MATH 4752/6752 – Mathematical Statistics II

Functions of Random Variables

Xiaojing Ye
Department of Mathematics & Statistics
Georgia State University

Let $X_1, \ldots, X_n$ be random variables (may not be iid) and $Y = u(X_1, \ldots, X_n)$ where u is a given function.

Question: what is the probability distribution/density of Y?

Three typical techniques:

- Probability distribution function technique.
- Transformation technique.
- Moment generating function technique.

Probability distribution function technique

If $X_1, \ldots, X_n$ are continuous RVs, then the probability distribution function of Y is

$$G(y) = \mathsf{P}(Y \le y) = \mathsf{P}(u(X_1, \ldots, X_n) \le y)$$

and the probability density function of Y is

$$g(y) = G'(y).$$

Example. If the pdf of X is

$$f(x) = \begin{cases} 6x(1-x), & \text{if } 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find the probability density function of $Y = X^3$.

Solution. Note that $X \in [0, 1]$, so $Y \in [0, 1]$. Denote $G(y)$ the distribution function of Y, then for any $y \in [0, 1]$, we have

$$\begin{aligned} G(y) &= \mathsf{P}(Y \le y) = \mathsf{P}(X^3 \le y) = \mathsf{P}(X \le y^{1/3}) \\ &= \int_0^{y^{1/3}} 6x(1-x)\,dx = 3y^{2/3} - 2y. \end{aligned}$$

Therefore, the pdf of Y is

$$g(y) = G'(y) = 2y^{-1/3} - 2$$

for $y \in [0, 1]$. If $y < 0$ or $y > 1$, then $g(y) = 0$.

Example. Let X be a random variable with pdf $f(x)$ and $Y = |X|$. Show that the pdf of Y is

$$g(y) = \begin{cases} f(y) + f(-y), & \text{if } y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. For $y > 0$, we have

$$\begin{aligned} G(y) &= \mathsf{P}(Y \le y) = \mathsf{P}(|X| \le y) = \mathsf{P}(-y \le X \le y) \\ &= \mathsf{P}(X \le y) - \mathsf{P}(X \le -y) \\ &= F(y) - F(-y). \end{aligned}$$

So $g(y) = G'(y) = f(y) + f(-y)$.

Example. Use the previous result to find the pdf of $Y = |X|$ where X is the standard normal RV.

Solution. Let $n(x; 0, 1)$ denote the pdf of X, i.e.,

$$n(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Then

$$g(y) = n(y; 0, 1) + n(-y; 0, 1) = 2n(y; 0, 1) = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-y^2/2}.$$

Example. Suppose the joint pdf of (X_1, X_2) is

$$f(x_1, x_2) = \begin{cases} 6e^{-3x_1 - 2x_2}, & \text{if } x_1, x_2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Find the pdf of $Y = X_1 + X_2$.

Solution. We notice that $Y \geq 0$. Furthermore,

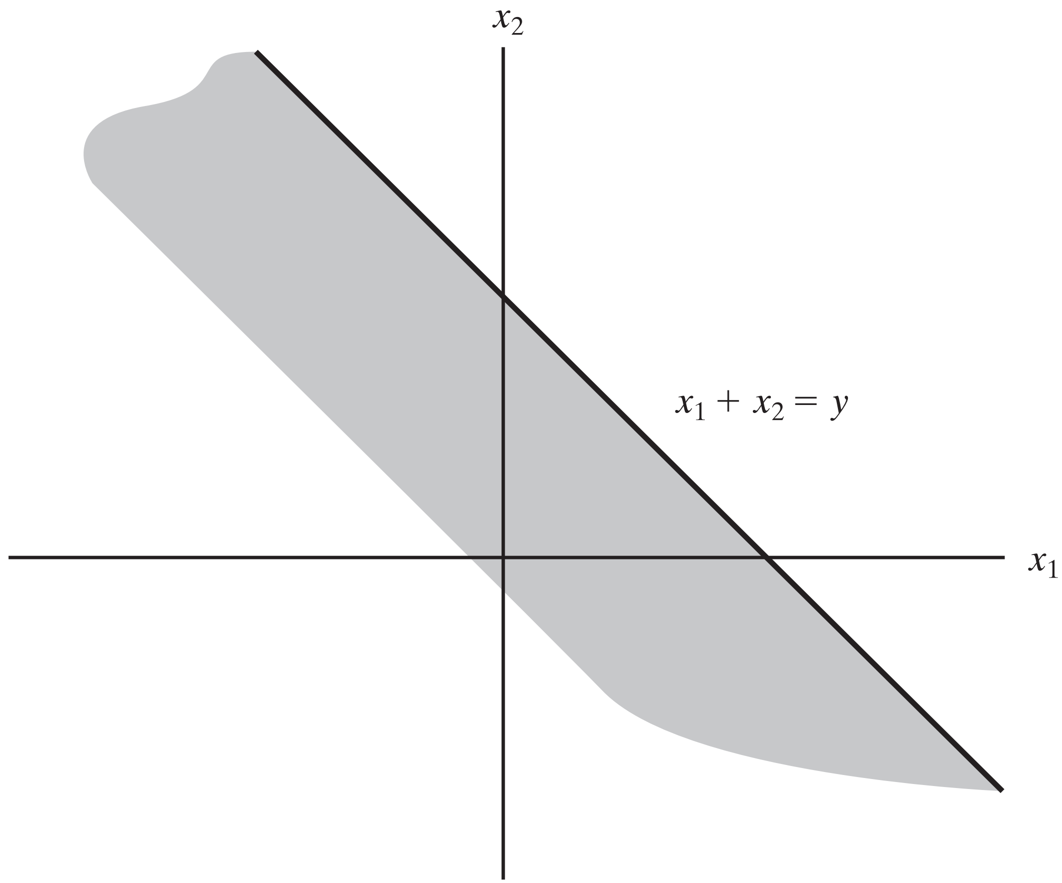
$$\begin{aligned} G(y) &= \mathsf{P}(Y \leq y) = \mathsf{P}(X_1 + X_2 \leq y) = \int_0^y \int_0^{y-x_2} 6e^{-3x_1 - 2x_2}\, dx_1 dx_2 \\ &= \cdots = 1 + 2e^{-3y} - 3e^{-2y}. \end{aligned}$$

