# Probability and Statistics (ENM 503)

Michael A. Carchidi

December 1, 2015

**Chapter 7 - Jointly Distributed Random Variables**

The following notes are based on the textbook entitled: *A First Course in Probability* by Sheldon Ross (9th edition) and these notes can be viewed at

https://canvas.upenn.edu/

after you log in using your PennKey user name and Password.

## 1. Joint Distribution Functions

We are often interested in probability statements concerning two or more random variables. In order to do deal with such probabilities, we define, for any two random variables $X$ and $Y$, the joint cumulative probability distribution function of $X$ and $Y$ by

$$F(a,b) = P((X \le a) \cap (Y \le b)) \tag{1}$$

for $-\infty < a, b < +\infty$. The cdf of just the random variable $X$ can then be obtained from Equation (1) by taking the limit of Equation (1) as $b \to +\infty$ since

$$\lim_{b \to +\infty} ((X \le a) \cap (Y \le b)) = (X \le a) \cap \lim_{b \to +\infty} (Y \le b) = (X \le a) \cap (Y \le +\infty)$$

which reduces to

$$\lim_{b \to +\infty} ((X \le a) \cap (Y \le b)) = X \le a.$$

Then, since probability is a continuous set function, we have

$$\lim_{b \to \infty} F(a,b) = \lim_{b \to \infty} P((X \le a) \cap (Y \le b)) = P(\lim_{b \to \infty} ((X \le a) \cap (Y \le b)))$$

or

$$\lim_{b\to\infty} F(a,b) = P(X \le a) = F_X(a). \tag{2a}$$

In a similar way, we have

$$\lim_{a\to\infty} F(a,b) = P(Y \le b) = F_Y(b). \tag{2b}$$

The separate joint distribution functions in Equations (2a,b) are sometimes referred to as the *marginal distributions* of $X$ and $Y$.

Note that

$$F((X > a) \cap (Y > b)) = 1 - F\left(\overline{(X > a) \cap (Y > b)}\right)$$

where

$$\overline{(X > a) \cap (Y > b)} = \overline{(X > a)} \cup \overline{(Y > b)} = (X \le a) \cup (Y \le b)$$

and so

$$F((X > a) \cap (Y > b)) = 1 - F((X \le a) \cup (Y \le b)).$$

But

$$F((X \le a) \cup (Y \le b)) = F(X \le a) + F(Y \le b) - F((X \le a) \cap (Y \le b))$$

which becomes

$$F((X \le a) \cup (Y \le b)) = F_X(a) + F_Y(b) - F(a,b).$$

Thus we find that

$$F((X > a) \cap (Y > b)) = 1 - F_X(a) - F_Y(b) + F(a,b). \tag{3}$$

The student should show that, in general

$$\begin{aligned} P((a_1 < X \le a_2) \cap (b_1 < Y \le b_2)) = \;& F(a_2,b_2) + F(a_1,b_1) \\ & -F(a_1,b_2) - F(a_2,b_1) \end{aligned} \tag{4}$$

whenever $a_1 < a_2$ and $b_1 < b_2$.

*When $X$ and $Y$ are Both Discrete Random Variables*

Although the earlier equations are valid for both discrete and continuous random variables, in the case when $X$ and $Y$ are both discrete random variables, it is more convenient to define the joint *probability mass function* of $X$ and $Y$ as

$$p(x,y) = P((X = x) \cap (Y = y)) \tag{5}$$

for all $x \in R_X$ and $y \in R_Y$. The probability mass functions of $X$ and $Y$ separately are then obtained from

$$p_X(x) = \sum_{y \in R_Y} p(x,y) \quad \text{and} \quad p_Y(y) = \sum_{x \in R_X} p(x,y), \tag{6}$$

respectively.

*Example #1: The Roll of Two Dice*

Suppose that a 5-sided die and a 6-sided die are rolled with $X$ being the outcome of the 5-sided die and $Y$ being the outcome of the 6-sided die. Then the values of $p(x,y)$ are

$$p(x,y) = P((X = x) \cap (Y = y)) = \frac{1}{5} \times \frac{1}{6} = \frac{1}{30}$$

for all $x = 1, 2, 3, 4, 5$ and $y = 1, 2, 3, 4, 5, 6$. ■

*Example #2: Balls and Urns*

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white and 5 blue balls. If we let $X$ and $Y$ denote, respectively, the number of red and white balls drawn, then $R_X = \{0, 1, 2, 3\}$ and $R_Y = \{0, 1, 2, 3\}$ with

$$p(x,y) = \frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}}{\binom{5+4+3}{3}} = \frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}}{\binom{12}{3}}$$

or

$$p(x,y) = \frac{1}{220}\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}.$$

