Functions of a Random Vector

Guy Lebanon

January 6, 2006

Assuming that we know the distribution of $\vec{\boldsymbol{X}} = (X_1, \ldots, X_n)$, and we have a mapping $g : \mathbb{R}^n \to \mathbb{R}^k$ yielding $\vec{\boldsymbol{Y}} = g(\vec{\boldsymbol{X}})$ (in other words $Y_1 = g_1(X_1, \ldots, X_n), \ldots, Y_k = g_k(X_1, \ldots, X_n)$). An important question is: can we find the distribution (in terms of cdf, pdf or pmf) of $\vec{\boldsymbol{Y}} = (Y_1, \ldots, Y_k)$ in terms of the distribution (cdf, pdf or pmf) of $\vec{\boldsymbol{X}}$?

This is a complicated question and we will answer it partially for some special cases. For the case $k = 1$, $\vec{\boldsymbol{Y}} = Y = g(X_1, \ldots, X_n)$ we have

$$F_Y(y) = P(g(X_1, \ldots, X_n) \leq y) = \begin{cases} \int_{\vec{\boldsymbol{x}}:g(\vec{\boldsymbol{x}})\leq y} f_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}})d\vec{\boldsymbol{x}} & \vec{\boldsymbol{X}} \text{ is continuous} \\ \sum_{\vec{\boldsymbol{x}}:g(\vec{\boldsymbol{x}})\leq y} p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) & \vec{\boldsymbol{X}} \text{ is discrete} \end{cases}$$

Important Example: For $\vec{\boldsymbol{X}} = (X_1, X_2)$ a continuous random vector, and $Y = X_1 + X_2$, we get

$$F_Y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y-w} f_{X_1,X_2}(w, z)\; dzdw$$

and by the fundamental theorem of calculus

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \int_{-\infty}^{\infty} f_{X_1,X_2}(w, y - w)\, dw.$$

If X_1 and X_2 are independent, $f_{X_1,X_2} = f_{X_1} f_{X_2}$ and the integral above representing the pdf f_Y becomes the convolution of the marginal pdfs

$$f_{X_1+X_2}(y) = \int_{-\infty}^{\infty} f_{X_1}(w) f_{X_2}(y - w)\, dw = (f_{X_1} \star f_{X_2})(y).$$

Similarly, for discrete random vector we obtain that the pmf of the sum of two independent random variables is the (discrete) convolution of their pmfs. By induction we have that $f_{X_1+\ldots+X_n}(y) = (f_{X_1} \star \cdots \star f_{X_n})(y)$. It is a good idea at this point to turn to signals and systems material and remind yourself of the basic properties of convolution (commutativity, associativity etc.).

In the case $k > 1$ we have $Y_1 = g_1(X_1, \ldots, X_n), \ldots, Y_k = g_k(X_1, \ldots, X_n)$ or written succinctly $\vec{\boldsymbol{Y}} = g(\vec{\boldsymbol{X}})$ where $g : \mathbb{R}^n \to \mathbb{R}^n$ is the mapping $g(\vec{\boldsymbol{x}}) = (g_1(\vec{\boldsymbol{x}}), \ldots, g_k(\vec{\boldsymbol{x}}))$. The cdf of the vector $\vec{\boldsymbol{Y}}$ is

$$F_{Y_1,\ldots,Y_k}(y_1, \ldots, y_k) = P(g_1(\vec{\boldsymbol{X}}) \leq y_1, \ldots, g_k(\vec{\boldsymbol{X}}) \leq y_1) = \begin{cases} \int_A f_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}})\, d\vec{\boldsymbol{x}} & \vec{\boldsymbol{X}} \text{ is continuous} \\ \sum_{\vec{\boldsymbol{x}}\in A} p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) & \vec{\boldsymbol{X}} \text{ is discrete} \end{cases}$$

where

$$A = \{\vec{\boldsymbol{x}} \in \mathbb{R}^n : g_j(\vec{\boldsymbol{x}}) \leq y_j \text{ for } j = 1, \ldots, k\}.$$

If $\vec{\boldsymbol{Y}}$ is discrete we can find its pmf by

$$p_{\vec{\boldsymbol{Y}}}(y_1, \ldots, y_k) = P(g_1(\vec{\boldsymbol{X}}) = y_1, \ldots, g_k(\vec{\boldsymbol{X}}) = y_k) = \begin{cases} \int_A f_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}})\, d\vec{\boldsymbol{x}} & \vec{\boldsymbol{X}} \text{ is continuous} \\ \sum_{\vec{\boldsymbol{x}}\in A} p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) & \vec{\boldsymbol{X}} \text{ is discrete} \end{cases}$$

where

$$A = \{\vec{\boldsymbol{x}} \in \mathbb{R}^n : g_j(\vec{\boldsymbol{x}}) = y_j \text{ for } j = 1, \ldots, k\}.$$

If $\vec{\boldsymbol{Y}}$ is continuous, we can obtain the pdf $f_{\vec{\boldsymbol{Y}}}$ either through the joint cdf (if it is available)

$$f_{\vec{\boldsymbol{Y}}} = \frac{\partial^k}{\partial y_1 \cdots \partial y_k} F_{\vec{\boldsymbol{Y}}}$$

or directly using a generalization of the formula to random vectors $f_{g(X)}(y) = \frac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y))$.

