrix}\right| = \left|\det\begin{pmatrix} \cos\theta_0 & \sin\theta_0 \\ \sin\theta_0 & -\cos\theta_0 \end{pmatrix}\right| = 1$$

de forma que una transformación rotacional como la descrita no cambia la función de densidad conjunta.

Sea

$$\begin{aligned} R^2 &= X^2 + Y^2 & 0 \leq R \leq \infty \\ \Theta &= \tan^{-1}(Y/X) & -\pi < \Theta < \pi \end{aligned}$$

$$X = R\cos\Theta \triangleq P(R,\Theta) \qquad Y = R\sin\Theta \triangleq Q(R,\Theta)$$

Luego,

$$|J(r,\theta)| = \left|\det\begin{pmatrix} \frac{\partial P}{\partial r} & \frac{\partial P}{\partial \theta} \\ \frac{\partial Q}{\partial r} & \frac{\partial Q}{\partial \theta} \end{pmatrix}\right| = \left|\det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}\right| = r$$

de forma que

$$\boxed{f_{R,\Theta}(r,\theta) = f_{X,Y}(r\cos\theta, r\sin\theta)r}$$

Distribución probabilística de la suma de variables aleatorias

Se va a analizar el problema de hallar las funciones de densidad y distribución para una suma de variables aleatorias **estadísticamente independientes**.

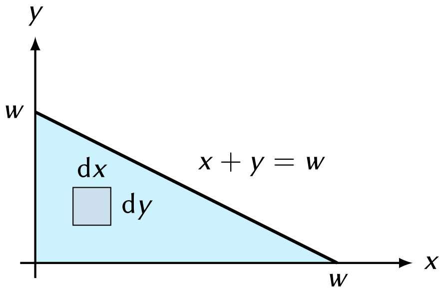
Sea W una variable aleatoria igual a la suma de dos variables aleatorias independientes X e Y:

$$W = X + Y \tag{14}$$

Nota: X pudiera representar una señal aleatoria de tensión e Y pudiera representar ruido aleatorio. La suma W pudiera representar entonces una tensión de señal más ruido, en algún receptor.

La función de distribución de probabilidad que se busca está definida por:

$$F_W(w) = P\{W \leq w\} = P\{X + Y \leq w\} \tag{15}$$



La probabilidad correspondiente a un área elemental $\mathrm{d}x\,\mathrm{d}y$ en el plano XY localizado en el punto (x, y) es $f_{X,Y}\,\mathrm{d}x\,\mathrm{d}y$. Si se suma todas las probabilidades sobre la región donde $x + y \leq w$ se obtendrá $F_W(w)$:

$$F_W(w) = \int_{-\infty}^{\infty} \int_{x=-\infty}^{w-y} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y \tag{16}$$

Como X y Y son independientes, es decir, $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, entonces,

$$F_W(w) = \int_{-\infty}^{\infty} f_Y(y) \int_{x=-\infty}^{w-y} f_X(x)\,\mathrm{d}x\,\mathrm{d}y \tag{17}$$

Después de derivar, usando la regla de Leibniz, se obtiene la función de densidad:

$$f_W(w) = \int_{-\infty}^{\infty} f_Y(y) f_X(w-y)\,\mathrm{d}y \tag{18}$$

que describe una **integral de convolución**, por tanto:

La función de densidad de la suma…

…de dos variables aleatorias estadísticamente independientes es la convolución de sus funciones de densidad individuales.

Encuentre la función de densidad de $W = X + Y$ donde las densidades respectivas son:

$$
\begin{aligned}
f_X(x) &= \left(\frac{1}{a}\right)[u(x) - u(x-a)] \\
f_Y(y) &= \left(\frac{1}{b}\right)[u(y) - u(y-b)]
\end{aligned}
$$

con $0 < a < b$.

Este ejercicio se puede resolver mediante la integral de convolución o mediante el método de la transformada de Laplace. Se escoge este último método al notarse que

ambas funciones de densidad están definidas a partir del origen, lo cual facilita un manejo algebraico del problema.
Se tiene entonces:

$$
\begin{aligned}
\mathcal{L}\{f_X(x)\} &= \int_0^a \frac{1}{a} e^{-sx} dx \\
&= \left(\frac{1}{a}\right) \left.\frac{e^{-sx}}{-s}\right|_0^a \\
&= \left(\frac{1}{as}\right) \left[1 - e^{-as}\right] \\
\mathcal{L}\{f_Y(y)\} &= \left(\frac{1}{bs}\right) \left[1 - e^{-bs}\right]
\end{aligned}
$$

La transformada de Laplace de una integral de convolución es igual al producto de las transformadas de Laplace de las funciones que conforman el integrando de tal integral. El próximo paso es entonces encontrar el producto de las dos transformadas de Laplace calculadas para luego encontrar la transformada de Laplace inversa y así hallar la nueva función de densidad, correspondiente a la nueva variable aleatoria W.

$$\mathcal{L}\{f_X(x)\}\mathcal{L}\{f_Y(y)\} = \left(\frac{1}{ab}\right)\left(\frac{1}{s^2}\right)\left[1 - e^{-bs} - e^{-as} + e^{-(a+b)s}\right]$$

Lo que sigue ahora es encontrar la transformada de Laplace inversa, para lo que es importante recordar la siguiente transformada:

$$\mathcal{L}\{t^n\} = \frac{n!}{s^n + 1}$$

Se obtiene entonces:

$$
\begin{aligned}
f_W(w) &= \mathcal{L}^{-1}\left\{\frac{1}{abs^2}\left[1 - e^{-as} - e^{-bs} + e^{-(a+b)s}\right]\right\} \\
&= \frac{1}{ab}wu(w) - \frac{1}{ab}(w-a)u(w-a) - \frac{1}{ab}(w-b)u(w-b) \\
&\quad + \frac{1}{ab}(w-a-b)u(w-a-b)
\end{aligned} \tag{19}
$$

Nótese que se ha hecho cambios de variable para emplear correctamente la transformada de Laplace. Finalmente, dado que el resultado debe ser una función de densidad, una prueba que se puede hacer para probar si el resultado es correcto, es comprobar que el área bajo la curva de la nueva función es en efecto igual a la unidad.

La función de densidad de la suma de varias VA

La función de densidad de $Y = X_1 + X_2 + \cdots + X_N$, donde las X_i son variables aleatorias estadísticamente independientes entre sí, es la convolución de las N funciones de densidad individuales:

$$f_Y(y) = f_{X_N}(x_N) * f_{X_{N-1}}(x_{N-1}) * \cdots * f_{X_1}(x_1) \tag{20}$$