Chapter 6: Functions of Random Variables

Professor Ron Fricker
Naval Postgraduate School
Monterey, California
Goals for this Chapter

• Introduce methods useful for finding the probability distributions of functions of random variables
  – Method of distribution functions
  – Method of transformations
• In OA3102 you’re going to learn to do inference
  – Using a sample of data, you’re going to make quantitative statements about the population
  – For example, you might want to estimate the mean $\mu$ of the population from a sample of data $Y_1, \ldots, Y_n$

• To estimate the mean $\mu$ of the population (distribution) you will use the sample average

$$\bar{Y} = \sum_{i=1}^{n} Y_i$$
Functions of Random Variables

• But in doing statistics, you will quantify the “goodness” of the estimate
  – This error of estimation will have to be quantified probabilistically since the estimates themselves are random
  – That is, the observations (the $Y_i$s) are random observations from the population, so they are random variables, and so are any statistics that are functions of the random variables
  – So, $\bar{Y}$ is a function of other random variables and the question then is how to find its probability distribution?
But Functions of Random Variables Arise in Other Situations As Well

• You’re analyzing a subsurface search and detection operation
  – Data shows that the time to detect one target is uniformly distributed on the interval (0, 2 hrs)
  – But you’re interested in analyzing the total time to sequentially detect \( n \) targets, so you need to know the distribution of the sum of \( n \) detection times

• You’re analyzing a depot-level repair activity
  – It’s reasonable to assume that the time to receipt of one repair part follows an exponential distribution with \( \beta = 5 \) days
  – But the end-item is not repaired until the last part is received, so you need to know the distribution of the maximum receipt time of \( n \) parts
Functions of Random Variables

• To determine the probability distribution for a function of \( n \) random variables, \( Y_1, Y_2, \ldots, Y_n \), we must find the joint probability distribution of the r.v.s

• To do so, we will assume observations are:
  – Obtained via (simple) random sampling
  – Independent and identically distributed

• Thus:
  \[
p(y_1, y_2, \ldots, y_n) = p(y_1)p(y_2) \cdots p(y_n)
  \]
or
  \[
f(y_1, y_2, \ldots, y_n) = f(y_1)f(y_2) \cdots f(y_n)
  \]
Methods for Finding Joint Distributions

• Three methods:
  – Method of distribution functions
  – Method of transformations
  – Method of moment-generating functions

• We’ll do the first two
  – Method of moment-generating functions follows from Chapter 3.9 (which we skipped)
    • If two random variables have the same moment-generating functions, then they also have the same probability distributions
  – Beyond the scope of this class (unless you’re an OR Ph.D. student)
Section 6.3: Method of Distribution Functions

• The idea: You want to determine the distribution function of a random variable $U$
  – $U$ is a function of another random variable $Y$
  – You know the pdf of $Y, f(y)$

• The methodology “by the numbers”:
  1. For $Y$ with pdf $f(y)$, specify the function $U$ of $Y$
  2. Then find $F_U(u) = P(U \leq u)$ by directly integrating $f(y)$ over the region for which $U \leq u$
  3. Then the pdf for $U$ is found by differentiating $F_U(u)$

• Let’s look at an example…
Textbook Example 6.1

• For clarity, let’s dispense with the story and get right to the problem: $Y$ has pdf

$$f(y) = \begin{cases} 2y, & 0 \leq y \leq 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find the pdf of $U, f_U(u)$, where $U = 3Y - 1$.

• Solution:
Can Also Apply the Method to Bivariate (and Multivariate) Distributions

• The idea:
  – In the bivariate case, there are random variables $Y_1$ and $Y_2$ with joint density $f(y_1, y_2)$
  – There is some function $U = h(Y_1, Y_2)$ of $Y_1$ and $Y_2$
    • I.e., for every point $(y_1, y_2)$ there is one and only one value of $U$
  – If there exists a region of values $(y_1, y_2)$ such that $U \leq u$, then integrating $f(y_1, y_2)$ over this region gives $P(U \leq u) = F_U(u)$
    – Then get $f_U(u)$ by differentiation
  
  • Again, let’s do some examples…

3/15/15
Textbook Example 6.2

• Returning to Example 5.4, we had the following joint density of $Y_1$ and $Y_2$:

$$f(y_1, y_2) = \begin{cases} 
3y_1, & 0 \leq y_2 \leq y_1 \leq 1 \\
0, & \text{elsewhere}
\end{cases}$$

Find the pdf of $U, f_U(u)$, where $U = Y_1 - Y_2$. Also find $E(U)$.

