

Chapter 8A: Transformations and Convolutions

MATH/STAT 394: Probability I

Arman Jahangiri

University of Washington
Department of Mathematics

Summer 2026

By the end of this lecture, students should be able to:

- ▶ compute distributions of transformed random variables;
- ▶ apply the change-of-variables formula;
- ▶ define convolutions;
- ▶ compute distributions of sums of independent random variables;

Transformations of Random Variables

Why Transform Random Variables?

Transformations appear everywhere in probability and statistics.

Examples:

- ▶ unit conversion;
- ▶ standardization;
- ▶ logarithms;
- ▶ sums and averages;
- ▶ squares and absolute values;
- ▶ maxima and minima.

Main question

If

$$Y = g(X),$$

what is the distribution of Y ?

Transformations in the Discrete Case

If X is discrete and

$$Y = g(X),$$

then, for every possible value y of Y ,

$$\mathbb{P}(Y = y) = \mathbb{P}(X \in g^{-1}(\{y\})) = \sum_{x: g(x)=y} \mathbb{P}(X = x).$$

In words: find all values of x that are mapped to y by g , and add their probabilities.

Example. Suppose

$$\mathbb{P}(X = -2) = 0.1, \quad \mathbb{P}(X = -1) = 0.2, \quad \mathbb{P}(X = 0) = 0.3, \quad \mathbb{P}(X = 1) = 0.25, \quad \mathbb{P}(X = 2) = 0.15,$$

and let

$$Y = X^2.$$

Then $Y \in \{0, 1, 4\}$, and

$$\mathbb{P}(Y = 0) = \mathbb{P}(X = 0) = 0.30,$$

$$\mathbb{P}(Y = 1) = \mathbb{P}(X \in \{-1, 1\}) = \mathbb{P}(X = -1) + \mathbb{P}(X = 1) = 0.45,$$

$$\mathbb{P}(Y = 4) = \mathbb{P}(X \in \{-2, 2\}) = \mathbb{P}(X = -2) + \mathbb{P}(X = 2) = 0.25.$$

Transformations of Continuous Random Variables

Change of Variables Theorem (Univariate Case)

Let X be a continuous random variable with PDF f_X and support S_X , and let

$$Y = g(X),$$

where g is differentiable and strictly monotone on S_X .

Then g has an inverse, and the support of Y is

$$S_Y = g(S_X).$$

For $y \in S_Y$,

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

and $f_Y(y) = 0$ for $y \notin S_Y$.

How to use it: solve for x in terms of y , find $|dx/dy|$, substitute into $f_X(x)$, and determine the support of Y .

Note: $x = g^{-1}(y)$, so by $\frac{dx}{dy}$, we mean $\frac{d}{dy} g^{-1}(y)$.

Important

The theorem requires g to be one-to-one on the support of X . When it is not, we can break the support into monotone cases, or follow the same steps of the proof and adjust the step for isolating x appropriately (examples below).

Example 1: A Monotone Transformation

Let

$$X \sim N(0, 1), \quad Y = e^X.$$

Since

$$g(x) = e^x$$

is strictly increasing on $\mathbb{R}$, the change-of-variables theorem applies.

Step 1: Find the support of Y .

Since $e^x > 0$ for every x ,

$$S_Y = (0, \infty).$$

Step 2: Solve $y = g(x)$ for x .

$$y = e^x \implies x = \log y.$$

Thus

$$g^{-1}(y) = \log y, \quad \left| \frac{dx}{dy} \right| = \frac{1}{y}.$$

Step 3: Apply the formula.

Since

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

we obtain

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y\sqrt{2\pi}} \exp \left\{ -\frac{(\log y)^2}{2} \right\}, \quad y > 0.$$

For $y \leq 0$, $f_Y(y) = 0$.

This is the **log-normal distribution**.

Example 2: When the Transformation Is Not One-to-One

Let

$$X \sim N(0, 1), \quad Y = X^2.$$

Here

$$g(x) = x^2$$

is **not one-to-one** on $\mathbb{R}$:

$$g(-x) = g(x).$$

Therefore, we cannot define a single inverse g^{-1} on all of $\mathbb{R}$. Instead, use the **CDF method**.

