# Transformations

## Examples of transformations

- Linear transformations of random variables and vectors $Y = aX + b$
- Non-linear invertible transformations of random variables $Y = g(X)$
- Sums $Y = X_1 + X_2$

## Transformations previously

We have a technique for computing the expectation of transformed random variable. What is its name?

LOTUS:

$$\mathbb{E}[g(X)] = \sum_x g(x)\mathbb{P}(X = x)$$

It works with continuous r.v.s:

$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

And it works in multiple dimensions:

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y)\mathbb{P}(X = x, Y = y)$$

## Transformations now

Now we want not just the expected value, but the whole distribution (CDF, PMF, PDF). The approach depends on whether the distribution is discrete or continuous.

## Discrete case

The formula:

$$\mathbb{P}(g(X) = y) = \sum_{x \text{ such that } g(x) = y} \mathbb{P}(X = x)$$

If $g(\cdot)$ is one-to-one, it simplifies to:

$$\mathbb{P}(g(X) = y) = \mathbb{P}(X = g^{-1}(y))$$

# Example 1

Let $X \sim Bin(n, p)$. Find the PDF of $Y = \exp(X)$.

# Solution 1

So $g(x) = \exp(x)$, it's one-to-one and the inverse is $g^{-1}(x) = \log x$.

$$\mathbb{P}(Y = y) = \mathbb{P}(X = g^{-1}(y)) = \mathbb{P}(X = \log y)$$

# Continuous case

In the continuous case, when additionally $g(\cdot)$ is one-to-one, continuous and strictly increasing, we have the following relation for CDF:

$$F_{g(X)}(y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y))$$

To get the PDF, we need to differentiate this relation at every point with the chain rule:

$$f_{g(X)}(g(x))d\left(g(x)\right) = f_X(x)dx$$

To account for the case when $g(\cdot)$ is strictly decreasing, we add the modulus

$$f_{g(X)}(g(x)) = f_X(x)\left|\frac{dg(x)}{dx}\right|^{-1}$$

This is called 1D change of variables formula. Similarly to sicrete case, we can extend it to non one-to-one $g(\cdot)$ using the sum over $x$ such that $g(x) = y$.

# Example 2

Let $X \sim Exp(1)$. Find the PDF of $Y = \exp(-X)$.

# Solution 2

So $g(x) = \exp(-x)$, it is one-to-one, and $g^{-1}(y) = -\log y$. Let's find

$$\frac{dg(x)}{dx} = -\exp(-x)$$

So,

$$f_Y(y) = f_X(x)\left|\frac{dg(x)}{dx}\right|^{-1} = f_X(-\log y)\exp(x) =$$
$$= f_X(-\log y)\exp(-\log y) = \frac{1}{y} f_X(-\log y)$$

# Example 3

Let $X \sim \mathcal{N}(0,1)$. Find the PDF of $Y = X^2$. This distribution is called chi-square distribution.

# Solution 3

So, $g(x) = x^2$. It is not one-to-one, so we need the sum:

$$f_Y(y) = \sum_{x=\{-\sqrt{y},\sqrt{y}\}} f_X(x)\left|\frac{dg(x)}{dx}\right|^{-1} = \sum_{x=\{-\sqrt{y},\sqrt{y}\}} f_X(x)2x =$$
$$= f_X(\sqrt{y})\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\frac{1}{2\sqrt{y}} =$$
$$= 2f_X(\sqrt{y})\frac{1}{2\sqrt{y}}$$

# Example 4

Let $X$ be a random variable with PDF $f_X$. Find PDF of $Y = aX + b$.

# Solution 4

So, $g(x) = ax + b$, it is one-to-one, and the inverse is $g^{-1}(y) = \frac{1}{a}(y-b)$. We need to calculate

$$\frac{dg(x)}{dx} = \frac{d(ax+b)}{dx} = a$$

$$f_Y(y) = f_X(x)\left|\frac{dg(x)}{dx}\right|^{-1} = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}$$

# Multivariate transformations

Consider $n$-dimensional random vector $X \in \mathbb{R}^n$ with continuous distribution with PDF $f_X$. Let $g : A_0 \to B_0$ be invertive one-to-one function from open subset $A_0$ containing support of $X$ to an open subset $B_0$ containing the range of $g(\cdot)$. Denote $Y = g(X)$ and $y = g(x)$. Suppose that all partial derivatives $\frac{\partial y_i}{\partial x_j}$ exist and are continuous. Then, we can form the Jacobian matrix:

$$\frac{\partial y}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}$$

Assume that this matrix is non-degenerate. Then, the PDF of $Y$ is:

$$f_Y(y) = f_X(g^{-1}(y)) \left| \det \frac{\partial y}{\partial x} \right|^{-1}$$

If the function is not one-to-one, we add sum.

