Chapter 2.3 Transformation Examples

1 General Comments on Transformations and Distributions

In these examples X is a continuous random variable. In a previous handout we considered X as a discrete r.v.

This section studies r.v.s $Y = g(X)$ where

$$g : D \mapsto E$$

where D, E are the domain and range of the function g. In particular the goal is to find the distribution of Y, in terms of the given information which is the distribution of X.

Properties of the solution depend on properties of the function g, in particular its domain and its range. The domain D and the range E are subsets of $\mathcal{R}$, but E may be a subset of the integers, and in some other special cases a categorical set, that is a finite set of categories. Depending on the properties of g, the r.v. Y may be continuous or discrete, or possibly neither (partly continuous and partly discrete). Some examples of g are

1. $g : (0, \infty) \mapsto \mathcal{R}$ given by $g(x) = \log(x)$ Aside : In this course we use log to mean the natural log unless otherwise specified.

   This function g is a 1 to 1‘ and onto function and has an inverse.

2. $g : \mathcal{R} \mapsto \mathcal{R}$ given by $g(x) = a + bx$ for some constants a, b. When $b \neq 0$ this is a 1 to 1 and onto function. It has an inverse.

3. $g : \mathcal{R} \mapsto \mathcal{R}$ given by $g(x) = \sin(x)$ (an into and not onto function)

4. $g : \mathcal{R} \mapsto [0, 1]$ given by $g(x) = \sin(x)$ (an onto function, but not 1 to 1)

5. $g : (0, \infty) \mapsto N_0$ where N_0 is the set of non-negative integers and $g(x) = [x] = \text{floor}(x) = \text{glb}(x)$ where $[x]$ is the greatest lower integer bound of x. In this case $Y = g(X)$ will be a discrete r.v.

6. $g : (0,1) \mapsto \{1,2,3\}$ where

$$g(x) = \begin{cases} 1 & \text{if } 0 < x \le \frac{1}{2} \\ 2 & \text{if } \le \frac{1}{2} < x \le \frac{5}{6} \\ 3 & \text{if } \frac{5}{6} < x < 1 \end{cases}$$

In this case Y is discrete. The function g is a many to one function, not a 1 to 1 function.

2 Linear Transformation

Suppose that $X \sim f$. Consider $Y = aX + b$, where $a > 0$. Then

$$\begin{aligned} F_Y(y) &= P(aX + b \le y) \\ &= P\left(X \le \frac{y-b}{a}\right) \text{ since } a > 0 \\ &= F_X\left(\frac{y-b}{a}\right) . \end{aligned}$$

We see that in this particular transformation the cdf of Y is easy to obtain in terms of the cdf of X. Our goal is to find the pdf of Y, that is the pdf f_Y, in terms of the given pdf f_X. We need to find this for every possible value of the argument, say y.

Thus by differentiation, using the chain rule, we obtain

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X\left(\frac{y-b}{a}\right)\frac{1}{a} .$$

An analogous result holds for $a < 0$. The student should work through the above steps and find that

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

Notice, if $a \neq 0$, then the function $g : R \mapsto R$ is a strictly monotone decreasing function, and hence is a 1-1 onto function and hence it has an inverse. If $y = g(x)$ then there is an inverse function $g^{-1} : R \mapsto R$ so that $x = g^{-1}(y) = \frac{y-b}{a}$.

Application to the normal family of distributions.

Suppose that $Z \sim \phi$ where

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

is the standard normal pdf.

If $X = \sigma Z + \mu$, where $\sigma > 0$ then X has pdf

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

The notation $f(x;\mu,\sigma^2)$ explicitly denotes the dependence of this distribution on the parameters μ and σ^2. This family is called the family of normal distributions. We also write

$$X \sim N(\mu, \sigma^2)$$

as a short hand for this. With this notation $Z \sim N(0,1)$.

If $X \sim N(\mu,\sigma^2)$, and $\sigma = \sqrt{\sigma^2} > 0$ then

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1) \ .$$

$$P\left(X \le x\right) \quad = \quad P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right)$$

Thus to solve for a q-th quantile

$$P(X \le x) = q$$

the solution is a number x_q such that

$$\frac{x_q - \mu}{\sigma} = z(q)$$

where $z(q)$ is the solution of

$$P(Z \le z(q)) = q$$

If $q = 0.975$ then $z(q) = 1.96$ and $x_q = \mu + 1.96 * \sigma$.

Application : Gamma family of distributions.

Consider the probability density, for a given *shape* parameter value $\alpha > 0$

$$g(x;\alpha) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} \mathrm{I}(x > 0) \ .$$

The normalizing constant is $\Gamma(\alpha)$, that is the number so that g integrates (with respect to its argument x) to 1. See the text for more information about the Gamma function.