The results of these can be summarized in the following table

| The Values of $220 \times p(x,y)$ | | | | | |
|---|---|---|---|---|---|
| $x \backslash y$ | 0 | 1 | 2 | 3 | $p_X(x) =$ Row Sums |
| 0 | 10 | 40 | 30 | 4 | 84 |
| 1 | 30 | 60 | 18 | 0 | 108 |
| 2 | 15 | 12 | 0 | 0 | 27 |
| 3 | 1 | 0 | 0 | 0 | 1 |
| $p_y(y) =$ Column Sums | 56 | 112 | 48 | 4 | Sums Equal 220 |

Note how $p_X(x)$and $p_y(y)$ are computed as marginal distributions by taking row and column sums, respectively. ■

*Example #3: Number of Boys and Girls*

Suppose that 15% of the families in a certain community have no children, 20% have 1 child, 35% have two children and 30% have three children. Suppose further that in each family, each child is equally likely to be a boy or a girl. If a family is chosen at random from this community, then $B$ (the number of boys) and $G$ (the number of girls) in this family will have the joint probability mass function, $p(b,g)$, given by

$$p(0,0) = P((X=0) \cap (Y=0)|0)P(0) = (1)(0.15) = 0.15$$

and

$$p(0,1) = P((X=0) \cap (Y=1)|1)P(1) = (0.5)(0.2) = 0.10 = p(1,0)$$

and

$$p(0,2) = P((X=0) \cap (Y=2)|2)P(2) = (0.5)^2(0.35) = 0.0875 = p(2,0)$$

and

$$p(0,3) = P((X=0) \cap (Y=3)|3)P(3) = (0.5)^3(0.3) = 0.0375 = p(3,0)$$

and

$$p(1,1) = P((X=1) \cap (Y=1)|2)P(2) = \binom{2}{1}(0.5)^1(0.5)^{2-1}(0.35) = 0.175$$

and

$$p(1,2) = P((X=1)\cap(Y=2)|3)P(3) = \binom{3}{1}(0.5)^1(0.5)^{2-1}(0.3) = 0.1125 = p(2,1).$$

The results of these can be summarized in the following table

The Values of $p(x,y)$

| $x\backslash y$ | 0 | 1 | 2 | 3 | $p_X(x) =$ Row Sums |
|---|---|---|---|---|---|
| 0 | 0.1500 | 0.1000 | 0.0875 | 0.0375 | 0.3750 |
| 1 | 0.1000 | 0.175 | 0.1125 | 0 | 0.3875 |
| 2 | 0.0875 | 0.1125 | 0 | 0 | 0.2000 |
| 3 | 0.0375 | 0 | 0 | 0 | 0.0375 |
| $p_y(y) =$ Column Sums | 0.3750 | 0.3875 | 0.2000 | 0.0375 | Sums Equal 1 |

Note how $p_X(x)$and $p_y(y)$ are computed as marginal distributions by taking row and column sums, respectively. ■

*Jointly Continuous Random Variables*

We say that $X$ and $Y$ are jointly continuous if there exists a function $f(x,y)$, defined for all real $x$ and $y$, having the property that for every set $C = R \times R$ of pairs of real numbers,

$$P((X,Y) \in C) = \iint_{(x,y)\in C} f(x,y)dxdy. \tag{7}$$

The function $f(x,y)$ is called the joint probability density function of $X$ and $Y$. If $A$ and $B$ are any sets of real numbers, then by defining

$$C = \{(x,y)|x \in A, y \in B\},$$

we see that

$$P((X \in A) \cap (Y \in B)) = \int_B \int_A f(x,y)dxdy \tag{8a}$$

and hence

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy \qquad \text{and} \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx. \tag{8b}$$

Since

$$F(a,b) = P((X \le a) \cap (Y \le b)) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)dydx$$

we see that

$$f(x,y) = \frac{\partial^2 F(a,b)}{\partial a \partial b} \tag{9}$$

and since

$$P((a_1 < X \le a_2) \cap (b_1 < Y \le b_1)) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f(x,y)dydx \tag{10a}$$

we see that

$$P((a < X \le a + \Delta a) \cap (b < Y \le b + \Delta)) = \int_{a}^{a+\Delta} \int_{b}^{b+\Delta b} f(x,y)dydx$$

which reduces to

$$P((a < X \le a + \Delta a) \cap (b < Y \le b + \Delta b)) = f(a,b)\Delta a \Delta b \tag{10b}$$

showing that $f(a,b)$ is the probability per unit area that $(X,Y)$ is near $(a,b)$.

