ECE 6010
Lecture 4 – Change of Variables

Reading from G&S: Section 4.7, 4.8, 4.9, 4.10, 4.11

Changing variables: One dimension

A simply invertible function

Let $Y = g(X)$, where X is a continuous r.v. and g is a one-to-one, onto, measurable function. Then

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

So we can determine the distribution of Y. Let us now take a different point of view that will allow us to generalize to higher dimensions and develop and understand a commonly used formula.

Consider an interval along the X axis,

$$P(x \leq X \leq x + dx) \approx f_X(x)dx$$

Suppose the function $g(x)$ has a positive derivative. The interval along the Y axis, when $Y = g(X)$ is $dy \approx \frac{dy}{dx}dx$ at the point x. The probability that X falls in its interval is the probability that Y falls in its interval:

$$P(x \leq X \leq x + dx) = P(y \leq Y \leq y + dy)$$

where $y = g(x)$, or equivalently, $x = g^{-1}(y)$. Then

$$f_X(x)dx \approx f_Y(y)dy$$

That is

$$f_Y(y) = f_X(x)\frac{dx}{dy}\bigg|_{x=g^{-1}(y)} = f_X(g^{-1}(y))\frac{dx}{dy}.$$

If we take the other case that $g(x)$ has a negative derivative, we have to take $f_X(x)dx = f_Y(y)(-dy)$. Combining these together we obtain

$$\boxed{f_Y(y) = f_X(g^{-1}(y))\left|\frac{dx}{dy}\right| = \frac{f_X(g^{-1}(y))}{|dy/dx|} = \frac{f_X(g^{-1}(y)}{|g'(g^{-1}(x))}.}$$

Example 1 Let $Y = aX + b$. Then

$$f_Y(y) = \frac{1}{|a|}f_X((y-b)/a).$$

□

Example 2 Suppose $f_X(x) = \frac{a/\pi}{x^2+a^2}$ (Cauchy). Let $Y = 1/X$. Then

$$f_Y(y) = \frac{1/a\pi}{y^2 + 1/a^2}$$

(Cauchy) □

Example 3 Suppose $X \sim \mathcal{U}(a,b)$, with $0 < a < b$. Then $f_X(x) = \frac{1}{b-a}$ for $x \in [a,b]$. Let $Y = 1/X$. Then

$$f_Y(y) = \frac{1}{(b-a)y^2} \text{ for } \frac{1}{b} < y < \frac{1}{a}.$$

□

Example 4 Let $Y = e^X$. Then

$$f_Y(Y) = \frac{1}{y} f_X(\ln y) \qquad y > 0.$$

If $X \sim \mathcal{N}(\mu, \sigma^2)$ then

$$f_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-(\ln y - \mu)^2/2\sigma^2}$$

This density is **log-normal.** □

Example 5 Suppose $y = g(x) = \tan x$, or $x = \tan^{-1} y$. This has an infinite number of solutions. However, if we take $x \in (-\pi/2, \pi/2)$, then there is a unique inverse. We have

$$g'(x) = \frac{1}{\cos^2 x} = 1 + y^2.$$

Now let $X \sim \mathcal{U}(-\pi/2, \pi/2)$. Then

$$f_Y(y) = \frac{1}{\pi(1+y^2)}$$

(Cauchy). □

Example 6 Suppose X has continuous distribution $F_X(x)$, and

$$Y = F_X(X).$$

That is, the function we use to transform is actually the c.d.f of X. Then

$$g'(x) = F_X'(x) = f_X(x)$$

and

$$f_Y(y) = \frac{f_X(x)}{|g'(x)|} = \frac{f_X(x)}{f_X(x)} = 1 \qquad 0 < y < 1$$

That is, Y is uniformly distributed.

In the image processing literature, this is called *histogram equalization*. □

Example 7 Let $X \sim \mathcal{U}(0,1)$, and let Y have a specified c.d.f $F_Y(y)$, which we take to be continuous. Let

$$g(x) = F_Y^{-1}(x), \qquad 0 < x < 1,$$

so

$$x = F_Y(y).$$

Then

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dx}{dy} \right|$$

Since f_X is 1 for all values of y, and since

$$\left| \frac{dx}{dy} \right| = |f_Y(y)| = f_Y(y),$$

putting all the pieces together we find

$$f_Y(y) = f_Y(y).$$

That is, Y has the desired distribution!

The point of this is that if we can generate $\mathcal{U}(0, 1)$, we can (in principle!) transform it to produce any other continuous distribution. □

Multiple inverses

It may happen that g is not a uniquely invertible function. That is, for a given y there may be more than one value of x such that $y = g(x)$. For example, $y = g(x) = x^2$: then $x = \sqrt{y}$ and $x = -\sqrt{y}$ are both inverses.

