Chapter 6

Dr Wayne Stewart

Bivariate Probability Distributions and Sampling Distributions

OBJECTIVE

To introduce the concepts of a bivariate probability distribution, covariance, and independence; to show you how to find the expected value and variance of a linear function of random variables; to find and identify the probability distribution of a statistic (a sampling distribution)

CONTENTS

$$P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)$$

$$
\begin{aligned}
P(A \cap B) &= P(X = x, Y = y) = p(x, y) \\
&= p_1(x) p_2(y \mid x) \\
&= p_2(y) p_1(x \mid y)
\end{aligned}
$$

Definition 6.1

The **joint probability distribution** $\boldsymbol{p(x,y)}$ for two discrete random variables, X and Y—called a **bivariate distribution**—is a table, graph, or formula that gives the values of $p(x, y)$ for every combination of values of X and Y.

Requirements for a Discrete Bivariate Probability Distribution for *X* and *Y*

1. $0 \leq p(x, y) \leq 1$ for all values of X and Y
2. $\sum_{y} \sum_{x} p(x, y) = 1$

(*Note*: The symbol $\sum_{y} \sum_{x}$ denotes summation over all values of both X and Y.)

TABLE 6.1 Bivariate Probability Distribution for X and Y

		$X = x$			
		1	2	3	4
	0	0	.10	.20	.10
$Y = y$	1	.03	.07	.10	.05
	2	.05	.10	.05	0
	3	0	.10	.05	0

TABLE 6.1 Bivariate Probability Distribution for X and Y

		$X = x$			
		1	2	3	4
	0	0	.10	.20	.10
$Y = y$	1	.03	.07	.10	.05
	2	.05	.10	.05	0
	3	0	.10	.05	0

The marginal probability distribution $p_1(x)$ is given in the following table:

x	1	2	3	4
$p_1(x)$	.08	.37	.40	.15

Note from the table that $\Sigma_{x=1}^{4} p_1(x) = 1$

Definition 6.2

Let X and Y be discrete random variables and let $p(x, y)$ be their joint probability distribution. Then the **marginal (unconditional) probability distributions** of X and Y are, respectively,

$$p_1(x) = \sum_y p(x, y) \quad \text{and} \quad p_2(y) = \sum_x p(x, y)$$

(*Note*: We will use the symbol $\sum_y$ to denote summation over all values of Y.)

Definition 6.3

Let X and Y be discrete random variables and let $p(x, y)$ be their joint probability distribution. Then the **conditional probability distributions** for X and Y are defined as follows:

$$p_1(x \mid y) = \frac{p(x, y)}{p_2(y)} \quad \text{and} \quad p_2(y \mid x) = \frac{p(x, y)}{p_1(x)}$$

TABLE 6.1 Bivariate Probability Distribution for X and Y

		$X = x$			
		1	2	3	4
	0	0	.10	.20	.10
$Y = y$	1	.03	.07	.10	.05
	2	.05	.10	.05	0
	3	0	.10	.05	0

Therefore, the conditional probability distribution of X, given that $Y = 2$, is as shown in the following table:

x	1	2	3	4
$p_1(x \mid 2)$	.25	.50	.25	0

6.7 *Red lights on truck route.* A special delivery truck travels from point A to point B and back over the same route each day. There are three traffic lights on this route. Let X be the number of red lights the truck encounters on the way to delivery point B and let Y be the number of red lights the truck encounters on the way back to delivery point A. A traffic engineer has determined the joint probability distribution of X and Y shown in the table.

		$X = x$			
		0	1	2	3
	0	.01	.02	.07	.01
$Y = y$	1	.03	.06	.10	.06
	2	.05	.12	.15	.08
	3	.02	.09	.08	.05

Definition 6.4

The **bivariate joint probability density function** $f(x, y)$ for two continuous random variables X and Y is one that satisfies the following properties:

1. $f(x, y) \geq 0 \quad \text{for all values of } X \text{ and } Y$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) = 1$
3. $P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)\, dx\, dy \quad \text{for all constants } a, b, c, \text{ and } d$

Definition 6.5

Let f(*x*, *y*) be the joint density function for *X* and *Y*. Then the **marginal density functions** for *X* and *Y* are

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \quad \text{and} \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

Definition 6.6

Let *f*(*x*, *y*) be the joint density function for *X* and *Y*. Then the **conditional density functions** for *X* and *Y* are

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} \quad \text{and} \quad f_2(y \mid x) = \frac{f(x, y)}{f_1(x)}$$

Example 6.4

Joint Density Function for Continuous Random Variables

Suppose the joint density function for two continuous random variables, *X* and *Y*, is given by

$$f(x, y) = \begin{cases} cx & \text{if } 0 \leq x \leq 1;\ 0 \leq y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

Determine the value of the constant *c*.

Example 6.4

Joint Density Function for Continuous Random Variables

Suppose the joint density function for two continuous random variables, X and Y, is given by

$$f(x, y) = \begin{cases} cx & \text{if } 0 \le x \le 1; 0 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Determine the value of the constant c.

Solution

A graph of $f(x, y)$ traces a three-dimensional, wedge-shaped figure over the unit square ($0 \le x \le 1$ and $0 \le y \le 1$) in the (x, y)-plane, as shown in Figure 6.1. The value of c is chosen so that $f(x, y)$ satisfies the property

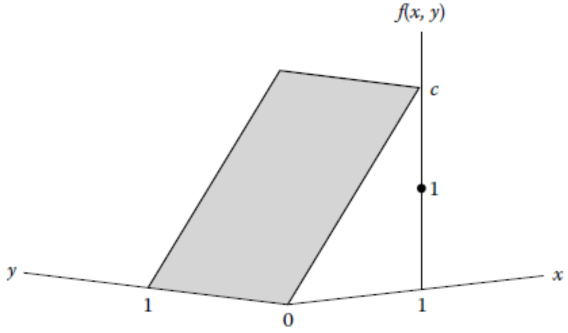
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

Performing this integration yields

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = \int_{0}^{1}\int_{0}^{1} cx\, dx\, dy$$

$$= c\int_{0}^{1}\int_{0}^{1} x\, dx\, dy = c\int_{0}^{1} \frac{x^2}{2}\bigg]_0^1 dy$$

$$= c\int_{0}^{1} \frac{1}{2}\, dy = \left(\frac{c}{2}\right)y\bigg]_0^1 = \frac{c}{2}$$



Setting this quantity equal to 1 and solving for c, we obtain

$$1 = \frac{c}{2} \quad \text{or} \quad c = 2$$

Therefore,

$$f(x, y) = 2x \quad \text{for } 0 \le x \le 1 \quad \text{and} \quad 0 \le y \le 1$$

Go through all examples!!

Example 6.7

Finding a Continuous Density Function

Refer to Examples 6.4–6.6. Find the conditional density function for X given Y, and show that it satisfies the property

$$\int_{-\infty}^{\infty} f_1(x \mid y)\, dx = 1$$

Solution

Using the marginal density function $f_2(y) = 1$ (obtained in Example 6.6) and Definition 6.6, we derive the conditional density function as follows:

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{2x}{1} = 2x, \quad 0 \leq x \leq 1$$

We now show that the integral of $f_1(x \mid y)$ over all values of X is equal to 1:

$$\int_0^1 f_1(x \mid y)\, dx = 2 \int_0^1 x\, dx = 2\left(\frac{x^2}{2}\right)\Bigg]_0^1 = 1$$

Example 6.8

Joint Density Function—Range of X Depends on Y

Suppose the joint density function for X and Y is

$$f(x, y) = \begin{cases} cx & \text{if } 0 \le x \le y;\ 0 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the value of c.

Solution

Refer to Figure 6.1. If we pass a plane through the wedge, diagonally between the points (0, 0) and (1, 1), and perpendicular to the (x, y)-plane, then the slice lying along the y-axis will have a shape similar to that of the given density function (graphed in Figure 6.2). The value of c will be larger than the value found in Example 6.4 because the volume of the solid shown in Figure 6.2 must equal 1.

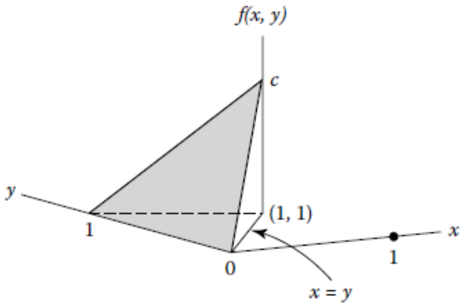


FIGURE 6.2
Graph of the joint density function for Example 6.8

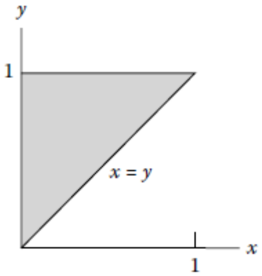


FIGURE 6.3
Region of integration for Example 6.8

We find c by integrating $f(x, y)$ over the triangular region (shown in Figure 6.3) defined by $0 \le x \le y$ and $0 \le y \le 1$, setting this integral equal to 1, and solving for c:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = \int_0^1\int_0^y cx\, dx\, dy = c\int_0^1 \frac{x^2}{2}\bigg]_0^y dy$$

$$= c\int_0^1 \frac{y^2}{2}\, dy = c\left(\frac{y^3}{6}\right)\bigg]_0^1 = \frac{c}{6}$$

Setting this quantity equal to 1 and solving for c yields $c = 6$; thus, $f(x, y) = 6x$ over the region of interest.

6.14 *Servicing an automobile.* The joint density of X, the total time (in minutes) between an automobile's arrival in the service queue and its leaving the system after servicing, and Y, the time (in minutes) the car waits in the queue before being serviced, is

$$f(x, y) = \begin{cases} ce^{-x^2} & \text{if } 0 \leq y \leq x; 0 \leq x < \infty \\ 0 & \text{elsewhere} \end{cases}$$

a. Find the value of c that makes $f(x, y)$ a probability density function.

b. Find the marginal density for X and show that

$$\int_{-\infty}^{\infty} f_1(x)\, dx = 1$$

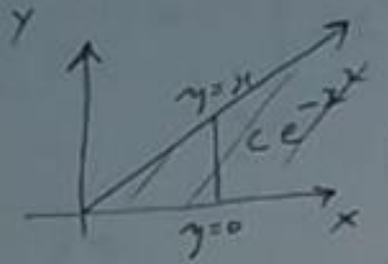
c. Show that the conditional density for Y given X is a uniform distribution over the interval $0 \leq Y \leq X$.