*Example #4: Joint Probability Density Functions*

Suppose that the joint probability density function for the random variables $X$ and $Y$ is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & \text{for} \quad 0 < x, y \\ 0, & \text{for} \quad \text{other values of } x \text{ and } y \end{cases}.$$

Then

$$P((X > 1) \cap (Y < 1)) = \int_{1}^{\infty} \int_{0}^{1} 2e^{-x}e^{-2y}dydx = e^{-1} - e^{-3} \simeq 0.3181.$$

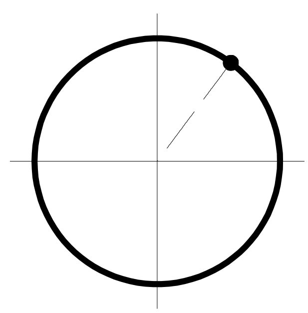
We also have

$$P(X < Y) = \int_{0}^{\infty} \int_{0}^{y} 2e^{-x}e^{-2y}dxdy,$$

which reduces to $P(X < Y) = 1/3$. ■

*Example #5: A Problem in Geometry*

Consider a circle of radius $R$, and suppose that a point within the circle is randomly chosen in such a manner that all regions within the circle of equal area are equally likely to contain the point. In other words, the point is uniformly distributed within the circle. Let the center of the circle be at the origin and define $X$ and $Y$ to be the coordinates of the point chosen, as shown below.



The point $(X, Y)$

Then since each $(X, Y)$ is equally likely to be near each point on the circle, it follows that the joint probability density function of $X$ and $Y$ is given by

$$f(x, y) = \begin{cases} c, & \text{for} \quad x^2 + y^2 \leq R^2 \\ 0, & \text{for} \quad x^2 + y^2 > R^2 \end{cases}$$

where $c$ is determined using

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

and this leads to

$$\int_{-R}^{+R} \int_{-\sqrt{R^2 - x^2}}^{+\sqrt{R^2 - x^2}} c dy dx = c\pi R^2 = 1$$

and so

$$f(x, y) = \frac{1}{\pi R^2} \times \begin{cases} 1, & \text{for} \quad x^2 + y^2 \leq R^2 \\ 0, & \text{for} \quad x^2 + y^2 > R^2 \end{cases} .$$

To determine the pdf ($f_X(x)$) for $X$, we have

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} \frac{1}{\pi R^2}dy = \frac{2}{\pi R^2}\sqrt{R^2-x^2}$$

for $-R \leq x \leq +R$. By symmetry, we also have

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_{-\sqrt{R^2-y^2}}^{+\sqrt{R^2-y^2}} \frac{1}{\pi R^2}dx = \frac{2}{\pi R^2}\sqrt{R^2-y^2}$$

for $-R \leq y \leq +R$, as the pdf for $Y$. Next, let $D$ be the random variable giving the distance the point $(X,Y)$ is from the origin, then

$$F_D(z) = P(D \leq z) = P(x^2+y^2 < z^2) = \frac{\pi z^2}{\pi R^2} = \left(\frac{z}{R}\right)^2$$

for $0 \leq z \leq R$. Using this, we have

$$f_D(z) = \frac{dF(z)}{dz} = \frac{2z}{R^2}$$

for $0 \leq z \leq R$. ■

*Example #6: Computing the pdf of $Z = X/Y$*

Suppose that the joint pdf of $X$ and $Y$ is given by

$$f(x,y) = \begin{cases} e^{-(x+y)}, & \text{for} \quad 0 < x, 0 < y \\ \\ 0, & \text{for} \quad \text{other values of } x \text{ and } y \end{cases},$$

to compute the pdf of $Z = X/Y$, we use the cdf (since this involves probabilities) and write

$$F(z) = P(Z \leq z) = P(X/Y \leq z) = \iint_{x/y \leq z} f(x,y)dxdy$$

or

$$F(z) = \int_0^\infty \int_0^{yz} f(x,y)dxdy = \int_0^\infty \int_0^{yz} e^{-(x+y)}dxdy$$