The value of the normalizing constant depends on the parameter α, and this normalizing constant is in itself the value of the Gamma function evaluated at α.

Consider $Y = \frac{1}{\lambda} X$ where $X \sim g$, and where $\lambda > 0$. Since $\lambda > 0$ then $y = x/\lambda > 0$ if and only if $x > 0$. Then Y has pdf

$$f(y;\alpha,\lambda) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha y^{\alpha-1} e^{-\lambda y} \mathrm{I}(y > 0)$$

Since the argument y is just a dummy variable we may write

$$f(x;\alpha,\lambda) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} \mathrm{I}(x > 0) \ . \tag{1}$$

The pdf (1) is a valid pdf for any parameter value $(\alpha, \lambda) \in R^+ \times R^+ = \Theta$. We refer to the family of distributions with pdf (1) and parameter space Θ as the Gamma distribution family.

Notice that the Gamma function plays a role in this only as the normalizing constant.

The pdf (1) may also be written in the form

$$f(x; \alpha, \lambda) = \begin{cases} \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

3 A Special Non-invertible Transform

Suppose $X \sim f$ and $Y = X^2$. The function $g : R \mapsto R$, given by the rule $g(x) = x^2$ does not have an inverse since it is not 1 to 1.

It is not invertible even if we change it to $g : R \mapsto [0, \infty)$, in which case it is an onto function, but still not 1 - 1. If however the domain and range are changed so that $g : [0, \infty) \mapsto [0, \infty)$, then the function g is 1-1 and onto and does have an inverse.

Here we consider g as a mapping of R to R. Notice that the method used in the previous application does not apply since our function g does not have an inverse. However we can still try to work with the cdf of Y.

Consider the cdf of Y. For $y < 0$ of course $F_Y(y) = 0$ and hence for $y < 0$ we have $f_Y(y) = 0$. For $y \geq 0$ we have

$$\begin{aligned} F_Y(y) &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - P(X < -\sqrt{y}) \\ &= F_X(\sqrt{y}) - P(X \leq -\sqrt{y}) \text{ since } F_X \text{ is continuous} \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

The pdf of Y, for $y > 0$, is then obtained by differentiation, using the chain rule.

$$\begin{aligned} f_Y(y) &= f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})(-1) \frac{1}{2\sqrt{y}} \\ &= f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} + f_X(-\sqrt{y}) \frac{1}{2\sqrt{y}} \end{aligned}$$

This formula is not defined when $y = 0$. Since $f_Y(y) = 0$ for $y < 0$ we can then conveniently complete the pdf by setting $f_Y(0) = 0$. Thus

$$f_Y(y) = \left\{ f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} + f_X(-\sqrt{y}) \frac{1}{2\sqrt{y}} \right\} I(y > 0)$$

Application to $Z \sim N(0,1)$

Z has pdf ϕ which is symmetric about 0. Thus $X = Z^2$ has pdf where for $x > 0$

$$\begin{aligned} f(x) &= \phi(\sqrt{x})\frac{1}{2\sqrt{x}} + \phi(-\sqrt{x})\frac{1}{2\sqrt{x}} \\ &= \phi(\sqrt{x})\frac{1}{\sqrt{x}} \\ &= \frac{1}{\sqrt{2\pi x}}e^{-\frac{x^2}{2}} \end{aligned}$$

More formally the pdf is

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}}e^{-\frac{x^2}{2}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

This density is also called the χ^2 distribution with 1 degree of freedom and we write

$$X = Z^2 \sim \chi^2_{(1)} \ .$$

4 General 1 to 1 Invertible Transforms

Suppose that $g : x \mapsto g(x)$ is 1 to 1, invertible and differentiable. Specifically suppose that $y = g(x)$ has inverse $x = h(y)$. We also write $h = g^{-1}$. Then

1. If h is monotone increasing

   $$\{x : g(x) \leq y\} = \{x : x \leq h(y)\}$$

   and

   $$F_Y(y) = F_X(h(y))$$

2. If h is monotone decreasing

   $$\{x : g(x) \leq y\} = \{x : x \geq h(y)\}$$

   and

   $$F_Y(y) = 1 - F_X(h(y))$$

In either case, by applying the chain rule we obtain

$$f_Y(y) = f_X(h(y))\left|\frac{dh(y)}{dy}\right|$$

where the absolute value is needed as h in increasing or decreasing. Note also that if $y \in \mathcal{R}$ is such that there is no $x \mapsto y = g(x)$ then for such y we have $f_Y(y) = 0$.

If we formally replace $h(y) = g^{-1}(y)$ we obtain the same formula as Rice, Section 2.3, Proposition B.

