Chapter 2
Transformations
Distributions of Functions of a Random Variable

Outline

A Single Random Variable

For a random variable X having state space S, consider the transformed random variable $Y = g(X) = g \circ X$ where the function

$$g : S \to \mathbb{R}.$$

The inverse image of a set A,

$$g^{-1}(A) = \{x \in S; g(x) \in A\}.$$

In other words,

$$x \in g^{-1}(A) \text{ if and only if } g(x) \in A.$$

For example, if $g(x) = x^3$, then $g^{-1}([1, 8]) = [1, 2]$.

A Single Random Variable

- For the singleton set $A = \{y\}$, we sometimes write $g^{-1}(\{y\}) = g^{-1}(y)$.
- For $y = 0$ and $g(x) = \sin x$, then $g^{-1}(0) = \{k\pi; k \in \mathbb{Z}\}$.
- If g is a one-to-one function, then the inverse image of a singleton set is itself a singleton set. In this case, the inverse image naturally defines an inverse function.
- For $g(x) = x^3$, this inverse function is the cube root.
- For $g(x) = \sin x$, we must limit the domain to obtain the inverse sine function. The typical choice for the domain of sine is $[-\pi/2, \pi/2]$.

Discrete Random Variables

For X a discrete random variable with probability mass function f_X, then the probability mass function f_Y for $Y = g(X)$ is

$$f_Y(y) = \sum_{x \in g^{-1}(y)} f_X(x)$$

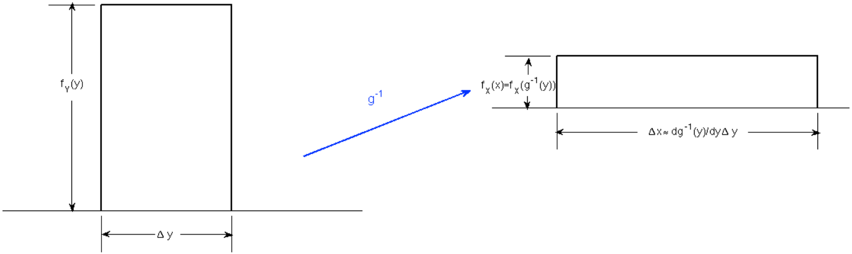
Example Let X be a uniform random variable on $\{1, 2, \ldots n\}$, i. e., $f_X(x) = 1/n$ for each x in the state space. Then $Y = X + a$ is a uniform random variable on $\{a+1, 2, \ldots a+n\}$.

Example Let X be a uniform random variable on $\{-n, -n+1, \ldots, n-1, n\}$. Then $Y = |X|$ has mass function

$$f_Y(y) = \begin{cases} \frac{1}{2n+1} & \text{if } x = 0, \\ \frac{2}{2n+1} & \text{if } x \neq 0. \end{cases}$$

Continuous Random Variables

The easiest case for transformations of continuous random variables is the case of g one-to-one. We consider the case of g increasing on the range of the random variable X by first looking at a geometric explanation of the transformation.



Finding the density of $Y = g(X)$ from the density of X. The areas of the two rectangles should be the same. Consequently, $f_Y(y)\Delta y \approx f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)\Delta y$.

Continuous Random Variables

More formally, let g be increasing and differentiable on the range of the random variable X. In this case, g^{-1} is also an increasing function.

To compute the cumulative distribution of $Y = g(X)$ in terms of the cumulative distribution of X, note that

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \leq g^{-1}(y)\} = F_X(g^{-1}(y)).$$

Now use the chain rule to compute the density of Y

$$f_Y(y) = F_Y'(y) = \frac{d}{dy}F_X(g^{-1}(y)) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y).$$

Continuous Random Variables

For g decreasing on the range of X,

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \geq g^{-1}(y)\} = 1 - F_X(g^{-1}(y)),$$

and the density

$$f_Y(y) = F'_Y(y) = -\frac{d}{dy} F_X(g^{-1}(y)) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y).$$

For g decreasing, we also have g^{-1} decreasing and consequently the density of Y is indeed positive,

We can combine these two cases to obtain

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

Continuous Random Variables

Example. Let U be a uniform random variable on $[0,1]$ and let $g(u) = 1 - u$. Then $g^{-1}(v) = 1 - v$, and $V = 1 - U$ has density

$$f_V(v) = f_U(1-v)|-1| = 1$$

on the interval $[0,1]$ and 0 otherwise. Thus, V is a uniform random variable on $[0,1]$.