# Example 5

Consider $n$-dimensional random vector $X \in \mathbb{R}^n$ with continuous distribution, a non-degenerate matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$, then define random vector $Y = AX + b$. What is the expected value, the covariance matrix and the PDF of $Y$?

# Solution 5

Its expected value will be $\mathbb{E}[Y] = A\mathbb{E}[X] + b$ and its covariance matrix will be:

$$\Sigma_Y = \mathbb{E}\left[(AX - A\mathbb{E}[X])(AX - A\mathbb{E}[X])^\top\right] = A\mathbb{E}\left[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top\right] A^\top =$$

So,

- $g(x) = Ax + b$
- $\frac{\partial y}{\partial x} = A$
- $g^{-1}(y) = A^{-1}(y - b)$

Therefore,

$$f_Y(y) = f_X(g^{-1}(y))\left|\det \frac{\partial y}{\partial x}\right|^{-1} = f_X(A^{-1}(y-b))\frac{1}{|\det A|}$$

# Example 6

Consider $n$ independent standard normal r.v.s $X_1, \ldots, X_n \sim \mathcal{N}(0,1)$ and vector $X = (X_1, \ldots, X_n)^\top$. We know that

- $\mathbb{E}[X] = 0$
- $\Sigma_X = I$
- $f_X(x) = (2\pi)^{-n/2}\exp(-\frac{1}{2}x^\top x)$

Consider a non-degenerate matrix $A \in \mathbb{R}^{n\times n}$ and $Y = AX + m$. Find its expectation, covariance matrix and distribution.

# Solution 6

From the previous example, $\mathbb{E}[Y] = m$ and $\Sigma_Y = AA^\top$

$$\begin{aligned} f_X(x) = f_X(A^{-1}(y-m))\frac{1}{|\det A|} = \\ = \frac{1}{(2\pi)^{n/2}|\det A|}\exp\left(-\frac{(y-m)^\top A^{-\top}A^{-1}(y-m)}{2}\right) = \\ = \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma_Y}}\exp\left(-\frac{(y-m)^\top \Sigma_Y^{-1}(y-m)}{2}\right) \end{aligned}$$

What we just did, is we obtained a new random normal vector with controllable parameters from a standard normal random vector. This is a very useful property of a random normal vectors, but also demonstrates the power of linear transforms.

# Example 7

Let random vector $(X, Y)$ have PDF $f_{X,Y}(x, y)$. Find the density of $Z = X + Y$.

# Solution 7

In order to find this density, consider transform

$$\begin{pmatrix} Z \\ Y \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix}$$

So, matrix $A$ is...?

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Then,

$$f_{Z,Y}(z,y) = f_{X,Y}(A^{-1}(x\ y)^\top, y)|\det A|^{-1} = f_{X,Y}(z-y,y)$$

Finally,

$$f_Z(z) = \int f_{Z,Y}(z,y)dy = \int f_{X,Y}(z-y,y)dy$$

If $X \perp Y$,

$$f_Z(z) = \int f_X(z-y)f_Y(y)dy$$

This is the convolution rule that we studied on Seminar 4.

# Convolution rule

Consider independent r.v.s $X$ and $Y$. Then, their sum $Z = X + Y$ is distributed:

- If they are discrete,

$$\mathbb{P}(Z = n) = \sum_{k=-\infty}^{\infty} \mathbb{P}(X = k)\mathbb{P}(Y = n-k)$$

- If they are continuous,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$$

# Example 8

Let $X, Y \sim Exp(1)$ be independent. Find the PDF of $Z = \frac{X}{X+Y}$.

# Solution 8

This one will be a non-linear transform. Consider transform $(X, Y) \to (Z, U)$, where $U = X + Y$. The inverse transform is $X = UZ, Y = U - UZ$. Jacobian of the inverse transform is

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

The joint density is

$$f_{Z,U}(z, u) = f_{X,Y}(g^{-1}(z, u)) \left| \det \frac{\partial(x, y)}{\partial(z, u)} \right| = f_{X,Y}(uz, u - uz)u = ue^{u}, u > 0, 0 <$$

The joint density does not depend on $z$, it means that marginal density of $Z$ is Uniform with support $[0, 1]$.