• Solution:

Textbook Example 6.2 Solution (cont’d)
• Graphs of the cdf and pdf of $U$:

**FIGURE 6.2**
Distribution and density functions for Example 6.2

(a) Distribution Function

(b) Density Function

Textbook Example 6.3

• Let \((Y_1, Y_2)\) denote a random sample of size \(n = 2\) from the uniform distribution on \((0,1)\). Find the pdf for \(U = Y_1 + Y_2\).

• Solution:

FIGURE 6.4
The region
$y_1 + y_2 \leq u$ for
$0 \leq u \leq 1$
Textbook Example 6.3 Solution (cont’d)

• Using a geometry-based approach…
**Textbook Example 6.3 Solution (cont’d)**

![Diagram showing the region $y_1 + y_2 \leq u$, $1 < u \leq 2$]

**FIGURE 6.5**
The region $y_1 + y_2 \leq u$, $1 < u \leq 2$.

**Textbook Example 6.3 Solution (cont’d)**

![Diagram showing distribution and density functions for Example 6.3](image)

**Figure 6.6**
Distribution and density functions for Example 6.3

Empirically Demonstrating the Result

• Let’s use R to verify the Example 6.3 result:

```r
> y1 <- runif(100000)  # 100,000 random observations from a U(0,1) distribution
> y2 <- runif(100000)  # Another 100,000 observations independent from y1
> u <- y1 + y2         # Calculate u, a function of y1 and y2
> hist(u,breaks=100,freq=FALSE)  # An empirical estimate of the pdf
```

Very close. Of course, there’s “noise” related to the stochastic nature of the demonstration.
Distribution Function Method Summary

Let $U$ be a function of the random variables $Y_1, Y_2, \ldots, Y_n$. Then:

1. Find the region $U = u$ in the $(y_1, y_2, \ldots, y_n)$ space
2. Find the region $U \leq u$
3. Find $F_U(u) = P(U \leq u)$ by integrating $f(y_1, y_2, \ldots, y_n)$ over the region $U \leq u$
4. Find the density function $f_U(u)$ by differentiating $F_U(u): f_U(u) = dF_U(u)/du$
Illustrating the Method

- Consider the case $U = h(Y) = Y^2$, where $Y$ is continuous with cdf $F_Y(y)$ and pdf $f_Y(y)$

1. Find the region $U = u$ in the $(y_1, y_2, \ldots, y_n)$ space

Solution:
As Figure 6.7 shows, $Y^2 = u$

or $Y = \pm \sqrt{u}$

2. Find the region where \( U \leq u \)

Solution: By inspection of Figure 6.7, it should be clear that \( U \leq u \) whenever 
\[-\sqrt{u} \leq Y \leq \sqrt{u}\]

3. Find \( F_U(u) = P(U \leq u) \) by integrating \( f(y_1, y_2, \ldots, y_n) \) over the region \( U \leq u \)

Solution: There are two cases to consider.
First, if \( u \leq 0 \) then it must be (see Fig. 6.7)
\[F_U(u) = P(U \leq u) = P(Y^2 \leq u) = 0\]
Second, if \( u > 0 \) …
Illustrating the Method Continued

\[ F_U(u) = \int_{-\sqrt{u}}^{\sqrt{u}} f(y)dy = F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) \]

So, we can write the cdf of \( U \) as

\[
F_U(u) = \begin{cases} 
F_Y(\sqrt{u}) - F_Y(-\sqrt{u}), & u > 0 \\
0, & \text{otherwise}
\end{cases}
\]

4. Finally, find the pdf by differentiating \( F_U(u) \):

\[
f_U(u) = \begin{cases} 
\frac{1}{2\sqrt{u}} [f_Y(\sqrt{u}) + f_Y(-\sqrt{u})], & u > 0 \\
0, & \text{otherwise}
\end{cases}
\]
Textbook Example 6.4

• Let $Y$ have pdf

\[ f_Y(y) = \begin{cases} 
(y + 1)/2, & -1 \leq y \leq 1 \\
0, & \text{elsewhere} 
\end{cases} \]

Find the pdf for $U = Y^2$.

