Notes on probabilistic applications of the change of variables formulas from calculus

November 11, 2018

Abstract

These are some notes on applications of the change of variables formula from Calculus

0.1 One variable

Let f be a continuous function on some interval (a,b). Let u be continuously differentiable, strictly monotone function from (a,b) to (c,d). It is allowed that either a or c could be $-\infty$, and the either b or d could be ∞.

Then since u is strictly monotone, it is invertible, and with inverse function $x(u)$, which is continuously differentiable as a function of u. The change of variables formula then says that for any u_0, u_1 with $c < u_0 < u_1 < d$,

$$\int_{x(u_0)}^{x(u_1)} f(x)\mathrm{d}x = \int_{u_0}^{u_1} f(x(u))x'(u)\mathrm{d}u \ . \tag{0.1}$$

This has the following probabilistic interpretation. Suppose that f is the density function of a continuous random variable X. Define a new random variable U by $U = u(X)$. Then, assuming $x(u_0) < x(u_1)$, which is the case if u is monotone increasing, the integral on the left in (0.1) equals

$$P(x(u_0) < X < x(u_1)) \ .$$

But, again since u is monotone increasing,

$$x(u_0) < X < x(u_1) \iff u(x(u_0)) < u(X) < u(x(u_1)) \iff u_0 < U < u_1 \ .$$

Hence the integral on the right in (0.1) equals $P(u_0 < U < u_1)$. It follows immediately that

$$g(u) := f(x(u))x'(u) \tag{0.2}$$

is the probability density function of U, Things are similar if u is monotone decreasing: Then so is x as a function of u, so that $x(u_1) < x(u_2)$, and then the integral on the left equals,

$$-\int_{x(u_1)}^{x(u_0)} f(x)\mathrm{d}x = -P(x(u_1) < X < x(u_2)) \ .$$

Since u is monotone decreasing,

$$x(u_1) < X < x(u_0) \iff u(x(u_1)) > u(X) > u(x(u_0)) \iff u_1 > U > u_0 \ .$$

Hence the integral on the right in (0.1) equals $-P(u_0 < U < u_1)$. It follows immediately that

$$g(u) := -f(x(u))x'(u) \tag{0.3}$$

is the probability density function of U,

We can combine both (0.2) and (0.3) into a single formula: $g(u) = f(x(u)|x'(u)|$. We have proved:

0.1 THEOREM. *Let X be a continuous random variable with values in (a,b), and let f be the probability density function of X. Let u be a continuously differentiable strictly monotone function from (a,b) to (c,d). Define a new random variable $U = u(X)$. Then U has the probability density function g where*

$$g(u) := f(x(u))|x'(u)| \tag{0.4}$$

0.2 EXAMPLE. *Let X be uniform on $(0,1)$ so that $f(x) = 1$ for $x \in (0,1)$. Let $u(x) = -\log(x)$. As x ranges over $(0,1)$, u ranges over $(0,\infty)$, and note that u is strictly monotone decreasing. The inverse function is $x(u) = e^{-u}$, and so $|x'(u)| = e^{-u}$. Defining $U = u(X) = -\log(X)$, we then have that the density of U is the function e^{-u} on $(0,\infty)$. That is, if X is uniform on $(0,1)$, $U = -\log(X)$ is exponential with unit rate on $(0,\infty)$.*

0.2 Several variables

Let $\widehat{\Omega}$ be an open subset of the x,y plane with piecewise smooth boundary. Let $\mathbf{x} = (x,y)$ denote a generic point in the x,y plane. Suppose that $\mathbf{U}(\mathbf{x}) = (u(x,y), v(x,y))$ is a continuously differentiable function defined on $\widehat{\Omega}$ with values in the u,v plane. Suppose further that $\mathbf{U}$ is one-to one on $\widehat{\Omega}$, and let Ω denote the image of $\widehat{\Omega}$ under $\mathbf{U}$. Then $\mathbf{U}$ is an invertible, continuously differentiable transformation from $\widehat{\Omega}$ onto Ω. Let $\mathbf{X}(u,v)$ denote the inverse function.

For $A \subset \Omega$, define $\widehat{A} = \mathbf{U}^{-1}(A)$. Then for any continuous function f on $\widehat{\Omega}$, the change of variables formula for two variables gives us

$$\int_{\widehat{A}} f(x,y)\mathrm{d}x\mathrm{d}y = \int_A f(\mathbf{X}(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \mathrm{d}u\mathrm{d}v \ , \tag{0.5}$$

where $\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|$ is the absolute value of the Jacobian determinant of the transformation:

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \left|\det\begin{bmatrix} \frac{\partial x}{\partial u}(u,v) & \frac{\partial y}{\partial u}(u,v) \\ \\ \frac{\partial x}{\partial v}(u,v) & \frac{\partial y}{\partial v}(u,v) \end{bmatrix}\right| . \tag{0.6}$$