6.14

Y = time (min) for a car to wait before being serviced.
X = total time (min) between joining queue and finishing service

$$f(x,y) = \begin{cases} c e^{-x^2} & \text{if } 0 \le y \le x \,;\ 0 \le x \le \text{infty} \\ 0 & \text{else} \end{cases}$$

a)



use $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

$$= \iint c e^{-x^2}\,dy\,dx$$

$$= \int_{x=0}^{x=\infty}\left[\int_{y=0}^{y=x} c e^{-x^2}\,dy\right]dx = 1$$

$$= c\int_{x=0}^{x=\infty} e^{-x^2}\Big[y\Big]_{y=0}^{y=x} dx = c\int_{x=0}^{x=\infty} x e^{-x^2}\,dx = c\left[-\tfrac{1}{2}e^{-x^2}\right]_0^{\infty}$$

$$= c\,\tfrac{1}{2} = 1$$

$$c = 2$$

6.14 b)

b) $f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$

$= \int 2e^{-x^2}\,dy$

$= \int_{y=0}^{y=x} 2e^{-x^2}\,dy = 2e^{-x^2}\left[y\right]_{y=0}^{y=x} = 2xe^{-x^2}$

$\int_{-\infty}^{\infty} f_1(x)\,dx = \int 2xe^{-x^2}\,dx = \int_{x=0}^{x=\infty} 2xe^{-x^2}\,dx = \left[-e^{-x^2}\right]_{x=0}^{x=\infty} = 0 + e^0 = 1$

6.14 c)

$$f_2(y|x) = \frac{f(x,y)}{f_1(x)} \quad \text{by def}$$

$$f_1(x) = 2xe^{-x^2} \quad 0 \le x < \infty$$

$$f(x,y) = 2e^{-x^2} \quad 0 \le y \le x;\ 0 \le x < \infty$$

$$f_2(y|x) = \frac{2e^{-x^2}}{2xe^{-x^2}} = \frac{1}{x} \quad 0 \le y \le x$$

This is a uniform $(x = x - 0.)$ for y over the interval $\underline{0 \le y \le x}$

Definition 6.7

Let $g(X, Y)$ be a function of the random variables X and Y. Then the **expected value (mean)** of $g(X, Y)$ is defined to be

$$E[g(X, Y)] = \begin{cases} \displaystyle\sum_y \sum_x g(x, y)p(x, y) & \text{if } X \text{ and } Y \text{ are discrete} \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\, dx\, dy & \text{if } X \text{ and } Y \text{ are continuous} \end{cases}$$

THEOREM 6.1

Let c be a constant. Then the expected value of c is

$$E(c) = c$$

THEOREM 6.2

Let c be a constant and let $g(X, Y)$ be a function of the random variables X and Y. Then the expected value of $cg(X, Y)$ is

$$E[cg(X, Y)] = cE[g(X, Y)]$$

THEOREM 6.3

Let $g_1(X, Y), g_2(X, Y), \ldots, g_k(X, Y)$ be k functions of the random variables X and Y. Then the expected value of the sum of these functions is

$$E[g_1(X, Y) + g_2(X, Y) + \cdots + g_k(X, Y)]$$

$$= E[g_1(X, Y)] + E[g_2(X, Y)] + \cdots + E[g_k(X, Y)]$$

Definition 6.8

Let X and Y be discrete random variables with joint probability distribution $p(x, y)$ and marginal probability distributions $p_1(x)$ and $p_2(y)$. Then X and Y are said to be **independent** if and only if

$$p(x, y) = p_1(x)p_2(y) \quad \text{for all pairs of values of } x \text{ and } y$$

Definition 6.9

Let X and Y be continuous random variables with joint density function $f(x, y)$ and marginal density functions $f_1(x)$ and $f_2(y)$. Then X and Y are said to be **independent** if and only if

$$f(x, y) = f_1(x)f_2(y) \quad \text{for all pairs of values of } x \text{ and } y$$

Example 6.10

Demonstrating Dependence

Refer to Example 6.8 and determine whether X and Y are independent.

Solution

From Example 6.8, we determined that $f(x, y) = 6x$ when $0 \le x \le y$ and $0 \le y \le 1$. Therefore,

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_x^1 6x\, dy = 6xy\Big]_x^1$$

$$= 6x(1 - x) \quad \text{where } 0 \le x \le 1$$

Similarly,

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^y 6x\, dx = \frac{6x^2}{2}\Big]_0^y$$

$$= 3y^2 \quad \text{where } 0 \le y \le 1$$

You can see that $f_1(x)f_2(y) = 18x(1 - x)y^2$ is *not* equal to $f(x, y)$. Therefore, X and Y are *not* independent random variables.

THEOREM 6.4

If X and Y are independent random variables, then

$$E(XY) = E(X)E(Y)$$

Proof of Theorem 6.4 We will prove the theorem for the discrete case. The proof for the continuous case is identical, except that integration is substituted for summation. By the definition of expected value, we have

$$E(XY) = \sum_y \sum_x xyp(x, y)$$

But, since X and Y are independent, we can write $p(x, y) = p_1(x)p_2(y)$. Therefore,

$$E(XY) = \sum_y \sum_x xyp_1(x)p_2(y)$$

If we sum first with respect to X, then we can treat Y and $p_2(y)$ as constants and apply Theorem 6.2 to factor them out of the sum as follows:

$$E(XY) = \sum_y yp_2(y) \sum_x xp_1(x)$$

But,

$$\sum_x xp_1(x) = E(x) \quad \text{and} \quad \sum_y yp_2(y) = E(y)$$

Therefore,

$$E(XY) = E(X)E(Y)$$

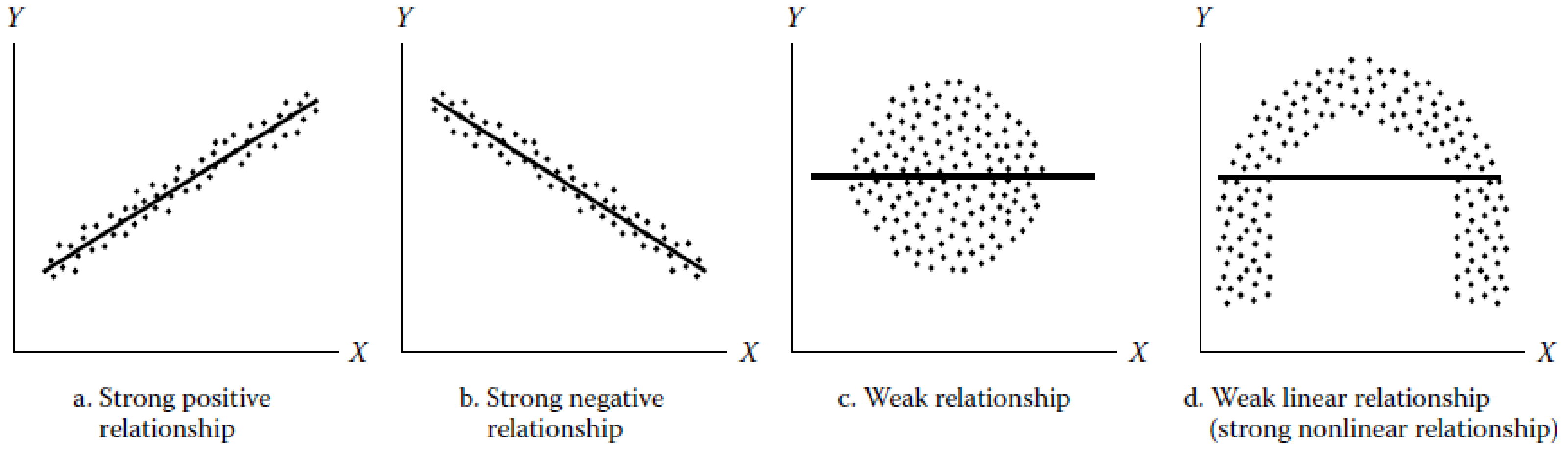
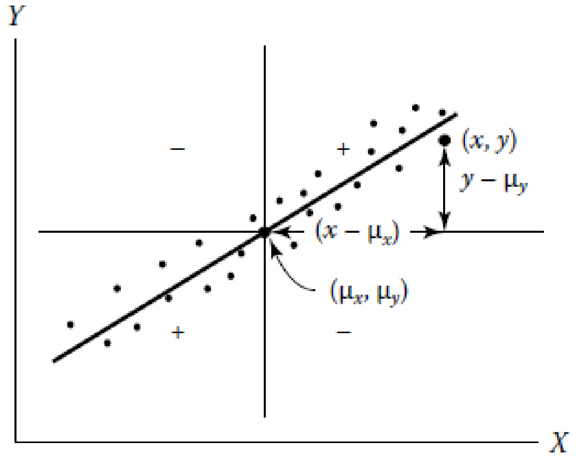


FIGURE 6.4
Linear relationships between X and Y

FIGURE 6.5
Signs of the cross-products $(x - \mu_x)(y - \mu_y)$



Definition 6.10

The **covariance** of two random variables, *X* and *Y*, is defined to be

$$\mathrm{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

THEOREM 6.5

$$\operatorname{Cov}(X, Y) = E(XY) - \mu_x\mu_y$$

Proof of Theorem 6.5 By Definition 6.10, we can write

$$\begin{aligned}\operatorname{Cov}(X, Y) &= E[(X - \mu_x)(Y - \mu_y)] \\ &= E(XY - \mu_x Y - \mu_y X + \mu_x\mu_y)\end{aligned}$$

Applying Theorems 6.1, 6.2, and 6.3 yields

$$\begin{aligned}\operatorname{Cov}(X, Y) &= E(XY) - \mu_x E(Y) - \mu_y E(X) + \mu_x\mu_y \\ &= E(XY) - \mu_x\mu_y - \mu_x\mu_y + \mu_x\mu_y \\ &= E(XY) - \mu_x\mu_y\end{aligned}$$

Example 6.11

Finding Covariance

Find the covariance of the random variables X and Y of Example 6.4.