• Solution:
Textbook Example 6.5

• Let $U$ be a uniform r.v. on $(0,1)$. Find a transformation $G(U)$ such that $G(U)$ has an exponential distribution with mean $\beta$.

• Solution:
Empirically Demonstrating the Result

• Let’s use R to verify the Example 6.5 result:

```r
> u <- runif(1000000)  # Generate a million observations from U(0,1) distribution
> beta <- 1  # Set beta to 1 for this example
> y <- -beta*log(1-u)  # Transform from U(0,1) to Exp(beta=1)
> hist(y,breaks=100,freq=FALSE)  # An empirical estimate of the pdf
> exp_rvs <- rexp(1000000,beta)  # Generate a million observations from Exp(beta=1)
> hist(exp_rvs,breaks=100,freq=FALSE)  # An empirical estimate of the pdf using Exp(1)
> plot(sort(y),sort(exp_rvs))  # Comparing the two via an empirical quantile-quantile plot
```
Simulating Exponential R.V.s Using $U(0,1)$s

• What the last example shows is that we can use computer subroutines designed to generate $U(0,1)$ random variables to generate other types of random variables
  – It works as long as the cdf $F(y)$ has a unique inverse $F^{-1}(\cdot)$
  – Not always the case; for those there are other methods
  – And, the beauty of R and other modern software packages is that most or all of this work is already done for you
Section 6.3 Homework

- Do problems 6.1, 6.4, 6.11
Section 6.4: Method of Transformations

- A (sometimes simpler) method to use, if $U = h(y)$ is an increasing / decreasing function
  - Remember, an increasing function is one where, if $y_1 < y_2$, then $h(y_1) < h(y_2)$ for all $y_1$ and $y_2$:

  ![Graph of an increasing function]

  - Similarly, a decreasing function is one where, if $y_1 > y_2$, then $h(y_1) > h(y_2)$ for all $y_1$ and $y_2$
Method of Transformations

• Note that if $h(y)$ is an increasing function of $y$ then $h^{-1}(u)$ is an increasing function of $u$
  
  – If $u_1 < u_2$, then $h^{-1}(u_1) = y_1 < y_2 = h^{-1}(u_2)$

• So, the set of points $y$ such that $h(y) \leq u_1$ is the same as the set of points $y$ such that $y \leq h^{-1}(u_1)$

• This allows us to derive the following result…

Method of Transformations

\[ F_U(u) = P(U \leq u) = P[h(Y) \leq u] \]
\[ = P \{ h^{-1}[h(Y)] \leq h^{-1}(u) \} \]
\[ = P[Y \leq h^{-1}(u)] \]
\[ = F_Y[h^{-1}(u)] \]

• And, from this result, it follows that

\[ f_U(u) = \frac{dF_U(u)}{du} \]
\[ = \frac{dF_Y[h^{-1}(u)]}{du} \]
\[ = f_Y(h^{-1}(u)) \frac{d[h^{-1}(u)]}{du} \]
Textbook Example 6.6

• Returning to Example 6.1, solve via the Method of Transformations, where $Y$ has pdf

$$f(y) = \begin{cases} 2y, & 0 \leq y \leq 1 \\ 0, & \text{elsewhere} \end{cases}$$

i.e., find the pdf of $U, f_U(u)$, where $U = 3Y - 1$.