Since $Y = X^2 \geq 0$, the support is

$$S_Y = [0, \infty).$$

For $y > 0$,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) = 2\Phi(\sqrt{y}) - 1. \quad (1)$$

Differentiate:

$$f_Y(y) = \frac{d}{dy} [2\Phi(\sqrt{y}) - 1] = 2\varphi(\sqrt{y}) \frac{1}{2\sqrt{y}} = \frac{\varphi(\sqrt{y})}{\sqrt{y}}, \quad y > 0. \quad (2)$$

Therefore, using

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

the PDF of Y is

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{-y/2}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Change of Variables Theorem (Multivariate Case)

Let $\mathbf{X} = (X_1, \dots, X_n)^T$ be a continuous random vector with joint PDF $f_{\mathbf{X}}$ and support $S_{\mathbf{X}} \subseteq \mathbb{R}^n$. Define $\mathbf{Y} = g(\mathbf{X})$, where $g : S_{\mathbf{X}} \rightarrow S_{\mathbf{Y}}$ is one-to-one and differentiable, with differentiable inverse $\mathbf{x} = g^{-1}(\mathbf{y})$. Then the support of $\mathbf{Y}$ is $S_{\mathbf{Y}} = g(S_{\mathbf{X}})$, and

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) \left| \det \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right) \right|, \quad \mathbf{y} \in S_{\mathbf{Y}}.$$

Here

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}$$

is the **Jacobian matrix**. For $\mathbf{y} \notin S_{\mathbf{Y}}$, $f_{\mathbf{Y}}(\mathbf{y}) = 0$.

Example: A Transformation of Two Variables

Let X_1, X_2 be independent $\text{Exp}(1)$ random variables, so

$$f_{X_1, X_2}(x_1, x_2) = e^{-(x_1 + x_2)}, \quad x_1 > 0, x_2 > 0.$$

Define

$$Y_1 = X_1 + X_2, \quad Y_2 = \frac{X_1}{X_1 + X_2}.$$

Question: What is $f_{Y_1, Y_2}(y_1, y_2)$?

Since

$$y_2 = \frac{x_1}{x_1 + x_2} = \frac{x_1}{y_1},$$

we obtain

$$x_1 = y_1 y_2, \quad x_2 = y_1(1 - y_2).$$

We now find the support:

Since $x_1 > 0$ and $x_2 > 0$, we have $y_1 y_2 > 0$, $y_1(1 - y_2) > 0$, which gives

$$y_1 > 0, \quad 0 < y_2 < 1.$$

We now find the Jacobian:

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{pmatrix} y_2 & y_1 \\ 1 - y_2 & -y_1 \end{pmatrix},$$

so

$$\left| \det \left(\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right) \right| = |-y_1| = y_1.$$

Example: Finding the Joint PDF

By the change of variables theorem,

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(y_1 y_2, y_1(1 - y_2)) \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right|.$$

Since

$$x_1 + x_2 = y_1 y_2 + y_1(1 - y_2) = y_1,$$

we have

$$f_{X_1, X_2}(y_1 y_2, y_1(1 - y_2)) = e^{-y_1}.$$

Therefore,

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} y_1 e^{-y_1}, & y_1 > 0, \quad 0 < y_2 < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that

$$f_{Y_1, Y_2}(y_1, y_2) = \underbrace{y_1 e^{-y_1}}_{f_{Y_1}(y_1)} \underbrace{1}_{f_{Y_2}(y_2)}.$$

Hence,

$$Y_1 \sim \text{Gamma}(2, 1), \quad Y_2 \sim \text{Unif}(0, 1), \quad Y_1 \perp Y_2.$$

Convolutions

What Is a Convolution?

A convolution is the distribution of a sum of independent random variables.

If

$$T = X + Y,$$

and X, Y are independent, we want the distribution of T .

Why sums matter

Sums appear everywhere:

- ▶ totals;
- ▶ averages;
- ▶ waiting times;
- ▶ counting processes.

Theorem

If X and Y are independent discrete random variables and

$$T = X + Y,$$

then

$$\mathbb{P}(T = t) = \sum_x \mathbb{P}(X = x) \mathbb{P}(Y = t - x).$$

Proof.

For a fixed value t , the event

$$\{X + Y = t\}$$

can be written as the disjoint union

$$\{X + Y = t\} = \bigcup_x \{X = x, Y = t - x\}.$$

Therefore,

$$\mathbb{P}(T = t) = \sum_x \mathbb{P}(X = x, Y = t - x).$$

Since X and Y are independent

Theorem

If X and Y are independent continuous random variables and

$$T = X + Y,$$

then

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x)f_Y(t-x) dx.$$

Proof of the Continuous Convolution Formula

Since

$$T = X + Y,$$

the CDF of T is

$$F_T(t) = \mathbb{P}(T \leq t) = \mathbb{P}(X + Y \leq t).$$

Using the joint density,

$$F_T(t) = \iint_{\{(x,y): x+y \leq t\}} f_{X,Y}(x,y) \, dx \, dy.$$

For each fixed x , the condition

$$x + y \leq t$$

is equivalent to

$$y \leq t - x.$$

Therefore,

$$F_T(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{t-x} f_{X,Y}(x,y) \, dy \, dx.$$

Since X and Y are independent,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y),$$

so

$$F_T(t) = \int_{-\infty}^{\infty} f_X(x) \left[\int_{-\infty}^{t-x} f_Y(y) \, dy \right] dx.$$