Solution

The variables have joint density function $f(x, y) = 2x$ when $0 \le x \le 1$ and $0 \le y \le 1$. Then,

$$E(XY) = \int_0^1 \int_0^1 (xy)2x\,dx\,dy$$

$$= \int_0^1 2\left(\frac{x^3}{3}\right)\Bigg]_0^1 y\,dy = \frac{2}{3}\int_0^1 y\,dy = \frac{2}{3}\left(\frac{y^2}{2}\right)\Bigg]_0^1 = \frac{1}{3}$$

In Examples 6.5 and 6.6, we obtained the marginal density functions $f_1(x) = 2x$ and $f_2(y) = 1$. Therefore,

$$\mu_x = E(X) = \int_0^1 xf_1(x)\,dx = \int_0^1 x(2x)\,dx = 2\left(\frac{x^3}{3}\right)\Bigg]_0^1 = \frac{2}{3}$$

Furthermore, since Y is a uniform random variable defined over the interval $0 \le y \le 1$ (see Example 6.6), it follows from Section 5.4 that $\mu_y = \frac{1}{2}$. Then,

$$\mathrm{Cov}(X, Y) = E(XY) - \mu_x\mu_y = \frac{1}{3} - \left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = 0$$

THEOREM 6.6

If two random variables X and Y are independent, then

$$\text{Cov}(X, Y) = 0$$

Definition 6.11

The **coefficient of correlation** $\boldsymbol{\rho}$ for two random variables X and Y is

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sigma_x \sigma_y}$$

where $\boldsymbol{\sigma}_X$ and σ_Y are the standard deviations of X and Y, respectively.

Property of the Correlation Coefficient

$$-1 \leq \rho \leq 1$$

Proof

- $var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) \geq 0$
- $var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) \geq 0$

$$\mathrm{Var}\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) \geq 0 \quad \text{because} \quad \mathrm{var}(r.v) \geq 0$$

$$= \mathrm{Var}\left(\frac{X}{\sigma_X}\right) + \mathrm{Var}\left(\frac{Y}{\sigma_Y}\right) + 2\,\mathrm{cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)$$

$$= \frac{1}{\sigma_X^2} \cdot \sigma_X^2 + \frac{1}{\sigma_Y^2} \sigma_Y^2 + \frac{2\,\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

$$= 1 + 1 + 2\rho$$

$$= 2 + 2\rho$$

$$= 2(1+\rho)$$

hence $2(1+\rho) \geq 0$

$$1+\rho \geq 0$$

$$\rho \geq -1$$

$$\mathrm{var}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) \geq 0$$

$$= \frac{\mathrm{Var}(X)}{\sigma_X^2} + \frac{\mathrm{Var}(Y)}{\sigma_Y^2} - 2\,\mathrm{cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)$$

$$= 1 + 1 - 2\,\frac{\mathrm{cov}(X, Y)}{\sigma_X \sigma_Y}$$

$$= 2 - 2\rho$$

$$= 2(1-\rho)$$

$$\therefore\ 2(1-\rho) \geq 0$$

$$1-\rho \geq 0$$

$$\rho \leq 1$$

$$\therefore\ \boxed{-1 \leq \rho \leq 1}$$

THEOREM 6.7

Let Y be a continuous random variable with density function $f(y)$ and cumulative distribution $F(y)$. Then the density function of $W = F(y)$ will be a uniform distribution defined over the interval $0 \leq w \leq 1$, i.e.,

$$g(w) = 1 \qquad (0 \leq w \leq 1)$$

Proof of Theorem 6.7 Figure 6.10 shows the graph of $W = F(y)$ for a continuous random variable Y. You can see from the figure that there is a one-to-one correspondence between y values and w values, and that values of Y corresponding to values of W in the interval $0 \leq W \leq w$ will be those in the interval $0 \leq Y \leq y$. Therefore,

$$P(W \leq w) = P(Y \leq y) = F(y)$$

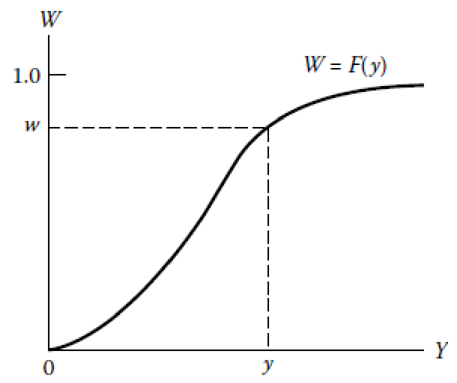
But since $W = F(y)$, we have $F(y) = w$. Therefore, we can write

$$G(w) = P(W \leq w) = F(y) = w$$

Finally, we differentiate over the range $0 \leq w \leq 1$ to obtain the density function:

$$g(w) = \frac{dF(w)}{dw} = 1 \; (0 \leq w \leq 1)$$

FIGURE 6.10
Cumulative distribution function $F(y)$



Example

Example 6.14

Generating a Random Sample

Use Theorem 6.7 to generate a random sample of $n = 3$ observations from an exponential distribution with $\beta = 2$.

Solution

The density function for the exponential distribution with $\beta = 2$ is

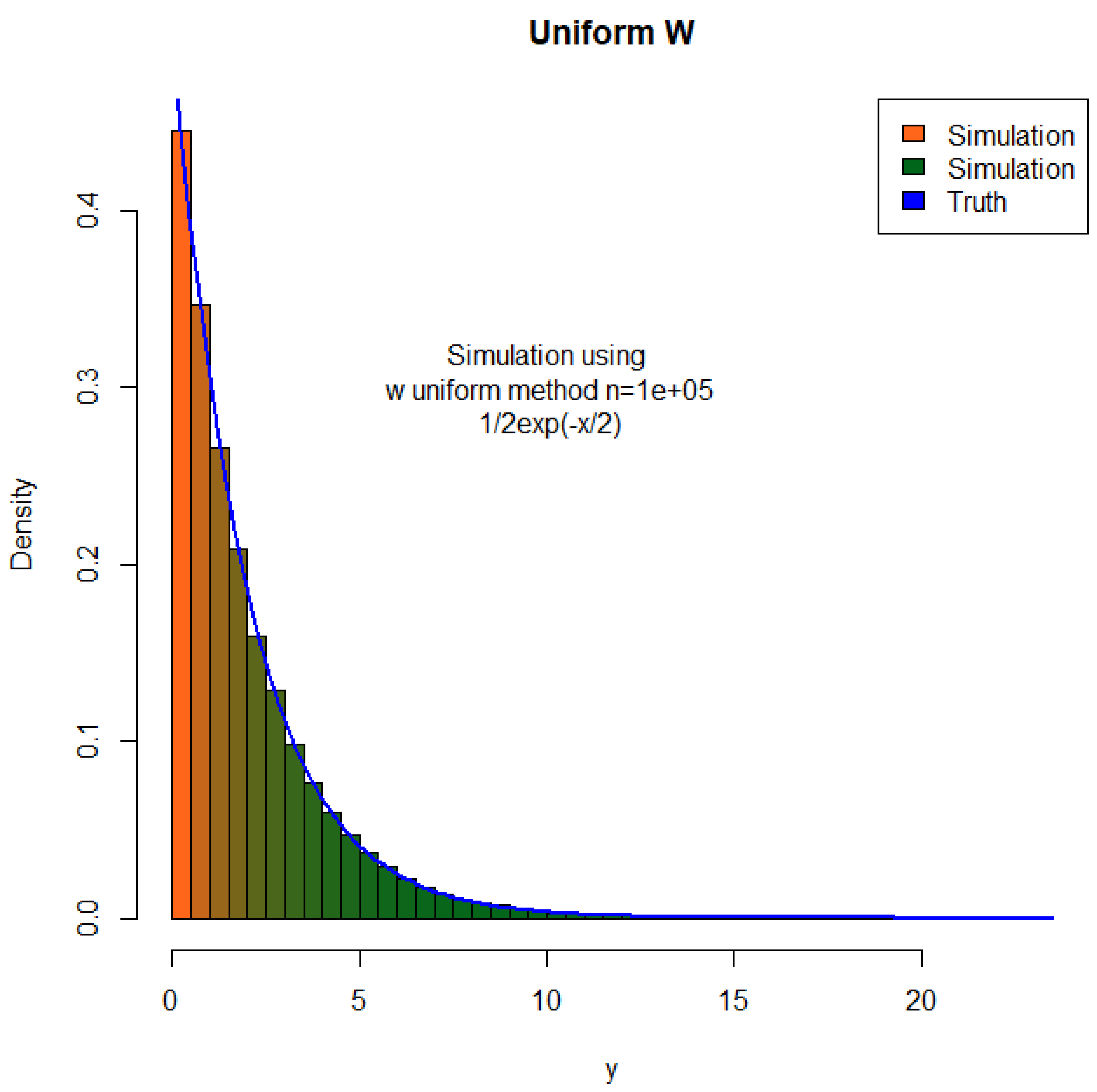
$$f(y) = \begin{cases} \dfrac{e^{-y/2}}{2} & \text{if } 0 \le y < \infty \\ 0 & \text{elsewhere} \end{cases}$$

$$F(y) = \int_{-\infty}^{y} f(t)\,dt = \int_{0}^{y} \frac{e^{-t/2}}{2}\,dt = -e^{-t/2}\Big]_{0}^{y} = 1 - e^{-y/2}$$

If we let $W = F(y) = 1 - e^{-y/2}$, then Theorem 6.7 tells us that W has a uniform density function over the interval $0 \le W \le 1$.

The true density

```
## w method of sampling
rmyexp=function(n, ...){
  graphics.off()
w=runif(n,0,1)
y=-2*log(1-w)
h=hist(y,plot=FALSE, ...)
coll = rgb(h$density/max(h$density),.4,.1)
windows()
hist(y,freq=FALSE,main="Uniform W",col = coll, ...)#
col=rainbow(length(h$mids)),...)
curve( exp(-x/2)/2,add=TRUE,col="Blue",lwd=2)
text(10,.3, paste0("Simulation using\n w uniform method ",
"n=",n,"\n ", "1/2exp(-x/2)"))
legend("topright", legend = c("Simulation",
"Simulation","Truth"), fill=c(coll[1],coll[length(coll)], "Blue"))
dev.new(noRStudioGD = TRUE)
df=data.frame(y)
library(ggplot2)
g = ggplot(df, aes(x=y)) + geom_histogram(aes(fill=..density..),
bins = 50)  + geom_density( col = "Red")
g = g + stat_function(fun = function(x) exp(-x/2)/2)
print(g)
}
rmyexp(100000, nclass = 40)
```


Uniform W
Simulation
Simulation
Truth
Simulation using
w uniform method n=1e+05
1/2exp(-x/2)
Density
0.0
0.1
0.2
0.3
0.4
0
5
10
15
20
y