This formula has a probabilisitic interpretation. Suppose that f is the joint probability density function of a pair of random variables (X,Y), where (X,y) takes values in $\widehat{\omega}$. Then left side of (0.5) equals $P((X,Y)\in\widehat{A})$. Define new random variables U and V by $U = u(X,Y)$ and $V = v(X,Y)$. Then by the definition of A,

$$(X,Y)\in\widehat{A} \iff (U,V)\in A\ .$$

and hence the integral on the right in (0.5) equals $P((U,V)\in A)$. It follows that

$$g(u,v) = f(x(u,v),y(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \tag{0.7}$$

is the joint probability density function of U and V. We have proved:

0.3 THEOREM. *Let f be the joint probability density of a pair (X,Y) of random variables taking values in $\widehat{\Omega}\subset\mathbb{R}^2$. Let $\mathbf{U}(\mathbf{x}) = (u(x,y),v(x,y))$ be a continuously differentiable function defined on $\widehat{\Omega}$. Suppose further that $\mathbf{U}$ is one-to one on $\widehat{\Omega}$, and let Ω denote the image of $\widehat{\Omega}$ under $\mathbf{U}$, so that $\mathbf{U}(x,y)$ has a continuously differentiable inverse $\mathbf{X}(u,v)$ defined on Ω. Define a new pair of random variables (U,V) by $U = u(X,Y)$ and $V = v(X,Y)$. Then the function $g(u,v)$ given in (0.7) is the joint probability density of (U,V).*

The generalization to more variables is straightforward.

0.4 EXAMPLE. *Let X and Y be independent and uniform on $(0,1)$. Define new random variables $U := X+Y$ and $V := X/Y$. Find the joint probability density of (U,V)? Are U and V independent?*

To apply the theorem, we note that $f(x,y) = 1$ for $(x,y)\in(0,1)\times(0,1)$ and $f(x,y) = 0$ elsewhere. Next define $u(x,y) = x+y$ and $V(x,y) = x/y$. Then $\mathbf{U}(x,y) = (x+y, x/y)$ which is defined and continuously differentiable on $\widehat{\Omega} = (0,1)\times(0,1)$.

To see that it is invertible, we seek to compute the inverse. Combining $u = x+y$ and $x = vy$ yields $u = y(v+1)$ so that

$$y = \frac{u}{1+v} \qquad \text{and then} \quad x = \frac{uv}{1+v}\ .$$

Hence the inverse transformation is

$$\mathbf{X}(u,v) = (x(u,v),y(u,v)) = \left(\frac{uv}{1+v}\ ,\ \frac{u}{1+v}\right)\ .$$

To find Ω, the domain of $\mathbf{X}$, we note first the by definition $u(x,y)$ and $v(x,y)$ are positive on $\widehat{\Omega} = (0,1)\times(0,1)$. By definition, $(u,v)\in\Omega$ if and only if $(x(u,v),y(u,v))\in\widehat{\Omega}$, which is the same as

$$0 \le \frac{uv}{1+v} \le 1 \qquad \text{and} \qquad 0 \le \frac{u}{1+v} \le 1 \ .$$

The region Ω is therefore bounded by

$$u = 0 \ , \qquad v = 0 \ , \quad v = \frac{1}{u-1} \quad \text{and} \quad v = u - 1 \ .$$

That is, Ω is the union of the rectangle $(0,1)\times(0,\infty)$, and the region above $(1,2)$ with $u-1 < v < \frac{1}{1-u}$.

Finally we compute

$$\begin{bmatrix} \frac{\partial x}{\partial u}(u,v) & \frac{\partial y}{\partial u}(u,v) \\ \\ \frac{\partial x}{\partial v}(u,v) & \frac{\partial y}{\partial v}(u,v) \end{bmatrix} = \begin{bmatrix} \frac{v}{1+v} & \frac{u}{(1+v)^2} \\ \\ \frac{1}{1+v} & \frac{-u}{(1+v)^2} \end{bmatrix}$$

and therefore

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = u(1+v)^{-2} \ .$$

Therefore, the joint density function of (U,V) is $g(u,v)$ given by

$$g(u,v) := \begin{cases} u(1+v)^{-2} & (u,v)\in\Omega \\ 0 & (u,v)\notin\Omega \end{cases} \ .$$

Although the function $u(1+v)^{-2}$ is a product function, $g(u,v)$ is not because Ω is not a rectangle. Therefore, U and V are not *independent. This can be seen without calculation: It is possible for V to be very large, but then Y must be very small, and then U cannot be much greater that 1, while in general U can be as large as 2. Hence U is not independent of V, and then neither is V independent of U.*

We have answered the questions posed at the beginning of the example, but let's check our work. It must be the case that

$$\int_\Omega g(u,v)\mathrm{d}u\mathrm{d}v = 1$$

since otherwise g would not be a probability density.

