Chapter 3
Functions and Transformations of Random Variables

Dr. Mohamed Rhaima

Department of Statistics and Operations Research

September, 30, 2019

1. Functions of one random variable

There are three main methods to find the distribution of a function of one or more random variables. These are to use the CDF, to transform the pdf directly or to use moment generating functions. We shall study these in turn and along the way find some results which are useful for statistics.

1.1 Method of distribution functions

I shall give an example before discussing the general method.

Example 1: Suppose the random variable Y has a pdf

$$f_Y(y) = \begin{cases} 3y^2, & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Suppose we want to find the pdf of $U = 2Y + 3$. The range of U is $3 < U < 5$. Now

$$F_U(u) = \mathbb{P}(U \leq u) = \mathbb{P}(2Y + 3 \leq u) = \mathbb{P}\left(Y \leq \frac{u-3}{2}\right).$$

Therefore

$$\mathbb{P}\left(Y \leq \frac{u-3}{2}\right) = \int_0^{\frac{u-3}{2}} f_Y(y)dy = \int_0^{\frac{u-3}{2}} 3y^2 dy = \left(\frac{u-3}{2}\right)^3$$

Then the CDF of U is given by

$$F_U(u) = \begin{cases} 0, & \text{if} \quad u < 3 \\ \left(\frac{u-3}{2}\right)^3, & \text{if} \quad 3 \leq u \leq 5 \\ 1, & \text{if} \quad u > 5 \end{cases}$$

and

$$f_U(u) = \frac{dF_U}{du} \begin{cases} \frac{3}{8}(u-3)^2, & \text{if} \quad 3 \leq u \leq 5 \\ 0, & \text{otherwise.} \end{cases}$$

The general method works as follows.:

1. Identify the domain of Y and U.
2. Write $F_U(u) = \mathbb{P}(U \leq u)$, the cdf of U, in terms of $F_Y(y)$, the cumulative distribution function of Y.
3. Differentiate $F_U(u)$ to obtain the pdf of U, $f_U(u)$.

The cdf method is useful for dealing with the squares of random variables. Suppose $U = X^2$, then

$$
\begin{aligned}
F_U(u) &= \mathbb{P}(U \leq u) \\
&= \mathbb{P}(X^2 \leq u) \\
&= \mathbb{P}(-\sqrt{u} \leq X \leq \sqrt{u}) \\
&= \int_{-\sqrt{u}}^{\sqrt{u}} f_X(x)dx \\
&= F_X(\sqrt{u}) - F_X(-\sqrt{u})
\end{aligned}
$$

So if we differentiate both sides with respect to u we find

$$\begin{aligned} f_U(u) &= f_X(\sqrt{u})\left(\frac{1}{2\sqrt{u}}\right) + f_X(-\sqrt{u})\left(\frac{1}{2\sqrt{u}}\right) \\ &= \frac{1}{2\sqrt{u}}\left(f_X(\sqrt{u}) + f_X(-\sqrt{u})\right) \end{aligned}$$

So, for example if

$$f_X(x) = \begin{cases} \frac{x+1}{2}, & \text{if} \quad -1 \leq x \leq 1 \\ 0, & \text{elsewhere.} \end{cases}$$

If $U = X^2$ then

$$\begin{aligned} f_U(u) &= \frac{1}{2\sqrt{u}}\left(f_X(\sqrt{u}) + f_X(-\sqrt{u})\right) \\ &= \frac{1}{2\sqrt{u}}\left(\frac{\sqrt{u}+1}{2} + \frac{-\sqrt{u}+1}{2}\right) \\ &= \frac{1}{2\sqrt{u}} \qquad 0 \leq u \leq 1. \end{aligned}$$

Example 2:

As a more important example suppose $Z \sim N(0,1)$.
Find the distribution of $U = Z^2$.
Solution: $Z \sim N(0,1)$ so that

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right) \qquad -\infty < x < +\infty$$

Then if $U = Z^2$

$$\begin{aligned} f_U(u) &= \frac{1}{2\sqrt{u}}\left(f_X(\sqrt{u}) + f_X(-\sqrt{u})\right) \\ &= \frac{1}{2\sqrt{u}}\left(\frac{1}{\sqrt{2\pi}} exp\left(-\frac{u}{2}\right) + \frac{1}{\sqrt{2\pi}} exp\left(-\frac{u}{2}\right)\right) \\ &= \frac{1}{\sqrt{2\pi u}} exp\left(-\frac{u}{2}\right), \qquad u > 0. \end{aligned}$$

Example 3:

Suppose that $Y \sim U(0,1)$. Find the distribution of $U = g(Y) = -\ln(Y)$.