Example. Let $g(x) = x^p,\ p \neq 0$. Then, $g^{-1}(y) = y^{1/p}$. If $p > 0$, then the range of g is $[0,1]$ and g is increasing.

$$\frac{d}{dy}g^{-1}(y) = \begin{cases} 0 & \text{if } y < 0, \\ \frac{1}{p}y^{1/p-1} & \text{if } 0 \leq y \leq 1, \\ 0 & \text{if } y > 1. \end{cases}$$

The density of $Y = g(U)$,

$$f_Y(y) = \frac{1}{p}y^{1/p-1} \text{ on } [0,1]$$

is *unbounded* near zero whenever $0 < p < 1$.

Continuous Random Variables

If $p < 0$, then g is *decreasing*. Its range is $[1, \infty)$, and

$$f_Y(y) = \begin{cases} 0 & \text{if } y < 1, \\ -\frac{1}{p} y^{1/p-1} & \text{if } y \geq 1, \end{cases}$$

In this case, Y is a Pareto distribution with $\alpha = 1$ and $\beta = -1/p$.

Exercise. Show that we can obtain a Pareto distribution with arbitrary α and β by taking X uniform on the interval $[0, 1]$

$$g(x) = \alpha x^{-1/\beta}.$$

Answer. $g^{-1}(y) = \left(\frac{y}{\alpha}\right)^{-\beta} = \alpha^\beta y^{-\beta}$

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

Continuous Random Variables

If the transform g is *not* one-to-one then special care is necessary to find the density of $Y = g(X)$. For example if we take $g(x) = x^2$,

$$F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\} = P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Thus,

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

Exercise. If the density f_X is symmetric about the origin, then

$$f_Y(y) = \frac{1}{\sqrt{y}} f_X(\sqrt{y}).$$

Continuous Random Variables

Example. A random variable Z is called a standard normal if its density is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}).$$

Exercise. Show that

$$\phi'(z) = -\frac{1}{\sqrt{2\pi}} z \exp(-\frac{z^2}{2}) = -z\phi(z), \quad \phi''(z) = \frac{1}{\sqrt{2\pi}}(z^2 - 1) \exp(-\frac{z^2}{2}) = (z^2 - 1)\phi(z).$$

Thus, ϕ has a global maximum at $z = 0$, it is concave down if $|z| < 1$ and concave up for $|z| > 1$. This show that the graph of ϕ has a bell shape.

Exercise. $Y = Z^2$ is called a χ^2 (chi-square) random variable with one degree of freedom. Show that its density is

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} \exp(-\frac{y}{2}).$$

Simulating Continuous Random Variables

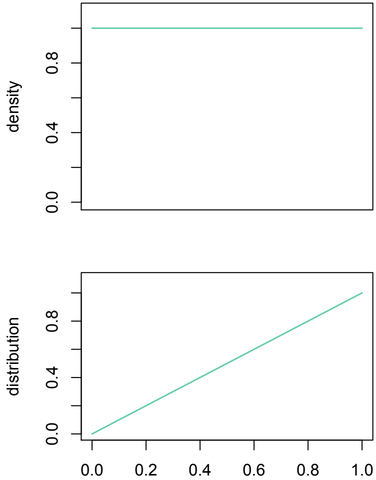
The `runif(n)` command is used to simulate n independent random variables U that are uniformly distributed on the interval $[0, 1]$.

The density function is

$$f_U(u) = \begin{cases} 0 & u < 0, \\ 1 & 0 \le u < 1, \\ 0 & 1 \le u. \end{cases}$$

The distribution function is

$$F_U(u) = \begin{cases} 0 & u < 0, \\ u & 0 \le u < 1, \\ 1 & 1 \le u. \end{cases}$$



Probability Transform

If X a continuous random variable with a density f_X that is positive everywhere in its domain, then

- the distribution function $F_X(x) = P\{X \leq x\}$ is strictly increasing.
- F_X has a inverse function F_X^{-1}, known as the quantile function.
  - For example, $F_X^{-1}(1/2)$ is the median, $F_X^{-1}(3/4)$ is the third quartile
- For values of u between 0 and 1, note that

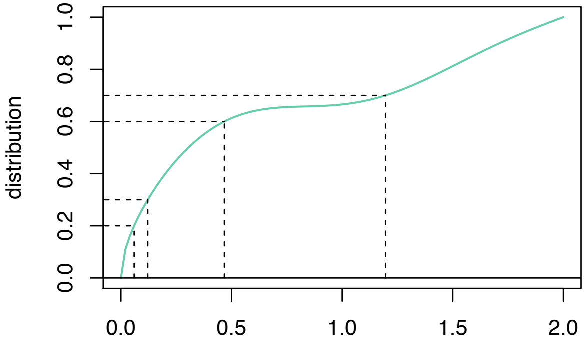
$$P\{F_X(X) \leq u\} = P\{X \leq F_X^{-1}(u)\} = F_X(F_X^{-1}(u)) = u.$$

  The distribution function for the random variable U.