We will prove the concept for two solutions, Let $y = g(x_1) = g(x_2)$, assuming to be specific that the slope is positive at x_1 and negative at x_2.

$$P(y < Y < y + dy) = P(x_1 < X < x_1 + dx_1) + P(x_2 + dx_2 < X < x_2)$$

That is

$$f_Y(y)dy = f_X(x_1)dx_1 + f_X(x_1)|dx_2|$$

From this,

$$f_Y(y) = f_X(x_1)\left|\frac{dx_1}{dy}\right| + f_X(x_2)\left|\frac{dx_2}{dy}\right|$$

This is sometimes written

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_1)}{|g'(x_1)|}$$

In general, with n solutions $x_1, x_2, \ldots, x_n$ we have

$$\boxed{f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_1)}{|g'(x_1)|} + \cdots + \frac{f_X(x_n)}{|g'(x_n)|}}$$

Example 8 Suppose $X \sim \mathcal{U}(-\pi, \pi)$ and $Y = a\sin(X + \theta)$. Generally the $\sin$ function has an infinite number of inverses, but there are only inverses in the stated range of X. We have

$$g'(x) = a\cos(x + \theta) = \sqrt{a^2 - y^2}.$$

The density of X is $f_X(x) = \frac{1}{2\pi}$, for $x \in [-\pi, \pi]$. Then

$$f_Y(y) = \frac{1}{2\pi}\frac{1}{\sqrt{a^2 - y^2}} + \frac{1}{2\pi}\frac{1}{\sqrt{a^2 - y^2}} = \frac{1}{\pi\sqrt{a^2 - y^2}}, \qquad |y| < a.$$

□

$g(X)$ constant in an interval

If the function $g(X)$ is constant over any interval, then there is no inverse, nor even multiple inverses. However, we can still compute the distribution. Let $g(x) = y_1$ for $x_0 < x \leq x_1$ (i.e., constant). Then

$$P(Y = y_1) = P(x_0 < X \leq x_1) = F_X(x_1) - F_X(x_0).$$

Hence, there is probability mass at the point y_1. This results in a c.d.f which is not continuous at that point.

Example 9 Suppose $g(x)$ is the limiter:

$$g(x) = \begin{cases} -b & x < -b \\ x & -b \leq x \leq b \\ b & x > b. \end{cases}$$

Then

$$P(Y = -b) = P(X \leq -b) = F_X(-b).$$

$$P(Y = b) = P(X > b) = 1 - F_X(b).$$

For $-b \leq Y < b$, $F_Y(y) = F_X(y)$. □

Example 10 Suppose

$$g(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0. \end{cases}$$

Then

$$P(Y = -1) = P(X \leq 0) = F_X(0)$$

$$P(Y = 1) = p(X > 0) = 1 - F_X(0).$$

We have a two-valued discrete random variable. □

Changing Variables: Multiple dimensions

Consider now multiple variables. Let $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ where

$$\mathbf{Y} = \mathbf{g}(\mathbf{X})$$

We go with the equal probability idea. The probability of falling in a region in $\mathbf{X}$ space should be the same as the probability of falling in the corresponding region in $\mathbf{Y}$ space. We'll draw pictures in two dimensions, but the concepts apply to higher dimensions.

$$P(x_1 < X_1 < x_1 + dx_1, x_2 < X_2 < x_2 + dx_2) \approx f_X(x_1, x_2) dx_1 dx_2$$

Suppose the region $dx_1 dx_2$ maps to the region dA in the y coordinates. Equating probabilities We have

$$f_Y(y_1, y_2) dA = f_X(x_1, x_2) dx_1 dx_2$$

We need to evaluate dA.

The region dA is a parallelepiped described by the vectors

$$\left(\frac{dy_1}{dx_1} dx_1, \frac{dy_2}{dx_1} dx_1\right) \qquad \text{and} \qquad \left(\frac{dy_1}{dx_2} dx_2, \frac{dy_2}{dx_2} dx_2\right)$$

Fact: Recall from calculus that that (signed) area of the parallelepiped described by the vectors $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ is obtained from the cross product. Let us express this in matrix form

$$\mathbf{v} \otimes \mathbf{w} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & 0 \\ w_1 & w_2 & 0 \end{bmatrix}$$

In our case, we have

$$\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{dy_1}{dx_1} dx_1 & \frac{dy_2}{dx_1} dx_1 & 0 \\ \frac{dy_1}{dx_2} dx_2 & \frac{dy_2}{dx_2} dx_2 & 0 \end{bmatrix} = \mathbf{k}\left(\frac{dy_1}{dx_1} dx_1 \frac{dy_2}{dx_2} dx_2 - \frac{dy_2}{dx_1} dx_1 \frac{dy_1}{dx_2} dx_2\right)$$

The (signed) area is then

$$\frac{dy_1}{dx_1}dx_1\frac{dy_2}{dx_2}dx_2 - \frac{dy_2}{dx_1}dx_1\frac{dy_1}{dx_2}dx_2 = (\frac{dy_1}{dx_1}\frac{dy_2}{dx_2} - \frac{dy_2}{dx_1}d\frac{dy_1}{dx_2})dx_1dx_2$$

Let

$$J = \det \begin{bmatrix} \frac{dy_1}{dx_1} & \frac{dy_1}{dx_2} \\ \frac{dy_2}{dx_1} & \frac{dy_2}{dx_2} \end{bmatrix}$$

This matrix of partial derivatives is called the **Jacobian** of the function $\mathbf{g}$.