which reduces to

$$F(z) = \int_0^\infty \int_0^{yz} e^{-(x+y)}dxdy = \frac{z}{z+1}$$

for $0 \leq z$. Then

$$f(z) = \frac{dF(z)}{dz} = \frac{d}{dz}\left(\frac{z}{z+1}\right) = \frac{1}{(z+1)^2}$$

for $0 \leq z$. ■

Of course, all of the above can be extended to several variables using

$$F(x_1, x_2, x_3, ..., x_n) = P\left(\bigcap_{k=1}^{n}(X_k \leq x_k)\right) \tag{11a}$$

and if $X_1$, $X_2$, $X_3$, ..., $X_n$ are all continuous random variables, we have

$$f(x_1, x_2, x_3, ..., x_n) = \frac{\partial^n F(x_1, x_2, x_3, ..., x_n)}{\partial x_1 \partial x_2 \partial x_3 \cdots \partial x_n} \tag{11b}$$

as the joint pdf function for $X_1$, $X_2$, $X_3$, ..., $X_n$.

*Example #7: The Multinomial Distribution*

Suppose that a sequence of $n$ independent and identical experiments are performed and suppose that each experiment can result in $r$ possible outcomes, with respective probabilities $p_1$, $p_2$, $p_3$, ..., $p_r$, where $p_1 + p_2 + p_3 + \cdots + p_r = 1$. If we let $X_j$ denote the number of the $n$ experiments that result in outcome number $n_j$ for $j = 1, 2, 3, ..., r$, then

$$p(n_1, n_2, n_3, ..., n_r) = P((X_1 = n_1) \cap (X_2 = n_2) \cap \cdots \cap (X_r = n_r))$$

leads to

$$p(n_1, n_2, n_3, ..., n_r) = \binom{n}{n_1\ n_2\ n_3\ \cdots\ n_r} p_1^{n_1} p_2^{n_2} p_3^{n_3} \cdots p_r^{n_r} \tag{12}$$

where $n_1 + n_2 + n_3 + \cdots + n_r = n$. For a specific example, suppose that a fair die is rolled 9 times. The probability that 1 appears 3 times, 2 and 3 appear twice, 4 and 5 appear once each and 6 does not appear at all (zero times) is

$$p(3, 2, 2, 1, 1, 0) = \binom{9}{3\ 2\ 2\ 1\ 1\ 0}\left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^0$$

or

$$p(3,2,2,1,1,0) = \frac{9!}{3!2!2!1!1!0!}\left(\frac{1}{6}\right)^9 = \frac{35}{23328}$$

or $p(3,2,2,1,1,0) = 0.0015 = 0.15\%$. ■

**2. Independent Random Variables**

Two random variables $X$ and $Y$ are said to be independent, if

$$P((X \in A) \cap (Y \in B)) = P(X \in A)P(Y \in B) \tag{13a}$$

for any two sets $A$ and $B$. This says that the events $E_A = X \in A$ and $E_B = Y \in B$ are independent. This also says that

$$F(x,y) = P((X \leq x) \cap (Y \leq y)) = P(X \leq x)P(Y \leq y) = F_X(x)F_Y(y) \tag{13b}$$

for all $x \in R_X$ and $y \in R_Y$. When $X$ and $Y$ are discrete random variables, we find that

$$p(x,y) = p_X(x)p_Y(x) \tag{14a}$$

for all $x \in R_X$ and $y \in R_Y$, and when $X$ and $Y$ are continuous random variables, we find that

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial y \partial x} = \frac{\partial^2 (F_X(x)F_Y(y))}{\partial y \partial x} = F'_X(x)F'_Y(y)$$

or

$$f(x,y) = f_X(x)f_Y(x) \tag{14b}$$

for all $x \in R_X$ and $y \in R_Y$. This says that $X$ and $Y$ are independent if knowing the value of one does not change the distribution of the other. Random variables that are not independent are said to be dependent.