Hence

$$g(y) = G'(y) = -6e^{-3y} + 6e^{-2y}$$

for $y \geq 0$. If $y < 0$, then $g(y) = 0$.

Remark. For each $y > 0$, the value of $G(y)$ is the integral of $f(x_1, x_2)$ over the triangle (intersection of the shaded area and the first quadrant).



Transformation technique: one variable case

We first consider the discrete case where u is one-to-one.

Example. Let X be the number of heads by tossing a fair coin for 4 times. Then the pmf f of X is

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Find the pmf g of $Y = \frac{1}{1+X}$.

Solution. Since $u(x) = \frac{1}{1+x}$ is one-to-one (because u is strictly monotone) for $x > 0$, we have

y	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
$g(y)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Remark. Notice that $f(x) = \frac{1}{16}\binom{4}{x}$, we have $g(y) = \frac{1}{16}\binom{4}{\frac{1}{y}-1}$.

Now we consider the discrete case where u is not one-to-one.

Example. Let X be the same as above and $Z = (X-2)^2$. Then

x	0	1	2	3	4
$z = (x-2)^2$	4	1	0	1	4

So we know the pmf h of Z is

$$h(0) = f(2) = \frac{6}{16}$$
$$h(1) = f(1) + f(3) = \frac{8}{16}$$
$$h(4) = f(0) + f(4) = \frac{2}{16}$$

We consider the continuous RV case where u is strictly monotone (increasing or decreasing) in supp(f).

Theorem. Let X be continuous with pdf $f(x)$ and $Y = u(X)$ where u is strictly monotone in supp(f), then pdf $g(y)$ of Y is

$$g(y) = f(w(y))|w'(y)|$$

where w is the inverse of u (i.e., $y = u(x)$ iff $x = w(y)$).

Proof. We first consider the case where u is strictly increasing. In this case $u'(x) \geq 0$ and $w'(y) \geq 0$, and

$$G(y) = \mathsf{P}(Y \leq y) = \mathsf{P}(u(X) \leq y) = \mathsf{P}(X \leq w(y)) = F(w(y)).$$

Hence

$$g(y) = G'(y) = f(w(y))w'(y) = f(w(y))|w'(y)|.$$

Proof (cont). If u is strictly decreasing, then $u'(x) \leq 0$ and $w'(y) \leq 0$. Then

$$\begin{aligned} G(y) &= \mathsf{P}(Y \leq y) \\ &= \mathsf{P}(u(X) \geq y) \\ &= 1 - \mathsf{P}(X \leq w(y)) \\ &= 1 - F(w(y)). \end{aligned}$$

Hence

$$g(y) = G'(y) = -f(w(y))w'(y) = f(w(y))|w'(y)|.$$

In summary, we have

$$g(y) = f(w(y))|w'(y)|.$$

Example. Let X be the standard exponent RV. Find the pdf of $Y = \sqrt{X}$.