Back to probabilities. We have

$$f_Y(y_1, y_2)dA = f_X(x_1, x_2)dx_1dx_2$$

or

$$f_Y(y_1, y_2)|J|dx_1dx_2 = f_X(x_1, x_2)dx_1dx_2$$

or

$$f_Y(y_1, y_2) = |J|^{-1}f_X(x_1, x_2)$$

or, in general for an invertible function $\mathbf{g}$,

$$\boxed{f_Y(\mathbf{y}) = |J|^{-1}f_X(\mathbf{g}^{-1}(\mathbf{y}))}$$

Example 11 Box-Muller transformation. Let $X_1 \sim \mathcal{U}(0,1)$ and $X_2 \sim \mathcal{U}(0,1)$ (independent). Let

$$Y_1 = \sqrt{-2\ln X_1}\cos 2\pi X_2 \qquad Y_2 = \sqrt{-2\ln X_1}\sin 2\pi X_2$$

Then

$$Y_1 \sim \mathcal{N}(0,1) \qquad Y_2 \sim \mathcal{N}(0,1).$$

□

Many-to-one mappings

Let $Y = g(X_1, X_2)$ where X_1 and X_2 are jointly distributed random variables. For a given value of y, the inverse may form a curve in (x_1, x_2) space. Let A_y denote the region in the X_1X_2 plane such that $g(X_1, X_2) \leq y$. (This may not be a connected region.) Then

$$\{Y \leq y\} = \{g(X_1, X_2) \leq y\} = \{(X_1, X_2) \in A_y\}$$

so

$$F_Y(y) = \int\int_{A_y} f_{X_1X_2}(x_1, x_2)dx_1dx_2$$

Let ΔA_y denote the region of the X_1X_2 plane such that $y < g(x_1, x_2) \leq y + dy$. then

$$f_Y(y)dy = \int\int_{\Delta A_y} f_{X_1X_2}(x_1, x_2)dx_1dx_2$$

Example 12 Let $Z = X + Y$. The region in the xy plane such that $x + y \leq z$ is the part of plane below the line $x + y = z$. We have

$$F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{XY}(x, y)dxdy$$

Differentiating this with respect to z we have

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(z-y,y)dy.$$

Independence... □

Example 13 Let $Z = X/Y$. The region of the plane such that $x/y \leq z$ can be determined as follows. Fix z. For $y > 0$ we want the region where $x \leq yz$, and for $y < 0$ we want the region where $x \geq yz$. We obtain

$$F_Z(z) = \int_0^{\infty} \int_{-\infty}^{yz} f_{XY}(x,y)dxdy + \int_{-\infty}^{0} \int_{zy}^{\infty} f_{XY}(x,y)dxdy.$$

To get the density: The region ΔA_z is the triangular sector bounded by the lines $x = yz$ and $x = y(z+dz)$. The coordinates of a point are $x = zy$ and y. The area of a differential is $|y|dydz$. We obtain

$$f_Z(z)dz = \int_{-\infty}^{\infty} f_{XY}(zy,y)|y|dydz$$

Then cancel dz:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(zy,y)|y|dy.$$

□

Example 14 Let $Z = \sqrt{X^2+Y^2}$. The region A_z is the circle $x^2 + y^2 \leq z^2$. If $f_{XY}(x,y) = \frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}$ we have

$$F_Z(z) = \frac{1}{2\pi\sigma^2}\int_0^z re^{-r^2/2\sigma^2}\,dr = 1 - e^{-z^2/2\sigma^2} \qquad z > 0.$$

so

$$f_Z(z) = \frac{z}{\sigma^2}e^{-z^2/2\sigma^2},$$

$z > 0$. This is a *Rayleigh* distribution.

□

Sometimes it is helpful to introduce an auxiliary variable, then integrate it out.

Suppose $Z = XY$. Introduce the auxiliary variable $W = X$. Then the inverse functions are straightforward to compute,

$$X = W \text{ and } Y = Z/W.$$

The Jacobian is

$$J = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} = -x = -w.$$

The joint density is

$$f_{ZW}(z,w) = \frac{1}{|w|}f_{XY}(w,z/w).$$

Then the density of Z can be obtained by integration:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{|w|}f_{XY}(w,z/w)dw.$$