Solution: The CDF of $Y \sim U(0,1)$ is given by

$$F_Y(y) = \begin{cases} 0, & \text{if} \quad y \le 0 \\ y, & \text{if} \quad 0 < y \le 1 \\ 1, & \text{if} \quad y \ge 1 \end{cases}$$

The domain (domain is the region where the pdf is non-zero) for $Y \sim U(0,1)$ is $R_Y = \{y : 0 < y < 1\}$, thus, because $u = -\ln y$, it follows that the domain for U is $R_U = \{u : u > 0\}$. The cdf of U is:

$$\begin{aligned} F_U(u) &= \mathbb{P}(U \le u) = \mathbb{P}(-\ln Y \le u) \\ &= \mathbb{P}(\ln Y > u) \\ &= \mathbb{P}(Y > e^{-u}) = 1 - \mathbb{P}(Y \le e^{-u}) \\ &= 1 - F_Y(e^{-u}) \end{aligned}$$

Because $F_Y(y) = y$ for $0 < y < 1$ ie. for $u > 0$, we have

$$F_U(u) = 1 - F_Y(e^{-u}) = 1 - e^{-u}$$

Taking derivatives, we get, for $u > 0$,

$$f_U(u) = \frac{d}{du} F_U(u) = \frac{d}{du}\left(1 - e^{-u}\right) = e^{-u}$$

Summarizing,

$$f_U(u) = \begin{cases} e^{-u}, & \text{if} \quad u > 0 \\ 0, & \text{elsewhere} \end{cases}$$

The Gamma distribution

Definition

We say the random variable X has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$, which we shall write as $Y \sim Gamma(\alpha, \beta)$ if

$$f_X(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta x), \qquad 0 \leq x < +\infty$$

where

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} \exp(-t)\, dt$$

is the Gamma function.

Now we can see that

$$\Gamma(1) = \int_0^{+\infty} \exp(-t)\, dt = 1$$

Also if we integrate $\Gamma(\alpha)$ by parts we see that

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1) \tag{1}$$

The Chi-Square distribution

Notice that as relation (1), we have $\Gamma(2) = 1 \times \Gamma(1) = 1$, $\Gamma(3) = 2 \times \Gamma(2) = 2$, $\Gamma(4) = 3 \times \Gamma(3) = 6$ and so on so that if n is an integer

$$\Gamma(n) = (n-1)!$$

Exercise: Show that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Definition

We say that a random variable X with a $Gamma(\frac{\nu}{2}, \frac{1}{2})$ distribution where ν is an integer has a Chi-Square distribution with ν degrees of freedom and we write it as $X \sim \chi^2_\nu$. ν is the Greek letter nu.

We showed in Â¨**Example 2** that the square of a standard normal distribution had pdf

$$f_U(u) = \frac{1}{\sqrt{2\pi u}} exp\left(-\frac{u}{2}\right), \qquad u > 0.$$

We can rewrite this, using the results above as

$$f_U(u) = \frac{(1/2)^{1/2} u^{-1/2}}{\Gamma(1/2)} exp\left(-\frac{u}{2}\right), \qquad u > 0.$$

and so U has a $Gamma(1/2, 1/2)$ or χ_1^2 distribution.
So we have proved the following theorem.

Theorem

If the random variable $Z \sim N(0,1)$ then $Z^2 \sim \chi_1^2$.

1.2 Method of direct transformation

Theorem

Let Y be a continuous random variable with probability density function f_Y and support I, where $I = [a, b]$. Let $g : I \to \mathbb{R}$ be a continuous monotonic function with inverse function $g^{-1} : J \to I$ where $J = g(I)$. Then the probability density function f_U of $U = g(Y)$ satisfies

$$f_U(u) = \begin{cases} f_Y\left(g^{-1}(u)\right)\left|\frac{d}{du}g^{-1}(u)\right|, & \text{if} \quad u \in J \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Steps of the pdf technique::

1. Verify that the transformation $u = g(y)$ is continuous and one-to-one over R_Y.
2. Find the domains of Y and U.
3. Find the inverse transformation $y = g^{-1}(u)$ and its derivative (with respect to u).
4. Use the formula (2) above for $f_U(u)$.