We compute:

$$\begin{aligned} \int_\Omega g(u,v)\mathrm{d}u\mathrm{d}v &= \int_0^1 u\left(\int_0^\infty (1+v)^{-2}\mathrm{d}v\right)\mathrm{d}u + \int_1^2 u\left(\int_{u-1}^{(u-1)^{-1}} (1+v)^{-2}\mathrm{d}v\right)\mathrm{d}u \\ &= \frac{1}{2} + \int_1^2 u\left(\frac{2-u}{u}\right)\mathrm{d}u = 1 \ . \end{aligned}$$

It is also instructive to compute the left marginal density of g; i.e.,

$$g_U(u) = \int_{\mathbb{R}} g(u,v)\mathrm{d}v \ .$$

There are two cases to consider: For $0 < u < 1$*,*

$$\int_{\mathbb{R}} g(u,v)\mathrm{d}v = u\int_0^\infty (1+v)^{-2}\mathrm{d}v = u\ .$$

for $1 < u < 2$*,*

$$\int_{\mathbb{R}} g(u,v)\mathrm{d}v = u\int_{u-1}^{(u-1)^{-1}} (1+v)^{-2}\mathrm{d}v = 2-u\ .$$

Altogether,

$$g_U(u) = \begin{cases} u & u \in [0,1] \\ 2-u & u \in [1,2] \\ 0 & u \notin [0,2] \end{cases}\ ,$$

which is what we computed for the convolution of two uniform densities on $[0,1]$*.*

We close by computing the right marginal density of g*, I.e.,*

$$g_V(v) = \int_{\mathbb{R}} g(u,v)\mathrm{d}u\ .$$

There are two cases to consider: For $0 < v < 1$*,*

$$\int_{\mathbb{R}} g(u,v)\mathrm{d}u = (1+v)^{-2}\int_0^{1+v} u\mathrm{d}u = \frac{1}{2}\ .$$

for $1 < v < \infty$*,*

$$\int_{\mathbb{R}} g(u,v)\mathrm{d}u = (1+v)^{-2}\int_0^{1+v^{-1}} u\mathrm{d}u = \frac{1}{2v^2}\ .$$

Altogether,

$$g_V(v) = \begin{cases} \frac{1}{2} & v \in [0,1] \\ \frac{1}{2v^2} & v \in [1,\infty) \\ 0 & v < 0 \end{cases}\ .$$

Let's check this last computation: $\log(X/Y) = \log(X) - \log(Y)$*. Define* $W := \log(X)$ *and* $Z := -\log(Y)$*, By Theorem 0.3,*

$$f_W(w) = \begin{cases} e^w & w < 0 \\ 0 & w \geq 0 \end{cases} \quad \text{and} \quad f_Z(z) = \begin{cases} e^{-z} & z > 0 \\ 0 & z \leq 0 \end{cases}.$$

Since W *and* Z *are independent, the density o* $Z+W$ *is given by the convolution:*

$$f_{Z+W}(t) = \int_{\mathbb{R}} f_W(t-z)f_Z(z)\mathrm{d}z$$

and the integrand is non-zero if and only if both $z > 0$ *and* $t - z < 0$*. Hence if* $t > 0$*, we have*

$$\int_{\mathbb{R}} f_W(t-z)f_Z(z)\mathrm{d}z = \int_t^\infty e^{t-z}e^{-z}\mathrm{d}z = \frac{1}{2}e^{-t}\ ,$$

while for $t < 0$*,*

$$\int_{\mathbb{R}} f_W(t-z) f_Z(z) \mathrm{d}z = \int_0^\infty e^{t-z} e^{-z} \mathrm{d}z = \frac{1}{2} e^t \ .$$

Altogether,

$$f_{W+Z}(t) = \frac{1}{2} e^{-|t|} \ .$$

Then since $X/Y = \exp(W+Z)$*, we can apply Theorem 0.3 once more to obtain the density for* X/Y*: If* $v(t) = \exp(t)$*,* $t(v) = \log(v)$ *and then* $t'(v) = 1/v$ *for* $v > 0$*. Theorem 0.3 then gives*

$$f_{X/Y}(v) = f_{W+Z}(t(v)) t'(v) = \frac{1}{2} e^{-|\log v|} \frac{1}{v} \ .$$

For $v < 1$*,* $-|\log v| = \log(v)$ *and so for such* v*,* $f_{X/Y}(v) = \frac{1}{2}$*. For* $v > 1$*,* $-|\log v| = -\log(v)$ *and so for such* v*,* $f_{X/Y}(v) = \frac{1}{2} v^{-2}$*. Altogether,*

$$f_{X/Y}(v) = \begin{cases} \frac{1}{2} & 0 \le v \le 1 \\ \frac{1}{2v^2} & v \ge 1 \\ 0 & v < 0 \ . \end{cases}$$

This is exactly what we found above by computing the right margin of the joint probability density of $(X+Y, X/Y)$*.*