Solution. Since $u(x) = \sqrt{x}$ is strictly increasing when $x \geq 0$, we know its inverse w exists and $w(y) = y^2$.

Recall that the pdf of X is $f(x) = e^{-x}$ for $x \geq 0$, we have

$$g(y) = f(w(y))|w'(y)| = 2ye^{-y^2}$$

for $y \geq 0$. If $y < 0$, then $g(y) = 0$.

Example. If $F(x)$ is the distribution function of the continuous RV X. Find the pdf of $Y = F(X)$.

Solution. Notice that F is strictly increasing in supp(f). Let w be the inverse of F, then $F(w(y)) = y$ for all y. Hence

$$F'(w(y))w'(y) = 1.$$

This implies

$$w'(y) = \frac{1}{F'(w(y))} = \frac{1}{f(w(y))}$$

So we have

$$g(y) = f(w(y))|w'(y)| = f(w(y))\frac{1}{f(w(y))} = 1.$$

This means that Y is the uniform RV on $[0, 1]$.

Example. Let X be the standard normal random variable. Find the pdf of $Z = X^2$.

Solution. Since x^2 is not monotone, we turn to consider $y = |x|$ and then $z = y^2$ is monotone since $y \geq 0$.

From a previous example, we know $g(y) = 2n(y; 0, 1) = \frac{2}{\sqrt{2\pi}}e^{-y^2/2}$ for $y \geq 0$. Let $u(y) = y^2$ and $w(z) = \sqrt{z}$ be the inverse of u. Then

$$h(z) = g(w(z))|w'(z)| = \frac{1}{\sqrt{2\pi z}}e^{-z/2}$$

for $z \geq 0$.

Z is said to follow the chi-square distribution with degree of freedom 1.

Transformation technique with several random variables

Let $X_1, \ldots, X_n$ be continuous RVs with joint pdf $f(x_1, \ldots, x_n)$. What is the pdf of $Y = u(X_1, \ldots, X_n)$ for a given function $u : \mathbb{R}^n \to \mathbb{R}$?

Here is a general strategy:

- Step 1: Set $(Y_1, \dots, Y_n) = \boldsymbol{u}(X_1, \dots, X_n)$ where $\boldsymbol{u}$ is one-to-one on supp(f), and $Y = Y_1$. So the first component of $\boldsymbol{u}$ is u.

- Step 2: Find the joint pdf of $\boldsymbol{Y} = (Y_1, \dots, Y_n)$ which is

$$g_{\boldsymbol{Y}}(y_1, \dots, y_n) = f(\boldsymbol{w}(y_1, \dots, y_n))|J(y_1, \dots, y_n)|,$$

  where $\boldsymbol{w}$ is the inverse of $\boldsymbol{u}$, and $J(y_1, \dots, y_n)$ is the Jacobian matrix of $\boldsymbol{w}$, i.e., $J(y_1, \dots, y_n) = D\boldsymbol{w}(y_1, \dots, y_n)$.

- Find $g(y)$ as the marginal pdf of $Y = Y_1$ by taking integral:

$$g(y) = \int g_{\boldsymbol{Y}}(y, y_2, \dots, y_n)\, dy_2 \cdots dy_n.$$

Remark. This strategy is based on the change of variable in multivariable calculus: for any $E \subset \mathbb{R}^n$, we have

$$\begin{aligned} \mathsf{P}(\boldsymbol{Y} \in E) &= \mathsf{P}(\boldsymbol{u}(\boldsymbol{X}) \in E) \\ &= \mathsf{P}(\boldsymbol{X} \in \boldsymbol{w}(E)) \\ &= \int_{\boldsymbol{w}(E)} f(\boldsymbol{x}) \, d\boldsymbol{x} \\ &= \int_E f(\boldsymbol{w}(\boldsymbol{y})) |J(\boldsymbol{y})| \, d\boldsymbol{y}. \end{aligned}$$

Hence we have

$$g(\boldsymbol{y}) = f(\boldsymbol{w}(\boldsymbol{y})) |J(\boldsymbol{y})|.$$

Example. Let (X_1, X_2) have joint density

$$f(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)}, & \text{if } x_1, x_2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Find the pdf of $Y = \frac{X_1}{X_1+X_2}$.

