3.8 Functions of a Random Variable

This section introduces a new and important topic: determining the distribution of a function of a random variable. We suppose that there is a random variable $X$ and a function $r$ defined on the support of $X$. Interest lies in a new random variable constructed from $X$, namely $Y = r(X)$.

Example Suppose that $X \sim \text{Binom}(n, p)$ and

$$r(x) = \begin{cases} 0, & x \leq np, \\ 1, & np < x. \end{cases}$$

is an indicator function. The random variable $Y = r(X)$ is also binomial with parameters $n = 1$ and $p^* = \Pr(Y = 1) = \sum_{x>np} \Pr(X = x)$.

**Theorem 3.8.1. Function of a discrete random variable** Let $X$ have a discrete distribution with p.f. $f$, and let $Y = r(X)$ for some function $r$ defined on the set of possible values of $X$. For each possible value $y$ of $Y$, the p.f. $g$ of $y$ is

$$g(y) = \Pr[r(X) = y] = \sum_{\{x|x(x) = y\}} f(x).$$

Example 3.8.2. Suppose that $X$ has a uniform discrete distribution on $S = \{1, 2, \ldots, 9\}$, and that $Y$ is the distance of a realization of $X$ from the median (5) of the distribution. The random variable $Y$ is defined as follows:

$$Y = r(X) = |X - 5|.$$ 

Then $r(S) = \{0, 1, 2, 3, 4\}$, and

$$\Pr(Y = 0) = \Pr(X = 5) = 1/9,$$

whereas $\Pr(Y = 1) = \Pr(X = 4) + \Pr(X = 6) = 2/9$. The p.f. of $Y$ is

$$g(y) = \begin{cases} 1/9, & y = 0, \\ 2/9, & y \in \{1, 2, 3, 4\}, \\ 0, & \text{otherwise}. \end{cases}$$

**Functions of a continuous random variable.**

Determining the distribution of a function of a continuous random variable is somewhat more involved. There are several methods, none or which is best for all situations. The first approach determines the c.d.f. of $Y = r(X)$ from basic principles, and is illustrated by the
Example 3.8.3. Let $Z$ be a random variable identifying the rate at which customers in a queue are served\footnote{In this example, it is supposed that the rate varies continuously over time (for many reasons perhaps).}, and let $F$ denote the c.d.f. of $Z$. The support of $Z$ is $[0, \infty)$. Then $Y = r(Z) = 1/Z$ is the average waiting time of customers in the queue. The support of $Y$ is $(0, \infty)$. Since $Z$ is continuous, $\Pr(Z = 0) = 0$, and we need not be concerned with the possibility that $Y$ will be undefined.

We can determine the c.d.f. of $Y$ as follows:

$$G(y) = \Pr(Y \leq y)$$
$$= \Pr[r(Z) \leq y]$$
$$= \Pr(Z^{-1} \leq y)$$
$$= \Pr(Z \geq y^{-1}) = \Pr(Z > y^{-1})$$
$$= 1 - F(y^{-1}).$$

The general approach is described as follows. Suppose that the p.d.f. of $X$ is $f$ and that $Y = r(X)$. for each $y \in \mathbb{R}$, the c.d.f. of $Y$ is given by

$$G(y) = \Pr(Y \leq y)$$
$$= \Pr[r(X) \leq y]$$
$$= \int_{\{x \mid r(x) \leq y\}} f(x) \, dx.$$ 

If the random variable $Y$ is also continuous in distribution, then its p.d.f. can be computed according to

$$g(y) = \frac{dG(y)}{dy}$$

at every point $y$ where $G$ is differentiable.