Derivation:

Suppose that g is monotone increasing. Then for a given y in the range of g we have

$$\{x : g(x) \leq y\} = \{x : x \leq g^{-1}(y)\}$$

Thus for y in the range of $g(X)$

$$\begin{aligned} F_Y(y) &= P(g(X) \leq y) \\ &= P(X \leq g^{-1}(y)) \end{aligned}$$

By differentiation using the chain rule we obtain, for such y

$$f_Y(y) = f_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy} \ .$$

Exponential Distribution

Suppose that $U \sim U(0,1)$. Let $\lambda > 0$. Consider $y = g(u) = -\frac{1}{\lambda} \log(u)$ as a mapping on $(0, \infty)$; that is g has domain $(0, \infty)$. We can solve for u in terms of y as

$$u = e^{-\lambda y} \equiv h(y) \ .$$

Notice that h is the inverse function of g. We also write this as $h(y) = g^{-1}(y)$ for the appropriate domain of y. Since h is differentiable we obtain for $y > 0$

In our application we are really only interested in treating g as a function which maps $(0,1)$ into the reals. In this restricted case $g : (0,1) \mapsto (0, \infty)$, that is the range is now the positive reals, rather than all reals. g is still invertible, with inverse

$$h : (0, \infty) \mapsto (0,1)$$

given by

$$h(y) = e^{-\lambda y} \ .$$

Thus the methodology above applies and the rv $Y = g(U)$ has pdf, for argument $y > 0$ given by

$$\begin{aligned} f_Y(y) &= f_U(h(y)) \left| \frac{dh(y)}{dy} \right| \\ &= 1 \left| \lambda e^{-\lambda y} \right| \ . \end{aligned}$$

For completeness we have

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

Two important special cases are given in Rice, Section 2.3 as Propositions C and D.

Proposition C: Suppose that F is continuous, $X \sim F$ and $Z = F(X)$. Then $Z \sim U(0,1)$.

Proposition D: Suppose that F is continuous, and that $U \sim U(0,1)$. Let $X = F^{-1}(U)$. Then $X \sim F$.

Remarks:

Proposition C is useful in many areas of statistical inference. It is the basis for nonparametric inference that is used in medical studies, in the social sciences, reliability and other applications.

Proposition D is an important result since it allows one to transform $U(0,1)$ r.v.s into a r.v. with any other distribution. This is useful in computer programs that can generate $U(0,1)$ r.v.'s, and is the basis for many simulation methods.

As an example application for Proposition D consider the exponential parameter λ cdf

$$F(y) = \begin{cases} 1 - e^{-\lambda y} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases} .$$

For a given $u \in (0,1)$ we can solve for y in the equation $u = F(y) = 1 - e^{-\lambda y}$ Thus $y = -\frac{1}{\lambda}\log(1-u)$. Thus for $0 < u < 1$ we have

$$y = F^{-1}(u) = -\frac{1}{\lambda}\log(1-u) \ .$$

Thus if $U \sim \text{Unif}(0,1)$ then

$$Y = F^{-1}(U) = -\frac{1}{\lambda}\log(1-U) \sim \text{exponential}, \lambda \ .$$

Notice that $W = 1 - U$ also has a Unif(0,1) distribution and hence

$$Y = -\frac{1}{\lambda}\log(W) \sim \text{exponential}, \lambda \ .$$

This is the algorithm given before Propositions C and D.

5 An Example Where $Y = g(X)$ is discrete

In this example let $X \sim \text{Unif}(0,1)$ and let $g : (0,1) \mapsto R$ be given by

$$g(x) = \begin{cases} 0 & \text{if } 0 < x \le \frac{1}{4} \\ 1 & \text{if } \frac{1}{4} < x \le \frac{3}{4} \\ 2 & \text{if } \frac{3}{4} < x < 1 \end{cases}$$

Notice that $Y = g(X)$ can only take 3 values with positive probability, and hence Y is discrete. In this case it is then useful to write the distribution of Y in its pmf form.

$$\begin{aligned} P(Y = 0) &= P(X \in \{x : g(x) = 0\}) \\ &= P(0 < X < \frac{1}{4}) \\ &= \int_0^{\frac{1}{4}} 1dx \\ &= \frac{1}{4} \end{aligned}$$

The student should finish this example and find that the pmf of Y is given by

$$P(Y = y) = \begin{cases} \frac{1}{4} & \text{if } y = 0 \\ \frac{1}{2} & \text{if } y = 1 \\ \frac{1}{4} & \text{if } y = 2 \\ 0 & \text{otherwise} \end{cases}$$

Notice that Y has a Binomial(2, $\frac{1}{2}$) distribution. This is so even though the game above is a transformation or function of a Unif(0,1) r.v. and is not a coin tossing game.