Solution. Let $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1+X_2}$. (So $Y = Y_2$.) We define

$$\boldsymbol{u}(\boldsymbol{x}) = \begin{pmatrix} x_1 + x_2 \\ \frac{x_1}{x_1+x_2} \end{pmatrix}.$$

Then $\boldsymbol{Y} = \boldsymbol{u}(\boldsymbol{X})$. We deduce that $X_1 = Y_1 Y_2$ and $X_2 = Y_1 - Y_1 Y_2$. Therefore the inverse of $\boldsymbol{u}$ is

$$\boldsymbol{w}(\boldsymbol{y}) = \begin{pmatrix} y_1 y_2 \\ y_1 - y_1 y_2 \end{pmatrix}.$$

Solution (cont). Now we compute the Jacobian matrix:

$$D\boldsymbol{w}(\boldsymbol{y}) = \begin{pmatrix} y_2 & y_1 \\ 1 - y_2 & -y_1 \end{pmatrix}$$

and thus

$$|J(\boldsymbol{y})| = |\det(D\boldsymbol{w}(\boldsymbol{y}))| = |-y_1 y_2 - y_1 + y_1 y_2| = |-y_1| = y_1.$$

So the joint pdf of (Y_1, Y_2) is

$$g(y_1, y_2) = f(\boldsymbol{w}(\boldsymbol{y}))|J(\boldsymbol{y})| = y_1 e^{-y_1}, \quad (y_1, y_2) \in (0, \infty) \times (0, 1).$$

The marginal pdf h of Y_2 is

$$h(y_2) = \int_0^\infty g(y_1, y_2)\, dy_1 = \int_0^\infty y_1 e^{-y_1}\, dy_1 = 1$$

for any $y_2 \in (0, 1)$. Hence $Y = Y_2$ follows the standard uniform distribution.

Example. Let (X_1, X_2) be uniformly distributed in $(0,1)^2$. Find the pdf of $Y = X_1 + X_2$.

Solution. Let $Y = X_1 + X_2$, $Z = X_2$ and $\boldsymbol{u}(x_1, x_2) = (x_1 + x_2, x_2)$. Hence

$$(x_1, x_2) = \boldsymbol{w}(y, z) = (y - z, z)$$

Therefore the Jacobian is

$$D\boldsymbol{w}(y, z) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

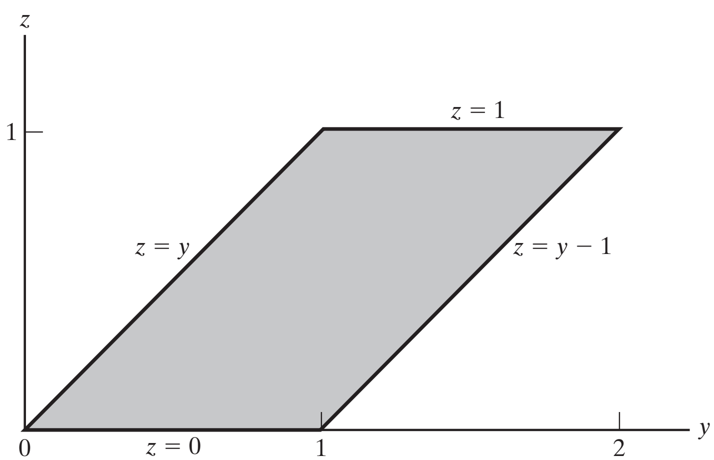
Hence $|J(y, z)| = |\det(D\boldsymbol{w}(y, z))| = 1$. The joint pdf of (Y, Z) is

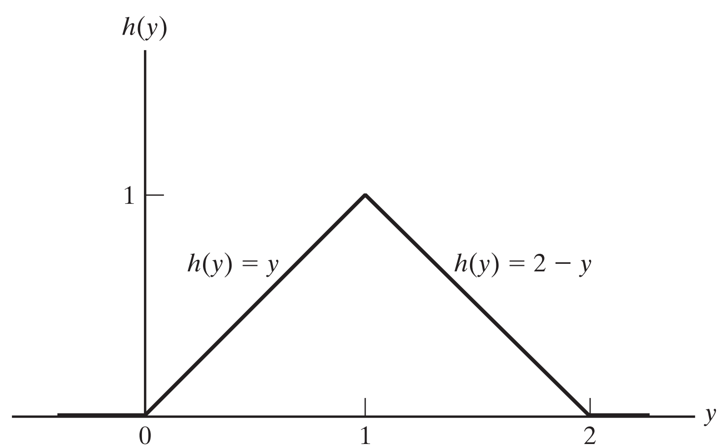
$$g(y, z) = f(\boldsymbol{w}(y, z))|J(y, z)| = 1, \quad 0 < z < 1,\ 0 < y - z < 1.$$

Solution (cont). Thus the marginal pdf h of Y is

$$h(y) = \begin{cases} 0, & \text{for } y \le 0 \\ \int_0^y 1\,dz = y, & \text{for } 0 < y < 1 \\ \int_{y-1}^1 1\,dz = 2 - y, & \text{for } 1 < y < 2 \\ 0, & \text{for } y \ge 2 \end{cases}$$

Left figure: shaded area is the support of the joint pdf $g(y, z)$.
Right figure: the marginal pdf h of Y.