Theorem 1. *Using the above setup, let $k = n$, and write $g(\vec{\boldsymbol{x}}) = (g_1(\vec{\boldsymbol{x}}), \ldots, g_n(\vec{\boldsymbol{x}}))$ where $g : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible and differentiable mapping (in the range of $\vec{\boldsymbol{X}}$): there exists an inverse $g^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ such that $g^{-1}(g(\vec{\boldsymbol{x}})) = g(g^{-1}(\vec{\boldsymbol{x}})) = \vec{\boldsymbol{x}}$. Then*

$$f_{\vec{\boldsymbol{Y}}}(\vec{\boldsymbol{y}}) = \frac{1}{|\det J(g^{-1}(\vec{\boldsymbol{y}}))|} f_{\vec{\boldsymbol{X}}}(g^{-1}(\vec{\boldsymbol{y}}))$$

where $\det J(g^{-1}(\vec{\boldsymbol{y}}))$ *is the determinant of the Jacobian matrix at $\vec{\boldsymbol{x}} = g^{-1}(\vec{\boldsymbol{y}})$:*

$$J(\vec{\boldsymbol{x}}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\vec{\boldsymbol{x}}) & \cdots & \frac{\partial g_1}{\partial x_n}(\vec{\boldsymbol{x}}) \\ \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\vec{\boldsymbol{x}}) & \cdots & \frac{\partial g_n}{\partial x_n}(\vec{\boldsymbol{x}}) \end{pmatrix}$$

Proof. By a direct application of the change of variable formula from multivariate calculus. □

Important Example: Assume $g : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation expressed by an invertible matrix T of size $n \times n$: $g(\vec{\boldsymbol{x}}) = T\vec{\boldsymbol{x}}$ and $g^{-1}(\vec{\boldsymbol{y}}) = T^{-1}\vec{\boldsymbol{y}}$. The jacobian $J(g^{-1}(\vec{\boldsymbol{y}}))$ is simply T and we obtain

$$f_{\vec{\boldsymbol{Y}}}(\vec{\boldsymbol{y}}) = \frac{1}{|\det T|} f_{\vec{\boldsymbol{X}}}(T^{-1}\vec{\boldsymbol{y}}).$$

Example: if $Y_1 = X_1$ and $Y_2 = X_1 + X_2$, we have that the vector mapping g is represented by the matrix $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ whose inverse is $A^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ (since $X_1 = Y_1, X_2 = Y_2 - Y_1$) and

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{|1|} f_{X,Y}(y_1, y_2 - y_1).$$

Recall that we proved above that for two independent variables the pdf of $X_1 + X_2$ is the convolution $f_{X_1} \star f_{X_2}$. We can now reaffirm that by obtaining the marginal of the pdf f_{Y_1,Y_2}:

$$f_{X_1+X_2}(y_2) = f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(y_1, y_2 - y_1)\, dy_1 = \int_{-\infty}^{\infty} f_{X_1}(y_1) f_{X_2}(y_2 - y_1)\, dy_1.$$

Example: Let X_1, X_2 be independent exponential RV and consider the distribution of Y_1, Y_2 where $Y_1 = X_1/(X_1 + X_2)$ and $Y_2 = X_1 + X_2$. Thus $g_1(x_1, x_2) = x_1/(x_1 + x_2), g_2(x_1, x_2) = x_1 + x_2$ and the inverse transformation is $g_1^{-1}(y_1, y_2) = y_1 y_2$, $g_2^{-1}(y_1, y_2) = (1 - y_1)y_2$. The Jacobian of g at (x_1, x_2) is $J(x_1, x_2) = \begin{pmatrix} \frac{x_2}{(x_1+x_2)^2} & -\frac{x_1}{(x_1+x_2)^2} \\ 1 & 1 \end{pmatrix}$, and its determinant is $\frac{x_2}{(x_1+x_2)^2} + \frac{x_1}{(x_1+x_2)^2} = \frac{1}{x_1+x_2}$. The joint pdf of (Y_1, Y_2) is then

$$f_{Y_1,Y_2}(y_1, y_2) = |(y_1 y_2 + (1 - y_1)y_2)|\ \lambda e^{-\lambda y_1 y_2} \lambda e^{-\lambda(1-y_1)y_2} = y_2 \lambda^2 e^{-\lambda y_2}$$

for $0 < y_1 < 1, y_2 > 0$ and 0 otherwise.