Example 3.8.4 Suppose that $X \sim \text{Unif}[-1, 1]$. Thus, $f(x) = .5I_{[-1,1]}(x)$. Let $Y = X^2$. The support of $Y$ is $[0, 1]$, and for every $y \in [0, 1]$,

$$G(y) = \Pr(Y \leq y)$$
$$= \Pr[X^2 \leq y]$$
$$= \Pr(-\sqrt{y} \leq X \leq \sqrt{y})$$
$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2} \, dx$$
$$= \sqrt{y}.$$
Since $G(y) = \sqrt{y}$ is differentiable on $[0, 1]$,

$$g(y) = \frac{dG(y)}{dy} = \frac{1}{2\sqrt{y}}I_{[0,1]}(y).$$

**Theorem 3.8.2.** Suppose that the function $r$ is a linear function; specifically \( r(x) = ax + b \) where \( 0 \neq a, b, \in \mathbb{R} \). Then, the p.d.f. of $Y = r(X)$ is

$$g(y) = \frac{1}{|a|}f\left(\frac{y - b}{a}\right), y \in \mathbb{R}.$$ 

The proof is left as a homework exercise.

---

The *Probability Integral Transform* provides a method of generating realizations of a random variable knowing only the c.d.f. $F$ of the random variable $X$. The probability integral transformation only is useful if $F$ is strictly increasing on the support of $X$ (so that the inverse function $F^{-1}$ exists on $(0, 1)$). The transformation of $X$ is this function $y = r(x) = F(x)$.

**Example 3.8.5.** Let \( f(x) = \exp(-x)I_{(0,\infty)}(x) \), so that

$$F(x) = \begin{cases} 
1 - \exp(-x), & 0 < x \\
0, & \text{otherwise.} 
\end{cases}$$

**Theorem 3.8.3.** Suppose that $X$ has a continuous c.d.f. $F$ and let $Y = F(X)$. Then $Y \sim \text{Unif}[0, 1]$. The transformation from $X$ to $Y$ is called the *probability integral transformation*.

The text proves the theorem for the general case, but here for simplicity, suppose that $F$ is monotonic (increasing) and therefore invertible on $(0, 1)$.

Since $Y = F(X)$, the support of $Y$ is $[0, 1]$ and the c.d.f. of $Y$ is

$$G(y) = \Pr(Y \leq y) = \Pr[F(X) \leq y] = \Pr[X \leq F^{-1}(y)] = F[F^{-1}(y)] = y.$$ 

Also,

$$\frac{dG(y)}{dy} = I_{[0,1]}(y).$$
Thus, \( Y \sim \text{Unif}[0, 1] \).

The following corollary is used in practical applications to generate random variates from a particular distribution.

**Corollary 3.8.1.** Let \( Y \sim \text{Unif}(0, 1) \) and suppose that \( F \) is a c.d.f. with quantile function \( F^{-1} \). Then, \( X = F^{-1}(Y) \) has c.d.f. \( F \).

To prove the corollary, we set \( X = F^{-1}(Y) \) and proceed as above:

\[
G(x) = \Pr(X \leq x) \\
= \Pr[F^{-1}(Y) \leq x] \\
= \Pr[Y \leq F(x)] \\
= F(x),
\]

since \( Y \sim \text{Unif}(0, 1) \Rightarrow F(y) = y \) for \( 0 \leq y \leq 1 \).

Operationally, we generate random variates \( y_1, \ldots, y_n \) from the Unif(0, 1) distribution, and then, given some \( F^{-1} \), random variates following the c.d.f. \( F \) are generated by computing \( x_k = F^{-1}(y_k) \) where \( y_k \) is a random variate from the Unif(0, 1) distribution.

**Question 1** : Generate \( n = 10,000 \) independent observations from the c.d.f.

\[
F(x) = \begin{cases} 
1 - \exp(-x), & 0 < x \\
0, & \text{otherwise}.
\end{cases}
\]

Construct an approximation of \( F \) by plotting the sorted observations on the horizontal axis and \( p = (1/n, 2/n, \ldots, 1) \) on the vertical axis. The following code may be helpful.

```r
n = 10000
randomVariates = function(n) {
  x = runif(n)
  x = -log(1-x)
}
x = randomVariates(n)
y = seq(from=1/n,length = n,to=1)
plot(sort(x),y,type = 'l',xlab='x',ylab='F(x)')
```

**Remark** A continuous function \( r \) of a continuous random variable \( X \) need not generate a continuous random variable; for instance, \( Y = r(X) \) is not continuous if \( r(x) = c \forall x \). In fact, \( r(Y) \) is generally not even considered a random variable because \( \Pr[r(X) = c] = 1 \).
Alternatively, we may say that the distribution of $Y$ is degenerate with $\Pr(Y = c) = 1$.