Order Statistics

- Sample of size "n" what is the distribution of:

$$Y_{min}, Y_{max}?$$

Taken From Larsen and Marx

Theorem 3.10.1 *Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample of continuous random variables, each having pdf $f_Y(y)$ and cdf $F_Y(y)$. Then*

a. *The pdf of the largest order statistic is*

$$f_{Y_{\max}}(y) = f_{Y'_n}(y) = n[F_Y(y)]^{n-1} f_Y(y)$$

b. *The pdf of the smallest order statistic is*

$$f_{Y_{\min}}(y) = f_{Y'_1}(y) = n[1 - F_Y(y)]^{n-1} f_Y(y)$$

Proof Finding the pdfs of $Y_{\max}$ and $Y_{\min}$ is accomplished by using the now-familiar technique of differentiating a random variable's cdf. Consider, for example, the case of the largest order statistic, Y'_n:

$$\begin{aligned} F_{Y'_n}(y) &= F_{Y_{\max}}(y) = P(Y_{\max} \le y) \\ &= P(Y_1 \le y, Y_2 \le y, \cdots, Y_n \le y) \\ &= P(Y_1 \le y) \cdot P(Y_2 \le y) \cdots P(Y_n \le y) \qquad \text{(why?)} \\ &= [F_Y(y)]^n \end{aligned}$$

Therefore,

$$f_{Y'_n}(y) = d/dy[[F_Y(y)]^n] = n[F_Y(y)]^{n-1} f_Y(y)$$

Similarly, for the smallest order statistic ($i = 1$),

$$\begin{aligned} F_{Y'_1}(y) &= F_{Y_{\min}}(y) = P(Y_{\min} \le y) \\ &= 1 - P(Y_{\min} > y) = 1 - P(Y_1 > y) \cdot P(Y_2 > y) \cdots P(Y_n > y) \\ &= 1 - [1 - F_Y(y)]^n \end{aligned}$$

Therefore,

$$f_{Y'_1}(y) = d/dy[1 - [1 - F_Y(y)]^n] = n[1 - F_Y(y)]^{n-1} f_Y(y)$$

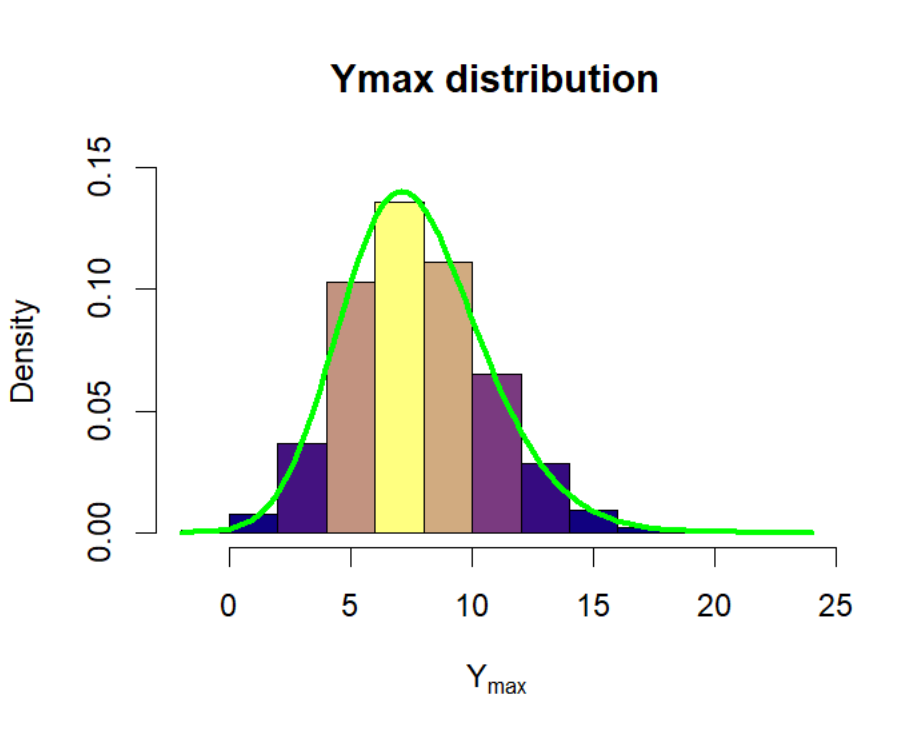
Make:
dpqr functions for order statistics

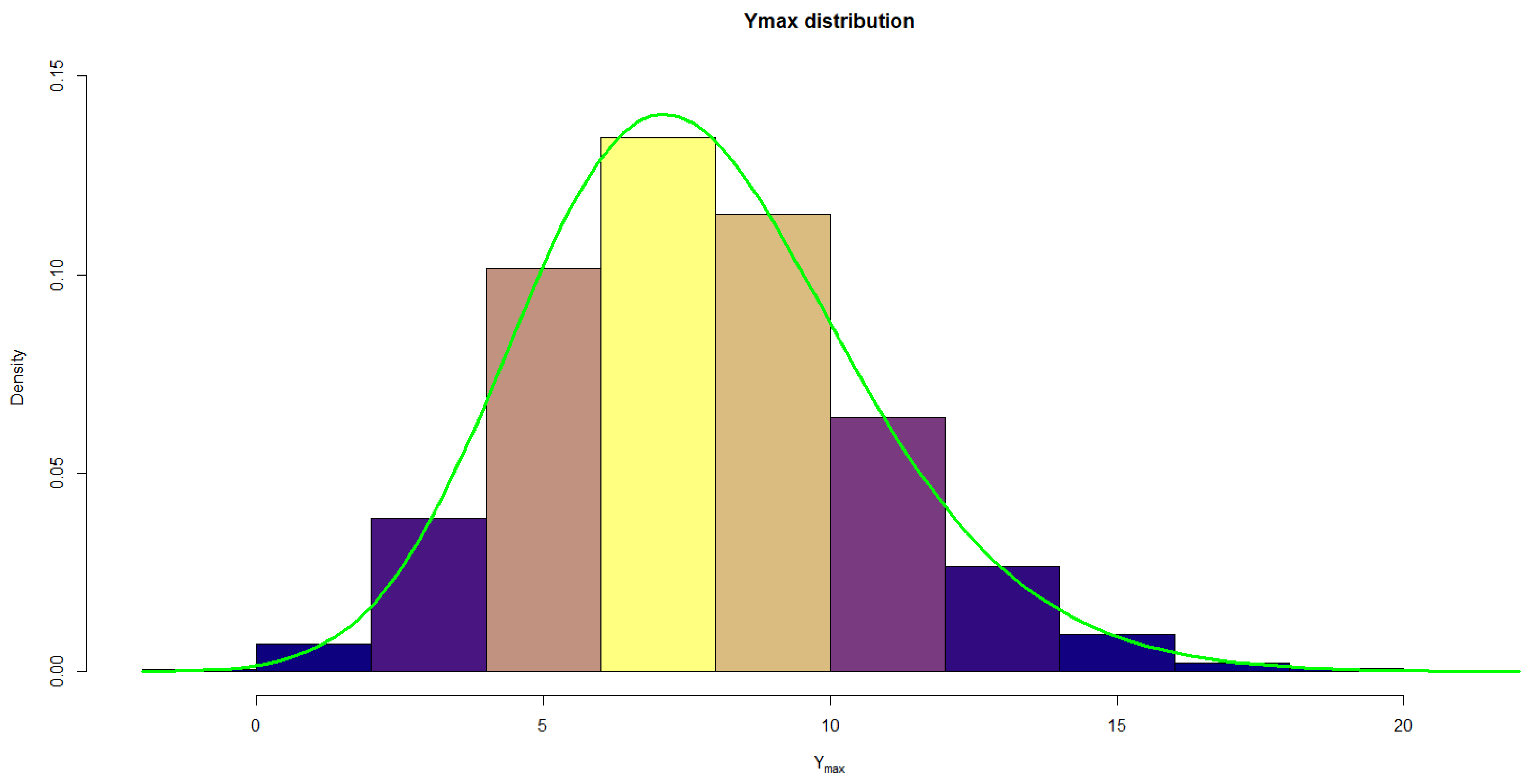
```r
#Distribution of Ymin,Ymax

dymax <- function(y,n, mu, sigma){
  n*(pnorm(y, mean = mu, sd = sigma))^(n-1)*dnorm(y,mean = mu, sd = sigma)
}

windows(); curve(dymax(x,10,0,5), from = -10, to = 20)

dymax(3:10,10,0,5)

pymax <- function(y,n,mu,sigma){
  pnorm(y,mean = mu, sd = sigma)^n
}

pymax(3:10,10,0,5)

qymax <- function(p,n,mu, sigma){
  qnorm(p^(1/n),mean = mu, sd = sigma)
}

qymax(0.5, 10,0,5)

rymax <- function(nsam, n, mu, sigma){
  qymax(runif(nsam), n, mu, sigma)
}
y <- rymax(10000,10,0,5)
h <- hist(y,plot = FALSE)
r <- h$density/max(h$density)
maxd <- max(h$density)
hist(y,freq = FALSE, col = rgb(r,r^2,0.5),
     ylim = c(0, 1.1*maxd), xlab = expression(Y[max]), main = "Ymax distribution")
curve(dymax(x,10,0,5), add = TRUE, lwd = 3, col = "green")
```

curve {graphics} R Documentation

Draw Function Plots

Description




Ymax distribution
Density
0.00
0.05
0.10
0.15
0
5
10
15
20
Y_{max}

W Find the density of $W = XY$ where $f_{X,Y}(x,y) = 2y$ on $[0,1]^2$

A colorful step-by-step derivation using the identity $F_W(w) = 1 - P(W > w)$.

1) Problem setup

We are given the joint density

$$f_{X,Y}(x,y) = 2y, \qquad 0 \le x \le 1,\ 0 \le y \le 1.$$

Let $W = XY$. We will compute the CDF via

$$F_W(w) = 1 - P(W > w).$$

2) Consider the range of w

If $w < 0$ then $F_W(w) = 0$. If $w > 1$ then $F_W(w) = 1$ because $0 \le XY \le 1$. So we only need to compute for $0 \le w \le 1$.

3) Write the probability $P(W > w)$ as a double integral

$$P(W > w) = P(XY > w) = \iint_{\{(x,y)\in[0,1]^2 : xy > w\}} 2y\,dx\,dy.$$

For a fixed y with \(0\le yw\). For $y \in [w,1]$ the condition is $x > w/y$. Thus the region is $y \in [w,1],\ x \in (w/y, 1]$.