Example 4: Suppose Y has the density

$$f_Y(Y) = \begin{cases} \frac{\theta}{y^{\theta+1}}, & \text{if} \quad y > 1 \\ 0, & \text{otherwise,} \end{cases}$$

where θ is a positive parameter. This is an example of a **Pareto distribution**. Find the density of $U = \ln Y$.

Solution:

R_Y is the domain (domain is the region where the pdf is non-zero) for Y, then $R_Y = \{y : y > 1\}$. As the domain of U (the domain on which the density is non-zero) is $R_U = \{u : u > 0\}$. The inverse transformation is $y = \exp(u)$ and $\frac{d}{du}\exp(u) = \exp(u)$. Therefore

$$\begin{aligned} f_U(u) &= \frac{\theta}{(\exp(u))^{\theta+1}} \times \exp(u) \\ &= \theta \exp(-u\theta) \qquad u > 0, \end{aligned}$$

and so U has an exponential distribution.

2. More-to-one transformation

Here we discuss transformations involving two random variable Y_1, Y_2. The **bivariate transformation** is

$$\begin{aligned} U_1 &= g_1(Y_1, Y_2) \\ U_2 &= g_2(Y_1, Y_2) \end{aligned}$$

Assuming that Y_1 and Y_2 are jointly continuous random variables, we will discuss the one-to-one transformation first. Starting with the joint distribution of $Y = (Y_1, Y_2)$, our goal is to derive the joint distribution of $U = (U_1, U_2)$.

Suppose that $Y = (Y_1, Y_2)$ is a continuous random vector with joint pdf $f_{Y_1Y_2}(y_1, y_2)$. Let $g : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous one-to-one vector-valued mapping from $R_{Y_1Y_2}$ to $R_{U_1U_2}$ where $U_1 = g_1(Y_1, Y_2)$ and $U_2 = g_2(Y_1, Y_2)$ and where $R_{Y_1Y_2}$ and $R_{U_1U_2}$ denote the two-dimensional domain of $Y = (Y_1, Y_2)$ and $U = (U_1, U_2)$, respectively. If $g_1^{-1}(u_1, u_2)$ and $g_2^{-1}(u_1, u_2)$ have continuous partial derivatives with respect to both u_1 and u_2, and the Jacobian J where, with "det" denoting "determinant",

$$J = det \begin{vmatrix} \frac{\partial}{\partial u_1} g_1^{-1}(u_1, u_2) & \frac{\partial}{\partial u_2} g_1^{-1}(u_1, u_2) \\ \frac{\partial}{\partial u_1} g_2^{-1}(u_1, u_2) & \frac{\partial}{\partial u_2} g_2^{-1}(u_1, u_2) \end{vmatrix}$$

then

$$f_{U_1 U_2}(u_1, u_2) = \begin{cases} f_{Y_1 Y_2}\left(g_1^{-1}(u_1, u_2), g_2^{-1}(u_1, u_2)\right) |J|, & \text{if } (u_1, u_2) \in R_{U_1 U_2} \\ 0, & \text{elsewhere.} \end{cases} \tag{3}$$

RECALL: The determinant of a 2×2 matrix, e.g.,

$$det \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Steps of the pdf technique:

1. Find $f_{Y_1Y_2}(y_1, y_2)$, the joint distribution of Y_1 and Y_2. This may be given in the problem. If Y_1 and Y_2 are independent, then $f_{Y_1Y_2}(y_1, y_2) = f_{Y_1}(y_1)f_{Y_2}(Y)_2)$.
2. Find $R_{U_1U_2}$, the domain of $U = (U_1, U_2)$.
3. Find the inverse transformations $y_1 = g_1^{-1}u_1, u_2$ and $y_2 = g_2^{-1}u_1, u_2$.
4. Find the Jacobian J, of the inverse transformation.
5. Use the formula (3) above to find $f_{U_1U_2}(u_1, u_2)$ the joint distribution of U_1 and U_2.