*Example #8*

A man and woman decide to meet at a certain location. If each of them independently arrives at a time uniformly distributed between 12 noon and 1 PM, find the probability that the first to arrive has to wait longer than 10 minutes. To answer this, $X$ and $Y$ denote, respectively, the time past 12 noon that the man and the woman arrive, then $X$ and $Y$ are independent random variables each of which is uniformly distributed over the interval from 0 to 60 minutes. The desired probability is

$$P = P(|Y - X| > 10) = P((Y - X > 10) \cup (Y - X < -10))$$

or

$$P = P(Y > X + 10) + P(Y < X - 10).$$

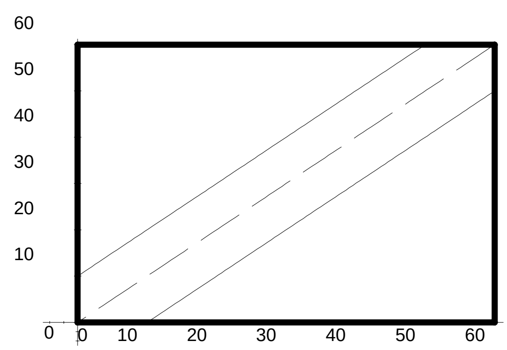
This leads to

$$P(Y > X + 10) = \int_0^{50} \int_{x+10}^{60} f_X(x) f_Y(y) dy dx = \frac{25}{72},$$

and

$$P(Y < X - 10) = \int_{10}^{60} \int_0^{x-10} f_X(x) f_Y(y) dy dx = \int_{10}^{60} \int_0^{x-10} \frac{1}{60} \times \frac{1}{60} dy dx = \frac{25}{72}$$

so that $P = 25/36 \simeq 0.694$. A plot of the region in which any one has to wait more than ten minutes for the other is shown as the upper left and lower right triangles in the figure below.




The region in which any one has to wait more than 10 minutes for the other is shown as the upper left and lower right triangles

The total area of these triangles compared to the total area of the square is

$$P = \frac{\frac{1}{2}(50)(50) + \frac{1}{2}(50)(50)}{(60)(60)} = \frac{25}{36}$$

which is the same answer obtained via integration. Note that just comparing areas when dealing with uniform distributions is valid and sometimes simpler than doing the integration. ■

*Example #9: Balls and Urns Example #2 Revisited*

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white and 5 blue balls. If we let $X$ and $Y$ denote, respectively, the number of red and white balls drawn, then $R_X = \{0,1,2,3\}$ and $R_Y = \{0,1,2,3\}$ with

$$p(x,y) = \frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}}{\binom{5+4+3}{3}} = \frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}}{\binom{12}{3}}$$

or

$$p(x,y) = \frac{1}{220}\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}.$$

The results of these can be summarized in the following table

The Values of $220 \times p(x,y)$

| $x\backslash y$ | 0 | 1 | 2 | 3 | $p_X(x) =$ Row Sums |
|---|---|---|---|---|---|
| 0 | 10 | 40 | 30 | 4 | 84 |
| 1 | 30 | 60 | 18 | 0 | 108 |
| 2 | 15 | 12 | 0 | 0 | 27 |
| 3 | 1 | 0 | 0 | 0 | 1 |
| $p_y(y) =$ Column Sums | 56 | 112 | 48 | 4 | Sums Equal 220 |

Note how $p_X(x)$and $p_y(y)$ are computed as marginal distributions by taking row and column sums, respectively. Note also that

$$p_X(1) = \frac{108}{220} \qquad \text{and} \qquad p_Y(2) = \frac{48}{220}$$

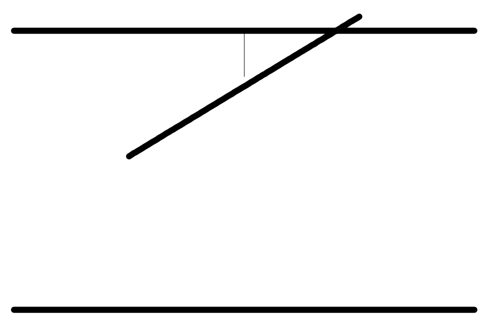
making

$$p_X(1)p_Y(2) = \frac{108}{220} \times \frac{48}{220} = \frac{324}{3025} \neq p(1,2) = \frac{18}{220}$$

and thereby showing that $X$ and $Y$ are not independent here. ■

*Example #10: A Geometry Problem - Buffon's Needle Problem*

A table is ruled with equidistant parallel lines a distance $D$ apart. A needle of length $L$ (where $L \leq D$) is randomly thrown on the table so that the perpendicular distance ($X$) from the center of needle to the closest of the parallel lines is uniform from 0 to $D/2$ and the angle ($\theta$) the needle makes with line connecting the center of the needle to the parallel line is also uniform from 0 to $\pi/2$, as shown below.