4) Integrate in x first

$$\begin{aligned} P(W > w) &= \int_{y=w}^{1} \int_{x=w/y}^{1} 2y\,dx\,dy \\ &= \int_{y=w}^{1} \left[2y\left(1 - \tfrac{w}{y}\right)\right] dy \\ &= \int_{y=w}^{1} (2y - 2w)\,dy. \end{aligned}$$

4) Integrate in x first

$$P(W > w) = \int_{y=w}^{1} \int_{x=w/y}^{1} 2y \, dx \, dy$$
$$= \int_{y=w}^{1} \left[2y \left(1 - \tfrac{w}{y}\right)\right] dy$$
$$= \int_{y=w}^{1} (2y - 2w) \, dy.$$

5) Finish the integral

$$P(W > w) = \int_{w}^{1} (2y - 2w) \, dy = \left[y^2 - 2wy\right]_{y=w}^{1}$$
$$= (1 - 2w) - (w^2 - 2w^2) = 1 - 2w + w^2 = (1 - w)^2.$$

6) CDF and PDF

Therefore for $0 \le w \le 1$

$$F_W(w) = 1 - P(W > w) = 1 - (1 - w)^2 = 2w - w^2.$$

Differentiate to obtain the density

$$f_W(w) = F'_W(w) = 2 - 2w = 2(1 - w), \qquad 0 \le w \le 1.$$

Final PDF $f_W(w) = 2(1 - w),\ 0 \le w \le 1$

And $f_W(w) = 0$ for $w < 0$ or $w > 1$. Check: $\int_0^1 2(1 - w) \, dw = 1$.

Suppose that the following is true:

$$F_W(w) = 2w - w^2$$

$$f_W(w) = 2 - 2w = 2(1 - w)$$

Create dpqr functions!

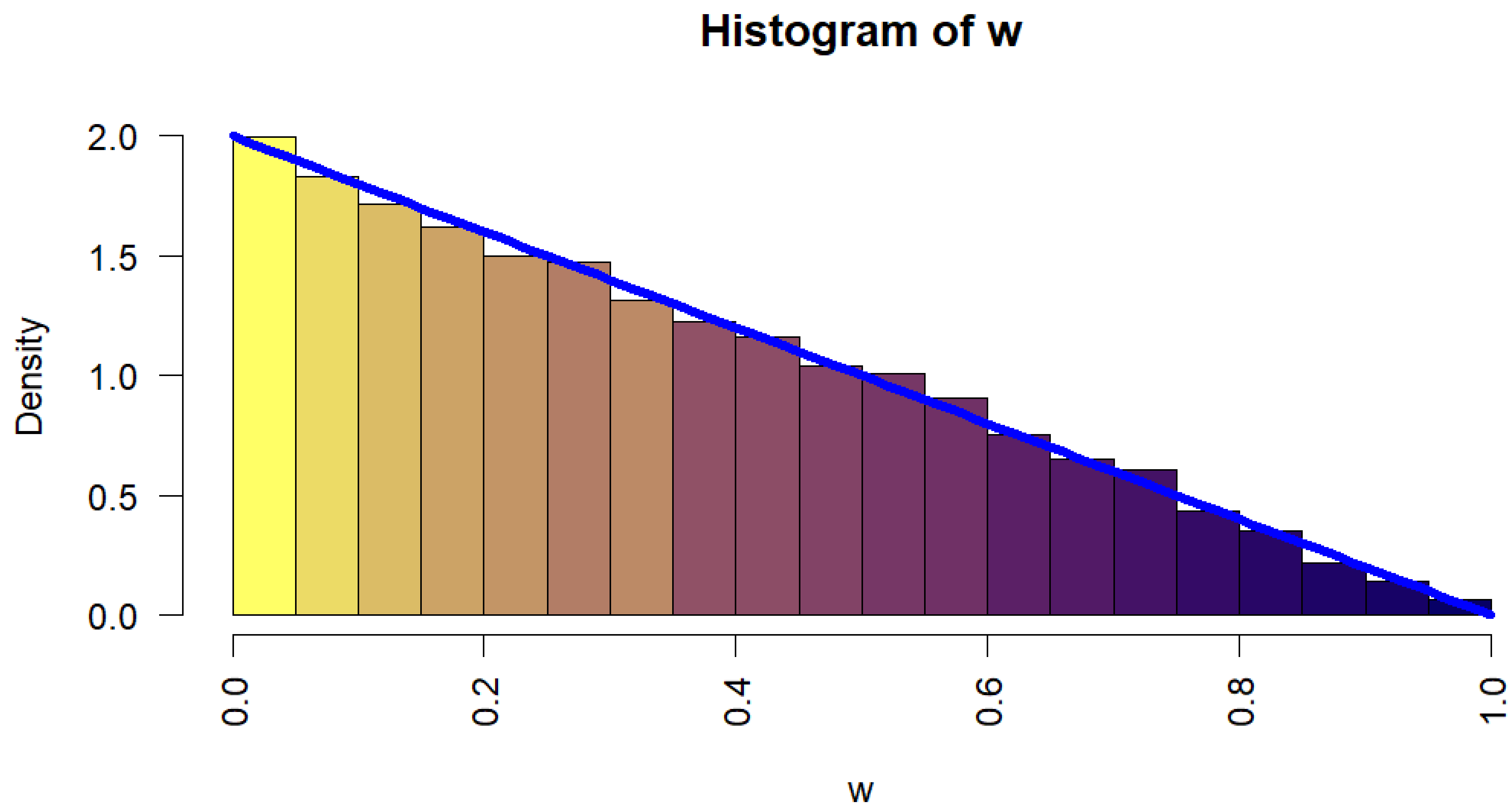
Follow the R code

```r
# dw

dw <- function(x){
  ifelse(x >= 0 & x <= 1, 2*(1-x), 0)
}

curve(dw, xlim = c(-1,2))

# pw

pw <- function(x){
  ifelse(x <0,0,
         ifelse(x >=0 & x <= 1,
                2*x - x^2, 1))
}

curve(pw, xlim = c(-1,2))
```

```r
# qw
# w^2 - 2w  + p = 0

qw <- function(p){
  ifelse(p >= 0 & p <= 1, 1 - sqrt(1-p),stop("p >= 0 and p <= 1"))
}

# OR
# This is done so that you can see how to make this work when an algebraic solution
is not obvious
# in our case we have w^2 - 2w  + p = 0
#Use quadratic eq to solve or find roots with uniroot

# experiment with map

purrr::map_vec(1:4, ~ .x^2 + 10) # replaces a for loop



p <- c(0.1,0.4)
map_vec(.x = p,.f = function(.x) {
  out <- uniroot(f = function(x) x^2 - 2*x + .x, interval = c(0,1), tol = 1e-15)
  out$root
})
```

```r
#Now make the function

qw2 <- function(p){
  ifelse(p >= 0 & p <= 1, {
    purrr::map_vec(.x = p,.f = function(.x) {
      out <- uniroot(f = function(x) x^2 - 2*x + .x,
                     interval = c(0,1), tol = 1e-15)
      return(out$root)
    })

  },stop("p >= 0 and p <= 1"))

}



# Does it work?
qw2(c(0.1,0.4))
qw(0.4)
qw(seq(0,1,by = 0.1)) - qw2(seq(0,1,by = 0.1)) #compare
```

```r
# rw
#This use the w-F theorem from book

rw <- function(n){
  ifelse(n >=1, {
    w <- runif(n)
    return(qw(w))
  }, stop("n >= 1"))
}

# Now lets test it!!

w <- rw(n = 10000)
h <- hist(w, plot = FALSE) # don't plot just obtain stats
rr <- h$density
rr <- rr/max(rr)# This will be a vector with components between 0 and 1
hist(r, freq = FALSE,
     col = rgb(rr,rr^2, 0.4),
     xlab = "w",
     main = "Histogram of w",
     las = 2)
curve(dw, add = TRUE,
      lwd = 4, col = "Blue")
```



Histogram of w
Density
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8
1.0
w

Definition 6.12

Let $Y_1, Y_2, \ldots, Y_n$ be random variables and let $a_1, a_2, \ldots, a_n$ be constants. Then ℓ is a **linear function** of $Y_1, Y_2, \ldots, Y_n$ if

$$\ell = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n$$

THEOREM 6.8

The Expected Value $E(\ell)$ and Variance $V(\ell)$* of a Linear Function of $Y_1, Y_2, \ldots, Y_n$

Suppose the means and variances of $Y_1, Y_2, \ldots, Y_n$ are $(\mu_1, \sigma_1^2), (\mu_2, \sigma_2^2), \ldots, (\mu_n, \sigma_n^2)$, respectively. If $\ell = a_1Y_1 + a_2Y_2 + \cdots + a_nY_n$, then

$$E(\ell) = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

and

$$\sigma_\ell^2 = V(\ell) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2$$
$$+ 2a_1a_2\text{Cov}(y_1, y_2) + 2a_1a_3\text{Cov}(y_1, y_3) + \cdots$$

Go through all examples

Example 6.15

Mean and Variance of a Function of Random Variables

Suppose Y_1, Y_2, and Y_3 are random variables with ($\mu_1 = 1, \sigma_1^2 = 2$), ($\mu_2 = 3, \sigma_2^2 = 1$), ($\mu_3 = 0, \sigma_3^2 = 4$), Cov($Y_1, Y_2$) = –1, Cov($Y_1, Y_3$) = 2, and Cov($Y_2, Y_3$) = 1. Find the mean and variance of

$$\ell = 2Y_1 + Y_2 - 3Y_3$$

Get these examples

Example 6.16

Expected Value of the Sample Mean

Let $Y_1, Y_2, \ldots, Y_n$ be a sample of n independent observations selected from a population with mean μ and variance σ^2. Find the expected value and variance of the sample mean, $\overline{Y}$.

Example 6.17

Probability Distribution of the Sample Mean

Suppose that the population of Example 6.16 has mean $\mu = 10$ and variance $\sigma^2 = 4$. Describe the probability distribution for a sample mean based on $n = 25$ observations.

Definition 6.13

The **sampling distribution** of a statistic is its probability distribution.

Definition 6.14

The **standard error** of a statistic is the standard deviation of its sampling distribution.

In R the standard error is the estimate of the standard deviation of the sampling statistic

THEOREM 6.9 The Central Limit Theorem

If a random sample of n observations, $Y_1, Y_2, \ldots, Y_n$, is drawn from a population with finite mean μ and variance σ^2, then, when n is sufficiently large, the sampling distribution of the sample mean $\overline{Y}$ can be approximated by a normal density function.