Differentiating with respect to t ,

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) \, dx.$$

Proof of Convolution Using the Jacobian Formula

Let

$$T = X + Y, \quad U = X.$$

Then the transformation is

$$t = x + y, \quad u = x.$$

Solving for the original variables gives

$$x = u, \quad y = t - u.$$

The Jacobian of the inverse transformation is

$$\frac{\partial(x, y)}{\partial(t, u)} = \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial u} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix},$$

and therefore

$$\left| \det \left(\frac{\partial(x, y)}{\partial(t, u)} \right) \right| = 1.$$

By the change-of-variables formula,

$$f_{T,U}(t, u) = f_{X,Y}(u, t - u).$$

Since X and Y are independent,

$$f_{X,Y}(u, t - u) = f_X(u)f_Y(t - u).$$

Hence,

$$f_{T,U}(t, u) = f_X(u)f_Y(t - u).$$

Proof of Convolution Using the Jacobian Formula (continued)

We have found the joint PDF

$$f_{T,U}(t, u) = f_X(u)f_Y(t - u).$$

To obtain the marginal PDF of

$$T = X + Y,$$

integrate out U :

$$\begin{aligned} f_T(t) &= \int_{-\infty}^{\infty} f_{T,U}(t, u) du \\ &= \int_{-\infty}^{\infty} f_X(u)f_Y(t - u) du. \end{aligned}$$

Therefore,

$$f_T(t) = \int_{-\infty}^{\infty} f_X(u)f_Y(t - u) du.$$

Renaming the dummy variable u as x gives the previous convolution formula:

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x)f_Y(t - x) dx.$$

Example 1: Sum of Two Exponential Random Variables

Let

$$X, Y \stackrel{\text{i.i.d.}}{\sim} \text{Exp}(\lambda), \quad T = X + Y.$$

Recall that

$$f_X(x) = f_Y(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Since $X, Y > 0$, we immediately know

$$T = X + Y > 0.$$

For independent continuous random variables, the convolution formula is

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) dx.$$

For the integrand to be nonzero, we need *both*

$$x > 0 \quad \text{and} \quad t - x > 0.$$

Thus, $0 < x < t$, which is possible only when $t > 0$.

Therefore, for $t > 0$,

$$f_T(t) = \int_0^t f_X(x) f_Y(t-x) dx = \int_0^t \lambda e^{-\lambda x} \lambda e^{-\lambda(t-x)} dx = \lambda^2 e^{-\lambda t} \int_0^t dx = \lambda^2 t e^{-\lambda t}. \quad (3)$$

Hence,

$$f_T(t) = \begin{cases} \lambda^2 t e^{-\lambda t}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

Thus, $T = X + Y \sim \text{Gamma}(2, \lambda)$.

Example 2: Sum of Two Uniform Random Variables

Let

$$X, Y \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(0, 1), \quad T = X + Y.$$

Recall that

$$f_X(x) = f_Y(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Since

$$0 < X < 1, \quad 0 < Y < 1,$$

we know

$$0 < T < 2.$$

The convolution formula gives

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) dx.$$

The integrand is nonzero exactly when

$$0 < x < 1 \quad \text{and} \quad 0 < t-x < 1.$$

The second inequality gives

$$t-1 < x < t.$$

Hence x must satisfy both

$$0 < x < 1 \quad \text{and} \quad t-1 < x < t.$$

Therefore,

$$\max(0, t-1) < x < \min(1, t).$$

So the integration bounds depend on t .

Example 2: Determining the PDF

We found that

$$f_T(t) = \int_{\max(0, t-1)}^{\min(1, t)} 1 \, dx, \quad 0 < t < 2.$$

Now consider the two possible regions.

Case 1: $0 < t \leq 1$.

Then

$$\max(0, t-1) = 0, \quad \min(1, t) = t,$$

so

$$f_T(t) = \int_0^t 1 \, dx = t.$$

Case 2: $1 < t < 2$.

Then

$$\max(0, t-1) = t-1, \quad \min(1, t) = 1,$$

so

$$f_T(t) = \int_{t-1}^1 1 \, dx = 2 - t.$$

Therefore,

$$f_T(t) = \begin{cases} t, & 0 < t \leq 1, \\ 2 - t, & 1 < t < 2, \\ 0, & \text{otherwise.} \end{cases}$$

This is a **triangular distribution**: values near 1 can be obtained from more pairs (X, Y) than values near 0 or 2.

Temporary page!

$\text{\LaTeX}$ was unable to guess the total number of pages correctly. As there was some unprocessed data have been added to the final page this extra page has been added to receive it.
If you rerun the document (without altering it) this surplus page will go away, because $\text{\LaTeX}$ now knows the number of pages to expect for this document.