**Direct derivation of the p.d.f. when $r$ is one-to-one and differentiable** In contrast to the last remark, computation of the c.d.f. and p.d.f. are direct when $r$ is one-to-one and differentiable (rather than only continuous). A function $r$ is differentiable on the open interval $(a, b)$ if the derivative of $r$ exists at each point in $(a, b)$. Suppose now that $r$ one-to-one, continuous, and either strictly increasing or strictly decreasing on $(a, b)$. Thus, the inverse function $s = r^{-1}$ exists on the image $r[(a, b)] = (\alpha, \beta)$. The inverse function is also continuous and its derivative is

$$
\frac{ds(y)}{dy} = \frac{dr^{-1}(y)}{dy} = \left( \frac{dr(x)}{dx} \right)_{x=s(y)}^{-1}.
$$

For example, $y = r(x) = e^x$ is one-to-one and differentiable on $\mathbb{R}$, and $s(y) = r^{-1}(y) = \log(y)$. The image is $r(\mathbb{R}) = (0, \infty)$. Then $ds(y)/dy = y^{-1}$ by direct calculation, and implicitly,

$$
\frac{ds(y)}{dy} = \left( \frac{dr(x)}{dx} \right)_{x=s(y)}^{-1} = \frac{1}{e^x} \bigg|_{x=\log(y)} = y^{-1}.
$$

**Theorem 3.8.4.** Let $X$ be a random variable with p.d.f. $f$ such that $\Pr(a < X < b) = 1$ ($a$ and $b$ may be infinite). Suppose that $r$ is one-to-one and differentiable on $(a, b)$ and the image of $r$ is $r[(a, b)] = (\alpha, \beta)$, and let $s = r^{-1}$ denote the inverse function of $r$ on $(\alpha, \beta)$. Then the p.d.f. of $Y = r(X)$ is

$$
g(y) = \begin{cases} 
  f[s(y)] \left| \frac{ds(y)}{dy} \right|, & \alpha < y < \beta, \\
  0, & \text{otherwise}.
\end{cases}
$$

The proof in the textbook treats two cases: $r$ is increasing on $(a, b)$ and $r$ is decreasing on $(a, b)$. Suppose that $r$ is decreasing on $(a, b)$. Then, $s$ is decreasing on $(\alpha, \beta)$, and

$$
G(y) = \Pr[r(X) \leq y] = \Pr[X \geq s(y)] = 1 - F[s(y)].
$$

Applying the chain rule yields

$$
g(y) = \frac{dG(y)}{dy} = -f[s(y)] \frac{ds(y)}{dy} = f[s(y)] \left| \frac{ds(y)}{dy} \right|,
$$

since $s$ decreasing implies $\frac{ds(y)}{dy} < 0$. 

Example 3.8.8. Suppose that in an resource-unlimited environment, the microbial population size after 5 time steps is \( r(x) = 10e^{5x} \) where \( x \) is the growth rate. Suppose further that the growth rate \( X \) is a random variable with p.d.f.

\[
f(x) = \begin{cases} 
3(1 - x)^2 & 0 < x < 1, \\
0, & \text{otherwise.}
\end{cases}
\]

The population size distribution is of interest, so let \( Y = r(X) = 10e^{5X} \) denote the population size random variable. \( r(x) \) is differentiable and continuous on \((0, 1)\), and the inverse function is \( s(y) = \log(y/10)/5 \Rightarrow ds(y)/dy = 1/(5y) \). Applying Theorem 3.8.4 leads to the p.d.f. of \( Y \):

\[
g(y) = \begin{cases} 
\frac{3(1 - \log(y/10)/5)^2}{5y} & 10 < x < 10e^5, \\
0, & \text{otherwise.}
\end{cases}
\]

The support of \( Y \) is determined by computing \( r(0) = 10 \) and \( r(1) = 10e^5 \).