• Solution:
Now, for Decreasing Functions…

• Now, if \( h \) is a decreasing function (e.g., Figure 6.9) then if \( u_1 < u_2 \), we have
  \[ h^{-1}(u_1) = y_1 > y_2 = h^{-1}(u_2) \]

• From this, it follows that
  \[
  F_U(u) = P(U \leq u) = P[Y \geq h^{-1}(u)] \\
  = 1 - F_Y[h^{-1}(u)]
  \]

• Differentiating gives…

Putting It All Together

\[ f_U(u) = -f_Y[h^{-1}(u)] \frac{d[h^{-1}(u)]}{du} \]

\[ = f_Y[h^{-1}(u)] \left| \frac{d[h^{-1}(u)]}{du} \right| \]

• **Theorem:** Let \( Y \) have pdf \( f_Y(y) \). If \( h(y) \) is either increasing or decreasing for all \( y \) such that \( f_Y(y) > 0 \),* then \( U = h(Y) \) has pdf

\[ f_U(u) = f_Y[h^{-1}(u)] \left| \frac{d[h^{-1}(u)]}{du} \right| \]

* Note that the function only has to be increasing or decreasing over the support of the pdf of \( y \): \( \{ y : f_Y(y) > 0 \} \).
Textbook Example 6.7

• Let $Y$ have pdf

$$f(y) = \begin{cases} 
2y, & 0 \leq y \leq 1 \\
0, & \text{elsewhere}
\end{cases}$$

Find the pdf of $U = -4Y + 3$.

• Solution:
Textbook Example 6.7 Solution (cont’d)
Let $Y_1$ and $Y_2$ have the joint pdf

$$f(y_1, y_2) = \begin{cases} e^{-(y_1+y_2)}, & y_1 \geq 0, y_2 \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the pdf for $U = Y_1 + Y_2$.

Solution:
Textbook Example 6.8 Solution (cont’d)
Textbook Example 6.9

• Example 5.19 had the joint pdf

\[ f(y_1, y_2) = \begin{cases} 2(1 - y_1), & 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 \\ 0, & \text{otherwise} \end{cases} \]

Find the pdf for \( U = Y_1Y_2 \) as well as \( E(U) \).

• Solution:
Distribution Function Method Summary

Let $U = h(Y)$, where $h(Y)$ is either an increasing or decreasing function of $y$ for all $y$ such that $f_Y(y) > 0$.

1. Find the inverse function $y = h^{-1}(u)$

2. Evaluate $\frac{d[h^{-1}(u)]}{du}$

3. Find $f_U(u)$ by

$$f_U(u) = f_Y[h^{-1}(u)] \left| \frac{d[h^{-1}(u)]}{du} \right|$$
Section 6.4 Homework

- Do problems 6.23, 6.24, 6.25
**Summary**

- We often need to know the probability distribution for functions of random variables
  - This is a common requirement for statistics and will arise throughout OA3102 and OA3103
  - It also comes up in lots of other ORSA probability problems
- Today we can often simulate and get an empirical estimate of a distribution, as we did here for a couple of examples, but a closed-form analytical result is better
Summary Continued

• We covered two methods
  – Distribution function method (chpt. 6.3)
  – Transformation method (chpt. 6.4)

and we skipped over two methods
  – Moment-generating function method (chpt. 6.5)
  – Multivariable transforms w/ Jacobians (chpt. 6.6)

• Which is better depends on the problem
  – Sometimes one or more are inapplicable
  – Sometimes one is easier than the other
  – Sometimes one or more works and others don’t
Summary Continued

• The sections we skipped prove some results you need to know for OA3102 and beyond:

1. If $Y \sim N(\mu, \sigma^2)$ then $Z = \frac{Y - \mu}{\sigma} \sim N(0, 1)$

2. If $Z \sim N(0, 1)$ then $Z^2 \sim \chi^2(\nu = 1)$

3. If $Z_i \sim N(0, 1), i = 1, \ldots, n$ then $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$
4. Let $Y_1, Y_2, \ldots, Y_n$ be independent normally distributed random variables with $E(Y_i) = \mu_i$ and $V(Y_i) = \sigma^2_i$ for $i = 1, 2, \ldots, n$, and let $a_1, a_2, \ldots, a_n$ be constants. Then

$$U = \sum_{i=1}^{n} a_i Y_i = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n$$

is a normally distributed random variable with

$$E(U) = \sum_{i=1}^{n} a_i \mu_i \quad \text{and} \quad V(U) = \sum_{i=1}^{n} a_i^2 \sigma^2_i$$
What We Have Just Learned

• Introduced methods useful for finding the probability distributions of functions of random variables
  – Method of distribution functions
  – Method of transformations