Lecture #5: The Transformation Theorem

Reference. §1.2.1 pages 20–23

Last class we discussed how to perform a change-of-variables in the one-dimensional case. In multiple dimensions, the idea is the same, although things get notationally messier.

Recall. The one-dimensional change-of-variables formula is usually written as

$$\int_a^b f(h(x))h'(x)\,\mathrm{d}x = \int_{h(a)}^{h(b)} f(u)\,\mathrm{d}u,$$

i.e., $u = h(x),\ \mathrm{d}u = h'(x)\,\mathrm{d}x$.

We can write this as

$$\int_{h(a)}^{h(b)} f(x)\,\mathrm{d}x = \int_a^b f(h(y))h'(y)\,\mathrm{d}y. \tag{$*$}$$

It is in this form that there is a resemblance to the multidimensional formula.

Let $\mathbf{X} = (X_1, \ldots, X_n)'$ be a random vector, and let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a sufficiently smooth function. That is, we write $g = (g_1, \ldots, g_n)$ with $g_i : \mathbb{R}^n \to \mathbb{R}$. Set $\mathbf{Y} = g(\mathbf{X})$ so that $(Y_1, \ldots, Y_n)' = (g_1(\mathbf{X}), \ldots, g_n(\mathbf{X}))'$ where $g_i(\mathbf{X}) = g_i(X_1, \ldots, X_n)$. Writing things out explicitly gives

$$Y_1 = g_1(X_1, \ldots, X_n)\,,\ \ldots\ ,\ Y_n = g_n(X_1, \ldots, X_n).$$

Let $h = g^{-1}$ so that $\mathbf{X} = h(\mathbf{Y})$, and let

$$J = \left|\frac{\mathrm{d}(\overline{x})}{\mathrm{d}(\overline{y})}\right| = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} & \cdots & \dfrac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \dfrac{\partial x_n}{\partial y_2} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}$$

be the *Jacobian*.

Let $B \subset \mathbb{R}^n$. Our goal is to compute $P\{\mathbf{Y} \in B\}$ which we will do in two ways.

The first way is as follows. Let $f_{\mathbf{Y}}(\overline{y})$ denote the density function for $\mathbf{Y}$ so that

$$P\{\mathbf{Y} \in B\} = \int \cdots \int_B f_{\mathbf{Y}}(\overline{y})\,\mathrm{d}\overline{y}. \tag{$**$}$$

The second way is by changing variables. Let $h = g^{-1}$ so that $h(B) = \{\overline{x} : g(\overline{x}) \in B\}$. Then, in analogy with the one-dimensional case, we have

$$P\{\mathbf{Y} \in B\} = P\{g(\mathbf{X}) \in B\} = P\{\mathbf{X} \in g^{-1}(B)\} = P\{\mathbf{X} \in h(B)\} = \int \underset{h(B)}{\cdots} \int f_{\mathbf{X}}(\overline{x})\,\mathrm{d}\overline{x}$$

and a change-of variables $\overline{y} = g(\overline{x})$ (i.e., $\overline{x} = h(\overline{y}),\ \mathrm{d}\overline{x} = |J|\,\mathrm{d}\overline{y}$) gives

$$\int \underset{h(B)}{\cdots} \int f_{\mathbf{X}}(\overline{x})\,\mathrm{d}\overline{x} = \int \underset{B}{\cdots} \int f_{\mathbf{X}}(h(\overline{y})) \cdot |J|\,\mathrm{d}\overline{y} \qquad (***)$$

Note that if we are in one dimension, then $(***)$ reduces to $(*)$.

We now see that $(*)$ and $(**)$ give two distinct expressions for $P\{\mathbf{Y} \in B\}$, namely

$$P\{\mathbf{Y} \in B\} = \int \underset{B}{\cdots} \int f_{\mathbf{Y}}(\overline{y})\,\mathrm{d}\overline{y} = \int \underset{B}{\cdots} \int f_{\mathbf{X}}(h(\overline{y})) \cdot |J|\,\mathrm{d}\overline{y}.$$

Since this is true for any $B \subset \mathbb{R}^n$ we conclude that the integrands must be equal; that is,

$$f_{\mathbf{Y}}(\overline{y}) = f_{\mathbf{X}}(h(\overline{y})) \cdot |J|. \qquad (\dagger)$$

In other words, this gives us a formula for the density of the random vector $\mathbf{Y} = g(\mathbf{X})$ in terms of the density of the random vector $\mathbf{X} = g^{-1}(\mathbf{Y}) = h(\mathbf{Y})$.

Example. Let $X, Y \in \mathcal{N}(0,1)$ be independent. Prove that $X+Y$ and $X-Y$ are independent $\mathcal{N}(0,2)$ random variables.

Solution. The fact that $X+Y$ and $X-Y$ each have a $\mathcal{N}(0,2)$ distribution can be proved using moment generating functions as done in Stat 251. It is the *independence* of $X+Y$ and $X-Y$ that is newly proved.

Let $U = X+Y$ and $V = X-Y$ so that

$$X = \frac{U+V}{2} \quad \text{and} \quad Y = \frac{U-V}{2}.$$

That is, $\mathbf{X} = (X,Y)'$ and $\mathbf{Y} = (U,V)'$. We also find

- $g(\mathbf{X}) = g(X,Y) = (X+Y, X-Y) = (U,V) = \mathbf{Y}$, and
- $h(\mathbf{Y}) = h(U,V) = \left(\frac{U+V}{2}, \frac{U-V}{2}\right) = (X,Y) = \mathbf{X}$.

We compute the Jacobian as

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = -1/2.$$

By the previous result (†), we conclude

$$\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}(h(u,v)) \cdot |J| \\
&= f_{X,Y}((u+v)/2, (u-v)/2) \cdot 1/2 \\
&= 1/2 \cdot f_X((u+v)/2) \cdot f_Y((u-v)/2) \quad \text{by the assumed independence of } X \text{ and } Y \\
&= \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(u+v)^2}{4}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(u-v)^2}{4}} \\
&= \frac{1}{\sqrt{2\pi \cdot 2}} e^{-\frac{1}{2}\frac{u^2}{2}} \cdot \frac{1}{\sqrt{2\pi \cdot 2}} e^{-\frac{1}{2}\frac{v^2}{2}}
\end{aligned}$$

for $-\infty < u, v < \infty$. In other words, since $f_{U,V}(u,v) = f_U(u) \cdot f_V(v)$ we conclude that $U = X + Y$ and $V = X - Y$ are independent $\mathcal{N}(0,2)$ random variables.