Definition 6.15

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of n observations from a population with finite mean μ and finite standard deviation σ. Then, the **mean and standard deviation of the sampling distribution** of $\overline{Y}$, denoted $\mu_{\overline{y}}$ and $\sigma_{\overline{y}}$, respectively, are

$$\mu_{\overline{y}} = \mu, \quad \sigma_{\overline{y}} = \sigma/\sqrt{n}$$

THEOREM 6.10

Let $a_1, a_2, \ldots, a_n$ be constants and let $Y_1, Y_2, \ldots, Y_n$ be n normally distributed random variables with $E(Y_i) = \mu_i$, $V(Y_i) = \sigma_i^2$, and $\text{Cov}(Y_i, Y_j) = \sigma_{ij}\,(i = 1, 2, \ldots, n)$. Then the sampling distribution of a linear combination of the normal random variables

$$\ell = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n$$

possesses a normal density function with mean and variance*

$$E(\ell) = \mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

and

$$\begin{aligned} V(\ell) = {} & a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2 \\ & + 2a_1a_2\sigma_{12} + 2a_1a_3\sigma_{13} + \cdots + 2a_1a_n\sigma_{1n} \\ & + 2a_2a_3\sigma_{23} + \cdots + 2a_2a_n\sigma_{2n} \\ & + \cdots + 2a_{n-1}a_n\sigma_{n-1,\,n} \end{aligned}$$

Example 6.20

Sampling Distribution of $(\overline{Y}_1 - \overline{Y}_2)$

Suppose you select independent random samples from two normal populations, n_1 observations from population 1 and n_2 observations from population 2. If the means and variances for populations 1 and 2 are (μ_1, σ_1^2) and (μ_2, σ_2^2), respectively, and if $\overline{Y}_1$ and $\overline{Y}_2$ are the corresponding sample means, find the distribution of the difference $(\overline{Y}_1 - \overline{Y}_2)$.

Solution Since $\bar{y}_1$ and $\bar{y}_2$ are both linear functions of normally distributed random variables, they will be normally distributed by Theorem 6.10. The means and variances of the sample means (see Example 6.16) are

$$E(\overline{Y}_i) = \mu_i \quad \text{and} \quad V(\overline{Y}_i) = \frac{\sigma_i^2}{n_i} \quad (i = 1, 2)$$

Then, $\ell = \overline{Y}_1 - \overline{Y}_2$ is a linear function of two normally distributed random variables, $\bar{y}_1$ and $\bar{y}_2$. According to Theorem 6.10, ℓ will be normally distributed with

$$E(\ell) = \mu_\ell = E(\overline{Y}_1) - E(\overline{Y}_2) = \mu_1 - \mu_2$$

$$V(\ell) = \sigma_\ell^2 = (1)^2 V(\overline{Y}_1) + (-1)^2 V(\overline{Y}_2) + 2(1)(-1)\text{Cov}(\overline{Y}_1, \overline{Y}_2)$$

But, since the samples were independently selected, $\overline{Y}_1$ and $\overline{Y}_2$ are independent and $\text{Cov}(\overline{Y}_1, \overline{Y}_2) = 0$. Therefore,

$$V(\ell) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

We have shown that $(\overline{Y}_1 - \overline{Y}_2)$ is a normally distributed random variable with mean $(\mu_1 - \mu_2)$ and variance $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$.

The Sampling Distribution of a Sum of Random Variables

If a random sample of n observations, $Y_1, Y_2, \ldots, Y_n$, is drawn from a population with finite mean μ and variance σ^2, then, when n is sufficiently large, the sampling distribution of the sum

$$\sum_{i=1}^{n} Y_i$$

can be approximated by a normal density function with mean $E(\Sigma Y_i) = n\mu$ and $V(\Sigma Y_i) = n\sigma^2$.

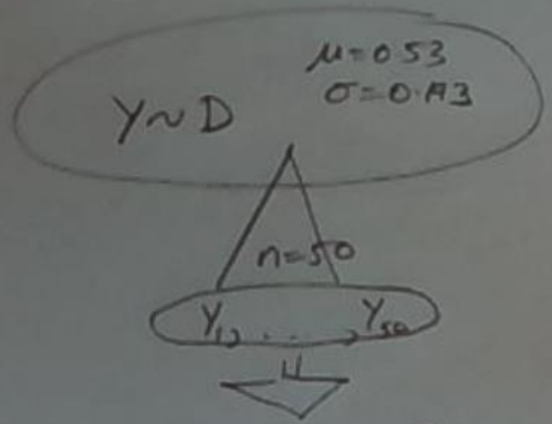
6.76 *Levelness of concrete slabs.* Geotechnical engineers use water-level "manometer" surveys to assess the levelness of newly constructed concrete slabs. Elevations are typically measured at eight points on the slab; of interest is the maximum differential between elevations. The *Journal of Performance of Constructed Facilities* (Feb. 2005) published an article on the levelness of slabs in California residential developments. Elevation data collected for over 1,300 concrete slabs *before tensioning* revealed that maximum differential, Y, has a mean of $\mu = .53$ inch and a standard deviation of $\sigma = .193$ inch. Consider a sample of $n = 50$ slabs selected from those surveyed and let $\overline{Y}$ represent the mean of the sample.

a. Fully describe the sampling distribution of $\overline{Y}$.
b. Find $P(\overline{Y} > .58)$.
c. The study also revealed that the mean maximum differential of concrete slabs measured *after tensioning and loading* is $\mu = .58$ inch. Suppose the sample data yields $\overline{Y} = .59$ inch. Comment on whether the sample measurements were obtained before tensioning or after tensioning and loading.

6.76

Before tensioning $\mu = 0.53$ $\sigma = 0.193$

After tensioning $\mu = 0.58$



$\bar{Y} \underset{n\to\infty}{\overset{CLT}{\sim}} N\left(\mu = 0.53, \sigma^2 = \frac{0.193^2}{50}\right)$

a) $\bar{Y} \sim N$

b) $P(\bar{Y} > 0.58) = 1 - \text{pnorm}\left(0.58, 0.53, \frac{0.193}{\sqrt{50}}\right)$
$= 0.03348444$

c) Say $\bar{Y} = 0.59$ is this a measurement Before or After tensioning?

$P(\bar{Y} > 0.59)$ Before $= 1 - \text{pnorm}\left(0.59, 0.53, \frac{0.193}{\sqrt{50}}\right) = 0.01396531$

After $= 1 - \text{pnorm}\left(0.59, 0.58, \frac{0.193}{\sqrt{50}}\right) = 0.3570421$

More likely the sample measurements came after tensioning

Normal approximation to the Binomial

$$Y = \sum_{i=1}^{n} Y_i \qquad \text{where } Y_i = \begin{cases} 1 & \text{if success} \\ 0 & \text{if failure} \end{cases}$$

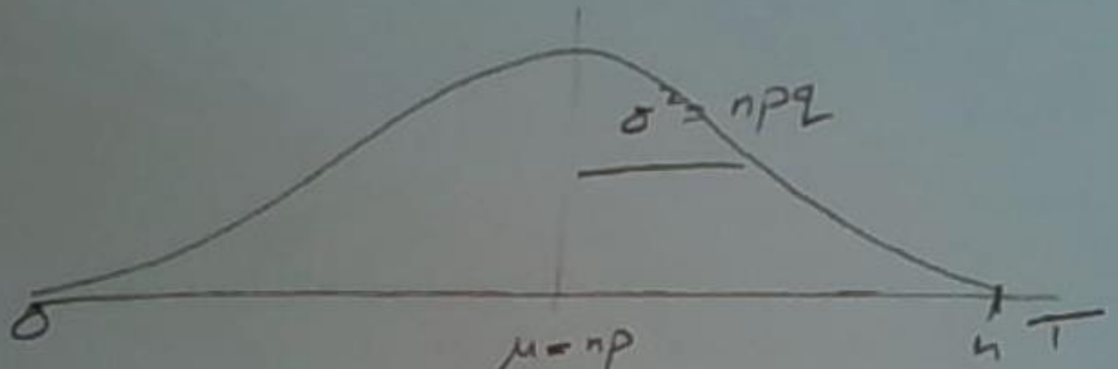
Condition to be satisfied

Condition Required to Apply a Normal Approximation to a Binomial Probability Distribution

The approximation will be good if both $\mu - 2\sigma = np - 2\sqrt{npq}$ and $\mu + 2\sigma = np + 2\sqrt{npq}$ lie between 0 and n. This condition will be satisfied if both $np \geq 4$ and $nq \geq 4$.

Why take $np \geq 4$, $nq \geq 4$

use empirical rule



95% of dist with $\pm 2\sigma$ of μ.

so

$$0 = \mu - 2\sigma$$
$$= np - 2\sqrt{npq}$$
$$2\sqrt{npq} = np$$
$$4npq = n^2p^2$$
$$4q = np \quad \text{L.H.S} \leq 4$$
$\therefore$ take $np \geq 4$

$$n = \mu + 2\sigma$$
$$n = np + 2\sqrt{npq}$$
$$n - np = 2\sqrt{npq}$$
$$n(1-p) = 2\sqrt{npq}$$
$$nq = 2\sqrt{npq}$$
$$n^2q^2 = 4npq$$
$$nq = 4p \quad \leq 4$$
take $nq \geq 4$

Continuity correction

Continuity Correction for the Normal Approximation to a Binomial Probability

Let Y be a binomial random variable with parameters n and p, and let Z be a standard random variable. Then,

$$P(Y \leq a) \approx P\left(Z < \frac{(a + .5) - np}{\sqrt{npq}}\right)$$

$$P(Y \geq a) \approx P\left(Z > \frac{(a - .5) - np}{\sqrt{npq}}\right)$$

$$P(a \leq Y \leq b) \approx P\left(\frac{(a - .5) - np}{\sqrt{npq}} < Z < \frac{(b + .5) - np}{\sqrt{npq}}\right)$$

This will be difficult to remember – please use first principles to perform continuity corrections

Sampling distributions – related to Normal

THEOREM 6.11

If a random sample of n observations, $Y_1, Y_2, \ldots, Y_n$, is selected from a normal distribution with mean μ and variance σ^2, then the sampling distribution of

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

has a **chi-square density function** (see Section 5.7) with $\nu = (n-1)$ degrees of freedom.

Note: The random variable s^2 represents the sample variance.

THEOREM 6.12

If χ_1^2 and χ_2^2 are independent chi-square random variables with ν_1 and ν_2 degrees of freedom, respectively, then the sum $(\chi_1^2 + \chi_2^2)$ has a **chi-square distribution** with $(\nu_1 + \nu_2)$ degrees of freedom.

Theorem 6.12

$\chi^2 \sim \text{chisq}(\nu)$ ν = degrees of freedom.