- If we can simulate U, we can simulate a random variable with distribution F_X via the quantile function

$$X = F_X^{-1}(U).$$

Probability Transform



$F^{-1}(u)$	u
x	$F(x)$
0.060	0.2
0.121	0.3
0.468	0.6
1.196	0.7

Figure: Cumulative Distribution Function and the Quantile Function

Probability Transform

For the dart board, for x between 0 and 1, the distribution function

$$u = F_X(x) = x^2 \quad \text{and thus the quantile function} \quad x = F_X^{-1}(u) = \sqrt{u}.$$

We can simulate independent observations of the distance from the center of the dart board

$$X_1, X_2, \ldots, X_n$$

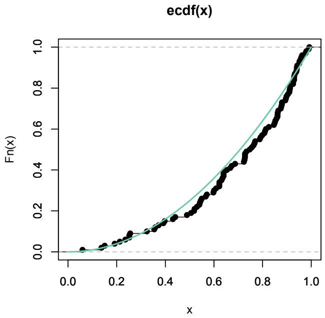
of the dart board by simulating independent uniform random variables $U_1, U_2, \ldots U_n$ and taking the quantile function

$$X_i = \sqrt{U_i}.$$

Probability Transform

```
> u<-runif(100)                    #simulate 100 uniform random variables
> x<-sqrt(u)                       #find the quantile functions
> plot(ecdf(x),xlim=c(0,1),ylim=c(0,1))
> par(new=TRUE)
> curve(x^2,0,1,ylim=c(0,1),col="aquamarine3",xlab="",ylab="",lwd=2)
```

Exercise. Perform the simulation of dart throws. Give the 0.25, 0,50, and 0.75 quantiles for both the distribution function F_X and for the simulated values.



Probability Transform

An exponential random variable T has distribution function

$$F_T(t) = \begin{cases} 0, & t < 0, \\ 1 - exp(-\beta t), & 0 \leq t. \end{cases}$$

For the inverse,

$$u = 1 - exp(-\beta t) \quad \exp(-\beta t) = 1 - u \quad -\beta t = \ln(1-u) \quad t = \frac{1}{\beta}\ln(1-u).$$

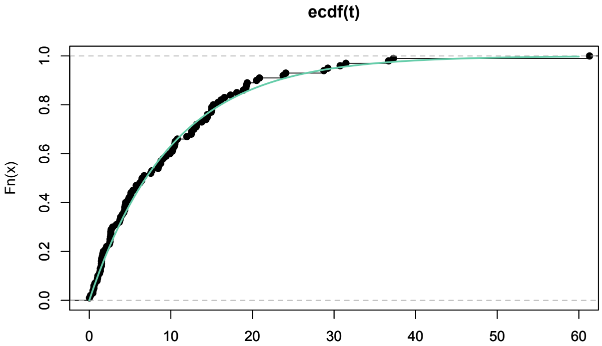
For the case $\beta = 1/10$,

```
> v<-runif(100)#simulate 100 uniform random variables
> t<--10*log(v)
> q<-1:3/4;  -10*log(1-q) #distributional quartiles
[1]  2.876821  6.931472 13.862944
> quantile(t,q)#simulation quantiles
      25%       50%       75%
 2.585831  6.609302 14.473724
```

Probability Transform

For the graphs

```
> plot(ecdf(t),xlim=c(0,60),ylim=c(0,1))
> par(new=TRUE)
> curve(1-exp(-x/10),0,60,ylim=c(0,1),
   col="aquamarine3",xlab="",ylab="",lwd=2)
```



Probability Transform

For the Pareto distribution

$$F_X(x) = 1 - x^{-3}, \quad x > 1$$

we have

$$u = 1 - x^{-3}, \quad 1 - u = x^{-3}, \quad (1-u)^{1/3} = x$$

```
> v<-runif(200) #v=1-u is uniform
> x<-v^-1/3     #use the probability transform
> q<-1:3/4      #check quartiles from distribution
> (1-q)^(-1/3)
[1] 1.100642 1.259921 1.587401
> quantile(x)   #against the simulation
      0%      25%      50%      75%     100%
1.002175 1.085188 1.278154 1.617686 7.716294
```