NOTE: If desired, marginal distributions $f_{U_1}(u_1)$ and $f_{U_2}(u_2)$ can be found by integrating the joint distribution $f_{U_1U_2}(u_1, u_2)$.

Example 5:

Suppose that Y_1 and Y_2 have joint pdf

$$f_{Y_1 Y_2}(y_1, y_2) = \exp\left(-\left(y_1 + y_2\right)\right) \qquad y_1 \geq 0,\ y_2 \geq 0$$

Consider the transformation $u_1 = y_1$ and $u_2 = y_1 + y_2$.

1. Find $f_{U_1 U_2}$ the pdf of U_1 and U_1.
2. Find the marginal distribution of U_1.
3. Find the marginal distribution of U_2.

Solution:

1. The transformation $u_1 = y_1$ and $u_2 = y_1 + y_2$ has the inverse $y_1 = u_1$ and $y_2 = u_2 - u_1$. Therefore, $R_{U_1U_2} = \{(u_1, u_2) : 0 \le u_1 \le u_2 \le +\infty\}$. The Jacobian is

$$J = det \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

Since, the joint pdf of U_1 and U_2 is given by

$$f_{U_1U_2}(u_1, u_2) = 1 \times \exp(-u_2) \qquad 0 \le u_1 \le u_2 \le +\infty$$

2. If we want the pdf of $U_2 = Y_1 + Y_2$ we must find the marginal pdf of U_2 by integrating out U_1.

$$f_{U_2}(u_2) = \int_0^{u_2} \exp(-u_2)\, du_1 = u_2 \exp(-u_2) \qquad 0 \le u_2 \le +\infty$$

3. The moment generating functions

Definition

1. The moment generating function of a random variable X, written as $M_X(t)$ is defined by
$$M_X(t) = \mathbb{E}\left(\exp\left(tX\right)\right) \tag{4}$$
and is defined for t in a region about 0, $-h < t < h$ for some h.
2. The moment generating function of a random variable $Y = U(X_1, X_2, \dots, X_n)$, written as $M_Y(t)$ is defined by
$$M_Y(t) = \mathbb{E}\left(\exp\left(tY\right)\right) = \mathbb{E}\left(\exp\left(tU(X_1, X_2, \dots, X_n)\right)\right)$$

Example 6:

Let X_1 and X_2 be independent random variables with uniform distributions on $\{1, 2, 3, 4\}$. Let $Y = X_1 + X_2$, find the moment-generating function of Y.

First note that $M_X(0) = 1$. Differentiating $M_X(t)$ in (4) with respect to t assuming X is continuous we have

$$
\begin{aligned}
M'_X(t) &= \frac{d}{dt}\int \exp(tx)f(x)dx \\
&= \int x\exp(tx)f(x)dx \\
M'_X(0) &= \int xf(x)dx \\
&= \mathbb{E}(X)
\end{aligned}
$$

Similarly

$$
\begin{aligned}
M"_X(t) &= \frac{d^2}{dt^2}\int \exp(tx)f(x)dx \\
&= \int x^2\exp(tx)f(x)dx \\
M"_X(0) &= \int x^2f(x)dx \\
&= \mathbb{E}(X^2)
\end{aligned}
$$

Theorem

$$V(X) = M''_X(0) - \left(M'_X(0)\right)^2 .$$

The uncentred moments of X are generated from this function by

$$\mathbb{E}\left(X^n\right) = M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t) \mid_{t=0}$$

Example 7: Suppose X is a discrete binomial random variable with probability mass function

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2 \ldots, n$$

1. Find the moment generating function of X.
2. Compute the variance of X.

Example 8:

Suppose X has a Gamma distribution, $Gamma(\alpha, \beta)$.

1. Prove that the mgf of X is given by

$$M_X(t) = \left(\frac{\beta}{\beta - t} \right)^{\alpha}$$

2. Compute the mean of X.
3. Compute the variance of X.

The following theorem, tells us why we can use the mgf to find the distributions of transformed variables.

Theorem

f X_1 and X_2 are random variables and $M_{X_1}(t) = M_{X_2}(t)$ then X_1 and X_2 have the same distribution.

Example 9: Suppose $Z \sim N(0,1)$ and $Y = Z^2$.