$M_{\chi^2}(t) = E(e^{\chi^2 t}) = (1-2t)^{-\nu/2}$ see ch 5

Suppose X & Y are indept r.vs.

$$E(e^{(X+Y)t}) = E(e^{Xt+Yt})$$
$$= E(e^{Xt} \cdot e^{Yt})$$
$$= E(e^{Xt})\, E(e^{Yt}) \quad \text{if } X \,\&\, Y \text{ are indept.}$$

Suppose $X \sim \text{chisq}(\nu_1)$ & $Y \sim \text{chisq}(\nu_2)$

Then

$$M_{X+Y}(t) = M_X(t)\, M_Y(t)$$
$$= (1-2t)^{-\nu_1/2} (1-2t)^{-\nu_2/2}$$
$$= (1-2t)^{-(\nu_1+\nu_2)/2}$$

= MGF of a χ^2 with $\nu_1+\nu_2$ degrees of freedom.

Problem

Find $f_Z(z)$

Suppose X and Y are independent random variables.

If $Z = X + Y$ and $f(x) = e^{-x},\ x > 0,\ f(y) = e^{-y},\ y > 0$

Proof a

Using MGF's prove that $Z \sim Gamma(\alpha = 2, \beta = 1)$

Solution

$M_X(t) = (1-t)^{-1}$ and $M_Y(t) = (1-t)^{-1}$ since both X and Y are exponential densities.

Therefore, $M_Z(t) = M_X(t)M_Y(t) = (1-t)^{-2}$ this is the MGF of a gamma $(1-\beta t)^{-\alpha}$

$\therefore Z \sim Gamma(\alpha = 2, \beta = 1)$

Proof b

Using Bivariate distributional theory and techniques (set up integral) prove that the density of Z is a $Gamma(\alpha = 2, \beta = 1)$

Solution

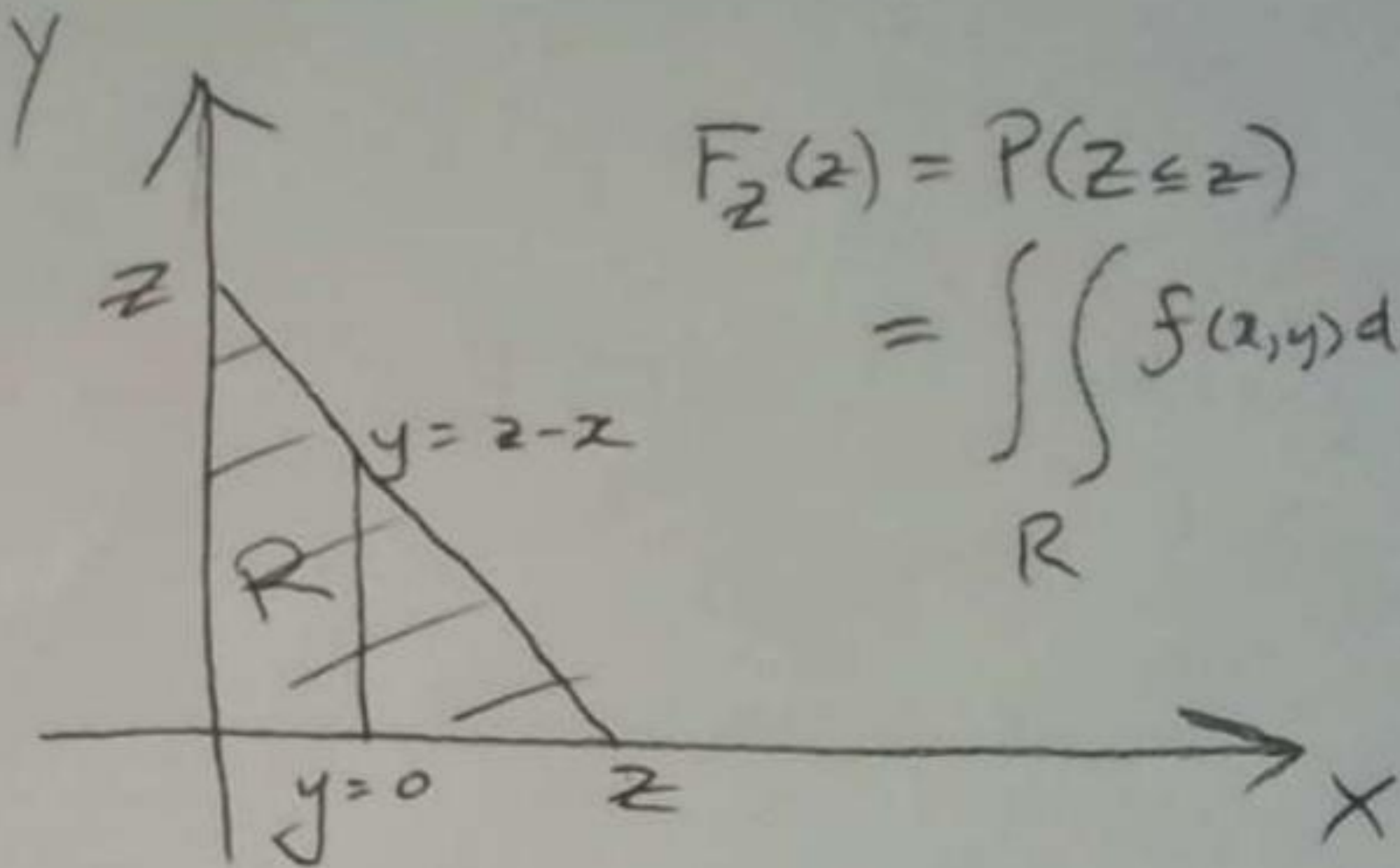
What we need to do is find the cumulative distributional function $F_Z(z)$ then differentiate to find the density.

$$f_Z(z) = \frac{dF_Z(z)}{dz}$$

Hopefully this will be the Gamma we desire!

We will start with $F_Z(z)$

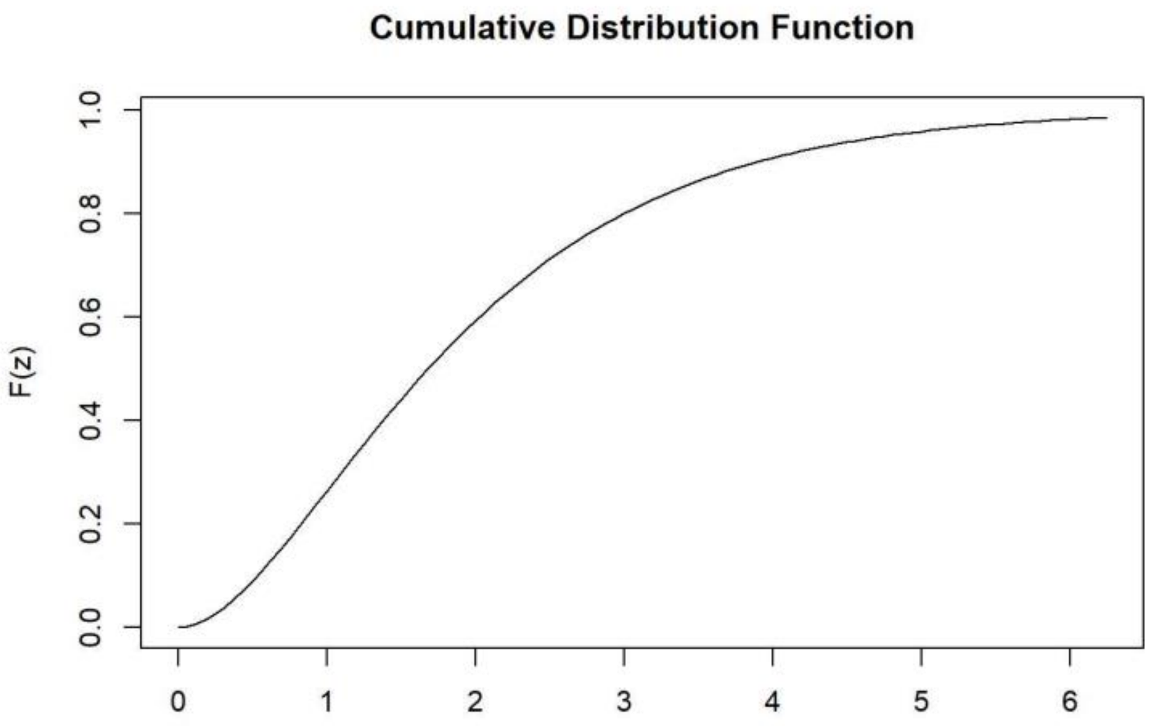
$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = P(Y \leq z - X)$$



$$F_Z(z) = P(Z \le z) = \iint_R f(x,y)\,dx\,dy$$

$$
\begin{aligned}
F_Z(z) &= \int_{x=0}^{x=z} \int_{y=0}^{z-x} e^{-x} e^{-y} dy dx \\
&= \int_{x=0}^{x=z} e^{-x} \int_{y=0}^{z-x} e^{-y} dy dx \\
&= \int_{x=0}^{x=z} e^{-x} \left[-e^{-y} \right]_0^{z-x} dx \\
&= \int_{x=0}^{x=z} e^{-x} (1 - e^{x-z}) dx \\
&= \int_{x=0}^{x=z} e^{-x} - e^{-z} dx \\
&= \left[-e^{-x} - x e^{-z} \right]_0^z \\
&= -e^{-z} - z e^{-z} + 1
\end{aligned}
$$