Example. Let (X_1, X_2, X_3) be RVs with joint pdf f as follows:

$$f(x_1, x_2, x_3) = \begin{cases} e^{-(x_1+x_2+x_3)}, & \text{if } x_1, x_2, x_3 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_2$, $Y_3 = X_3$. Find the marginal pdf of Y_1.

Solution. We have

$$\boldsymbol{y} = \boldsymbol{u}(\boldsymbol{x}) = (x_1 + x_2 + x_3, x_2, x_3)$$
$$\boldsymbol{x} = \boldsymbol{w}(\boldsymbol{y}) = (y_1 - y_2 - y_3, y_2, y_3)$$

Hence there is

$$J(\boldsymbol{y}) = \det(D\boldsymbol{w}(\boldsymbol{y}))| = \left| \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = 1.$$

Solution (cont). So $g(y_1, y_2, y_3) = f(\boldsymbol{w}(\boldsymbol{y}))|J(\boldsymbol{y})| = e^{-y_1}$ for $y_2, y_2 > 0$ and $y_1 > y_2 + y_3$.

The marginal pdf of Y_1 is

$$h(y) = \int_0^{y_1} \int_0^{y_1 - y_2} g(y_1, y_2, y_3)\, dy_3 dy_2 = \cdots = \frac{1}{2} y_1^2 e^{-y_1}.$$

Moment generating function (MGF) technique

Let $X_1, \dots, X_n$ be **independent** RVs. What is the pmf/pdf of

$$Y = X_1 + \cdots + X_n.$$

Since pmf/pdf is uniquely determined by MGF, it suffices to find the mgf $M_Y(t)$ of Y, and then check which pmf/pdf this $M_Y(t)$ corresponds to.

Theorem. The mgf of $Y = X_1 + \cdots + X_n$, where $X_1, \ldots, X_n$ are independent RVs with MGFs $M_{X_1}(t), \ldots, M_{X_n}(t)$ respectively, is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

Proof. Since $X_1, \ldots, X_n$ are independent, we know their joint pdf is

$$f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n)$$

where f_i is the pmf/pdf of X_i. Therefore, we have

$$\begin{aligned} M_Y(t) &= \mathbb{E}[e^{Yt}] = \mathbb{E}[e^{(X_1 + \cdots + X_n)t}] \\ &= \int e^{(x_1 + \cdots + x_n)t} f(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \\ &= \int e^{x_1 t} f_1(x_1) \, dx_1 \cdots \int e^{x_n t} f_n(x_n) \, dx_n \\ &= \mathbb{E}[e^{X_1 t}] \cdots \mathbb{E}[e^{X_n t}] \\ &= M_{X_1}(t) \cdots M_{X_n}(t). \end{aligned}$$

Remark. This result can be easily extended to $Y = c_1 X_1 + \cdots + c_n X_n$ where $X_1, \ldots, X_n$ are independent RVs and $c_1, \ldots, c_n$ are constants by noting that

$$M_{cX}(t) = \mathbb{E}[e^{cXt}] = M_X(ct).$$

Remark. The MGF technique may not work when $X_1, \ldots, X_n$ are not independent.

Example. Suppose $X_1, \dots, X_n$ are independent Poisson RVs with parameters $\lambda_1, \dots, \lambda_n$ respectively. Find the distribution of $Y = X_1 + \cdots + X_n$.

Solution. Since $M_{X_i}(t) = e^{\lambda_i(e^t-1)}$ for each i, we have

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = e^{(\sum_{i=1}^{n} \lambda_i)(e^t-1)}$$

which is the mgf of the Poisson RV with parameter $\sum_{i=1}^{n} \lambda_i$. Hence Y is Poisson with parameter $\sum_{i=1}^{n} \lambda_i$.

Example. Suppose $X_1, \ldots, X_n$ are independent exponential RVs with the same parameter θ. Find the distribution of $Y = X_1 + \cdots + X_n$.

Solution. Since $M_{X_i}(t) = (1 - \theta t)^{-1}$ for each i, we have

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = (1 - \theta t)^{-n}$$

which is the mgf of $\Gamma(n, \theta)$. Hence Y is Gamma with parameter (n, θ).

Remark. Recall that the mgf of $Z \sim \Gamma(\alpha, \beta)$ is $M_Z(t) = (1 - \beta t)^{-\alpha}$.