1. Find the mgf of Z.
2. Find the mgf of Y.
3. Find the mgf of χ_1^2.
4. Deduce the pdf of Y.

Theorem

Suppose $X_1, X_2, \ldots, X_n$ are independent rvs with mgf $M_{X_i}(t)$, $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} X_i$ then

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

Example 10: Let $X_1, X_2, \ldots, X_n$ denote the outcomes of n Bernoulli trials, each with probability of success p. Let $Y = \sum_{i=1}^{n} X_i$.

1. find the mgf of X_i, $i = 1, \ldots, n$.
2. Prove that the mgf of Y is given by

$$M_Y(t) = \prod_{i=1}^{n} (q + pe^t) = (q + pe^t)^n .$$

3. Deduce the pdf of Y.

Theorem

Suppose $X_1, X_2, \ldots, X_n$ are independent, normally distributed with mean $\mathbb{E}(X_i) = \mu_i$ and variance $V(X_i) = \sigma_i^2$. Let $Z_i = \frac{X_i - \mu_i}{\sigma_i}$ so that $Z_1, Z_2, \ldots, Z_n$ are independent and each has a $N(0,1)$ distribution. Then $\sum_i Z_i^2$ has a χ_n^2 distribution.

Proof: We have seen before that each Z_i^2 has a χ_1^2 distribution. So

$$M_{Z_i^2}(t) = (1-2t)^{-1/2}$$

Let $Y = \sum Z_i^2$. Then

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^n M_{Z_i^2}(t) \\ &= \frac{1}{(1-2t)^{n/2}} = \left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)^{\frac{n}{2}} \end{aligned}$$

but this is the mgf of a $Gamma(n/2, 1/2)$ random variable, that is a χ_n^2 random variable.

Characteristic function

The characteristic function of a random variable $g(X)$, defined as

$$\Phi_{g(X)}(t) = \mathbb{E}(\exp(itg(X))) = \int_{-\infty}^{+\infty} \exp(itg(x)) f(x) dx$$

where $f(x)$ is the density for X.

Features of characteristic function:

- The CF always exists. This follows from the equality $e^{itx} = \cos(tx) + i\sin(tx)$, and both the real and complex parts of the integrand are bounded functions.
- Consider a symmetric density function, with $f(-x) = f(x)$ (symmetric around zero). Then resulting $\phi(t)$ is real-valued, and symmetric around zero.
- The CF completely determines the distribution of X (every cdf has a unique characteristic function).

- Let X have characteristic function $\phi_X(t)$. Then $Y = aX + b$ has characteristic function $\phi_Y(t) = e^{ibt}\phi_X(at)$
- X and Y, independent, with characteristic functions $\Phi_X(t)$ and $\phi_Y(t)$. Then $\phi_{X+Y}(t) = \phi_X(t)\phi_X(t)$.
- $\phi_X(0) = 1$.
- For a given characteristic function $\phi_X(t)$ such that $\int_{-\infty}^{+\infty} |\phi_X(t)|dt < \infty$ the corresponding density $f_X(x)$ is given by the inverse Fourier transform, which is

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Phi_X(t)\exp(-itx)dt$$

- Characteristic function also summarizes the moments of a random variable. Specially, note that the h-th derivative of $\phi_X(t)$ is

$$\phi_X^{(h)}(t) = \int_{-\infty}^{+\infty} i^h g(x)^h \exp(itg(x)) f_X(x) dx \tag{5}$$

Hence, assuming the h-th moment, denoted $\mu_{g(X)}^h = \mathbb{E}\,(g(X))^h$ exists, it is equal to

$$\mu_{g(X)}^h = \frac{\phi_X^{(h)}(0)}{i^h}$$

Hence, assuming that the required moments exist, we can use Taylor's theorem to expand the characteristic function around $t = 0$ to get:

$$\phi_X(t) = 1 + \frac{it}{1}\mu_{g(X)}^1 + \frac{(it)^2}{2!}\mu_{g(X)}^2 + \cdots + \frac{(it)^k}{k!}\mu_{g(X)}^k + o(t^k)$$

- Cauchy distribution, cont'd: The characteristic function for the Cauchy distribution is

$$\phi_X(t) = \exp(-|t|)$$

This is not differentiable at $t = 0$, which by equation (5) is saying that its mean does not exist. Hence, the expansion of the characteristic function in this case is invalid.