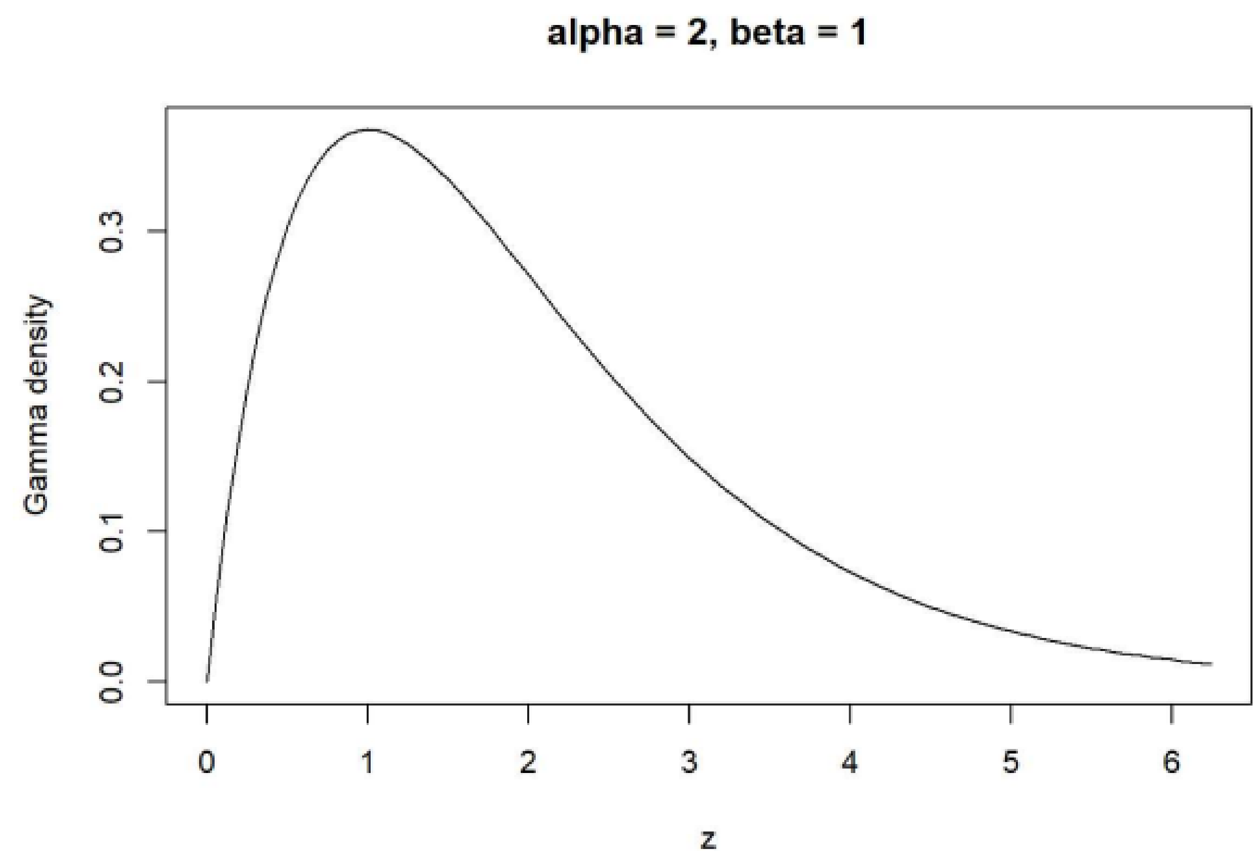
```
pg <- function(x){
  1-exp(-x) - x*exp(-x)
}

curve(expr = pg,
      xlim = c(0, 2 +3*sqrt(2)),
      main = "Cumulative Distribution Function",
      xlab = "z",
      ylab = "F(z)")
```



$$
\begin{aligned}
f_Z(z) &= \frac{dF_Z(z)}{dz} \\
&= \frac{d(-e^{-z} - ze^{-z} + 1)}{dz} \\
&= e^{-z} - e^{-z} + ze^{-z} \\
&= ze^{-z}
\end{aligned}
$$

$$
f(y) = \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)},\ Gamma\ Density\ \mu = \alpha\beta,\ \sigma^2 = \alpha\beta^2
$$

```
dg <- function(x) {
  x*exp(-x)
}

curve(expr = dg,
      xlim = c(0, 2 +3*sqrt(2)), # 0,mu + 3 sigma
      main = "alpha = 2, beta = 1",
      ylab = "Gamma density",
      xlab = "z")
```



So $Z \sim Gamma(\alpha = 2, \beta = 1)$. Note that:

$$f(z) = \frac{z^{2-1}e^{-z/1}}{1^2\Gamma(2)} = ze^{-z}$$

Definition of T random variable

Definition 6.16

Let Z be a standard normal random variable and χ^2 be a chi-square random variable with ν degrees of freedom. If Z and χ^2 are independent, then

$$T = \frac{Z}{\sqrt{\chi^2/\nu}}$$

is said to possess a **Student's *T* distribution** (or, simply, ***T* distribution**) with ν degrees of freedom.

Definition 6.17

Let χ_1^2 and χ_2^2 be chi-square random variables with ν_1 and ν_2 degrees of freedom, respectively. If χ_1^2 and χ_2^2 are independent, then

$$F = \frac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2}$$

is said to have an ***F* distribution** with ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom.

Go through all examples

Example 6.25

Derivation of Student's T-distribution

Suppose the random variables $\overline{Y}$ and S^2 are the mean and variance of a random sample of n observations from a normally distributed population with mean μ and variance σ^2. It can be shown (proof omitted) that $\overline{Y}$ and S^2 are statistically independent when the sampled population has a normal distribution. Use this result to show that

$$T = \frac{\overline{Y} - \mu}{S/\sqrt{n}}$$

possesses a T distribution with $\nu = (n - 1)$ degrees of freedom.*

TABLE 6.3a Sampling Distributions of Statistics Based on Independent Random Samples of n_1 and n_2 Observations, Respectively, from Normally Distributed Populations with Parameters (μ_1, σ_1^2) and (μ_2, σ_2^2)

Statistic	Sampling Distribution	Additional Assumptions	Basis of Derivation of Sampling Distribution
$\chi^2 = \frac{(n_1 + n_2 - 2)S_p^2}{\sigma^2}$ where $S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	Chi-square with $\nu = (n_1 + n_2 - 2)$ degrees of freedom	$\sigma_1^2 = \sigma_2^2 = \sigma^2$	Theorems 6.11–6.12
$T = \frac{(\overline{Y}_1 - \overline{Y}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ where $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	Student's T with $\nu = (n_1 + n_2 - 2)$ degrees of freedom	$\sigma_1^2 = \sigma_2^2 = \sigma^2$	Theorems 6.10–6.11 and Definition 6.15
$F = \left(\frac{S_1^2}{S_2^2}\right)\left(\frac{\sigma_2^2}{\sigma_1^2}\right)$	F distribution with $\nu_1 = (n_1 - 1)$ numerator degrees of freedom and $\nu_2 = (n_2 - 1)$ denominator degrees of freedom	None	Theorem 6.11 and Definition 6.17

TABLE 6.3b Sampling Distributions of Statistics Based on a Random Sample from a Single Normally Distributed Population with Mean μ and Variance σ^2

Statistic	Sampling Distribution	Additional Assumptions	Basis of Derivation of Sampling Distribution
$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$	Chi-square with $\nu = (n-1)$ degrees of freedom	None	Methods of Section 6.7
$t = \frac{\bar{y} - \mu}{S/\sqrt{n}}$	Student's T with $\nu = (n-1)$ degrees of freedom	None	Theorems 6.10–6.11 and Definition 6.15

Quick Review

Key Terms

[*Note: Items marked with an asterisk (*) are from the optional section in this chapter.*]

Bivariate density function 241

Bivariate probability distribution 236

Central limit theorem 283

Chi-square distribution 274

Conditional density function 242

Conditional probability distribution 237, 238

Continuity correction 272

Correlation 283

Covariance 251

*Cumulative distribution function method 253

Expected values 253

F distribution 274

Independent 247

Joint probability distribution 236

Linear function 257

Linear relationships 250

Marginal density function 242

Marginal probability distribution 236, 237

Monte Carlo simulation 262

Multivariate probability distribution 239

Sampling distribution 261

Sampling distribution of the mean 261

Sampling distribution of a sum 268

Standard error 261

T distribution 274

Key Formulas

Conditional probability distribution for discrete random variable:	$p(x \mid y) = p(x, y)/p(y)$ if X and Y are dependent $= p(x)$ if X and Y are independent	238 238
Conditional density function for continuous random variable:	$f(x \mid y) = f(x, y)/f(y)$ if X and Y are dependent $= f(x)$ if X and Y are independent	242 242
Expected values:	$E(c) = c$ $E[c \cdot g(X, Y)] = c \cdot E[g(X, Y)]$ $E[g_1(X, Y) + g_2(X, Y)] = E[g_1(X, Y)] + E[g_2(X, Y)]$ $E(XY) = E(X) \cdot E(Y)$ if X and Y are independent	246
Covariance:	$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$ if X and Y dependent $= 0$ if X and Y independent	251 251
Correlation:	$\rho = \dfrac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$ if X and Y dependent $= 0$ if X and Y independent	283
Normal approximation to binomial:	$P(a < Y < b) = P\left\{\dfrac{(a - .5) - np}{\sqrt{npq}} < Z < \dfrac{(b + .5) - np}{\sqrt{npq}}\right\}$	272
Sampling distribution of $\overline{Y}$:	Mean $= \mu$ Standard deviation $= \sigma/\sqrt{n}$	263
Sampling distribution of ΣY:	Mean $= n\mu$ Standard deviation $= \sqrt{n}\,\sigma$	268

LANGUAGE LAB

Symbol	Pronunciation	Description
$p(x \mid y)$	p of x given y	Conditional probability distribution for X given Y
$f(x \mid y)$	f of x given y	Conditional density function for X given Y
$\text{Cov}(X, Y)$	Covariance	Covariance of X and Y
ρ	rho	Correlation coefficient for X and Y
$\mu_{\overline{y}}$	mu of $\overline{Y}$	Mean of sampling distribution of $\overline{Y}$
$\sigma_{\overline{y}}$	sigma of $\overline{Y}$	Standard deviation of sampling distribution of $\overline{Y}$

Chapter Summary

- The **joint probability distribution** for two random variables is called a **bivariate distribution**.
- The **conditional probability distribution** for a random variable X, given Y, is the joint probability distribution for X and Y divided by the marginal probability distribution for Y.
- The **covariance** of X and Y: $\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$.
- The **correlation**: of X and Y: $\rho = \text{Cov}(X, Y)/(\sigma_x \sigma_y)$
- For two **independent** random variables, (1) the joint probability distribution is the product of the two respective marginal probability distributions, (2) $E(XY) = E(X) \cdot E(y)$, (3) covariance equals 0, and (4) correlation equals 0.
- The **sampling distribution** of a statistic is the theoretical probability distribution of the statistic in repeated sampling.
- The **standard error** of a statistic is the standard deviation of the sampling distribution.
- **Monte Carlo simulation** involves repeatedly generating observations on a statistic in order to approximate the sampling distribution.
- The **central limit theorem** states that the sampling distribution of $\overline{Y}$ is approximately normal for large n.
- Two properties of the sampling distribution of $\overline{Y}$: mean $= \mu$, standard deviation $= \sigma/\sqrt{n}$
- Normal distribution can be used to approximate a binomial probability when $\mu \pm 2\sigma$ falls within the interval $(0, n)$. This will be true when $np \geq 4$ **and** $nq \geq 4$.
- Some sampling distributions related to the normal distribution: **chi-square distribution, Student's *T* distribution**, and ***F* distribution**.


WOW!! IT'S ROCK'N'ROLL!!!