$X$ is the length of the dashed line and $\theta$ is the angle between the needle and the dashed line

Let us compute the probability that the needle will intersect one of the lines, the other possibility being that the needle will be completely contained in the strip between two lines. To answer this, let us determine the position of the needle by specifying (1) the distance $X$ from the midpoint of the needle to the nearest parallel line and (2) the angle $\theta$ between the needle and the projected line of length $X$. The needle will intersect the line if the hypotenuse of the right triangle in the above figure is less than $L/2$, that is if

$$X < \frac{L}{2}\cos(\theta).$$

As $X$ varies from 0 to $D/2$ and $\theta$ between 0 and $\pi/2$, it is reasonable to assume that they are independent, uniformly distributed random variables over these respective ranges. Hence

$$P = P(X < (L/2)\cos(\theta)) = \int_0^{\pi/2}\int_0^{(L/2)\cos(\theta)} \frac{1}{\pi/2} \times \frac{1}{D/2} dx d\theta$$

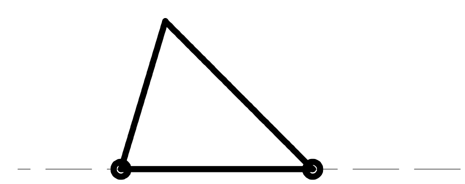
which reduces to $P = 2L/(\pi D)$. ■

*Example #11: A Geometry Problem - The Pipe Cleaner Problem*

Consider a pipe cleaner of length 1, which is simply a bendable wire of length 1. Two points are chosen at random on this wire and then the wire is bent at these points to see if it is possible to form a triangle, as shown in the following figure.

The pipe cleaner is bent at two points

The pipe cleaner is bend at two points to form a triangle as shown below.



The pipe cleaner is bent into a triangle

We would like to determine the probability that a triangle can be formed.

To solve this, let the two points be located at $R_1$ and $R_2$. We are given that $R_1 \sim U[0,1)$, $R_2 \sim U[0,1)$, so that their pdfs are $f_1(x) = 1$ for $0 \leq x < 1$ and $f_2(y) = 1$ for $0 \leq y < 1$ . We first note that the probability $P$ is given by

$$P \;=\; P((\text{Triangle with } R_1 \leq R_2) \cup (\text{Triangle with } R_2 \leq R_1))$$

$$= \quad P(\text{Triangle with } R_1 \le R_2) + P(\text{Triangle with } R_2 \le R_1)$$

since the two events $(R_1 \le R_2)$ and $(R_2 \le R_1)$ are disjoint. Assuming first that $R_1 \le R_2$, the three sides of the "possible" triangle being formed have lengths $R_1$, $R_2 - R_1$ and $1 - R_2$, and for these to form the sides of a triangle, we must have

$$R_1 \le (R_2 - R_1) + (1 - R_2) \quad \text{and} \quad R_2 - R_1 \le (R_1) + (1 - R_2)$$

and

$$1 - R_2 \le (R_1) + (R_2 - R_1)$$

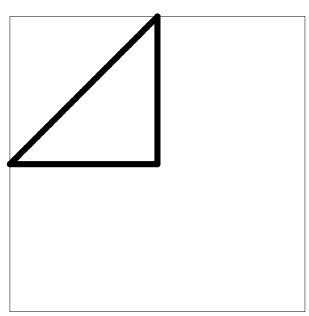
which all reduce to

$$R_1 \le 1/2 \qquad , \qquad R_2 - R_1 \le 1/2 \qquad , \qquad 1/2 \le R_2.$$

Thus we want to compute

$$P(\text{Triangle with } R_1 \le R_2) = P((R_1 \le 1/2) \cap (R_2 - R_1 \le 1/2) \cap (1/2 \le R_2))$$

which is a region in the unit square that looks like a right triangle with side lengths $1/2$ and $1/2$, as shown below.



The horizontal axis is $R_1$, the vertical axis is $R_2$ and the right triangle is shown in bold

Using areas, we then have

$$P(\text{Triangle with } R_1 \le R_2) = \frac{1}{2}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$$

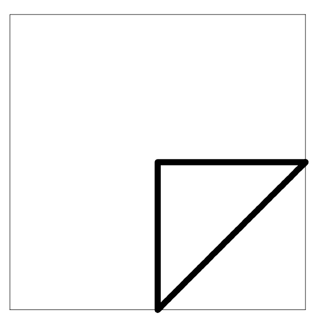
We can also compute this probability using

$$P(\text{Triangle with } R_1 \leq R_2) = \int_0^{1/2} \int_{1/2}^{x+1/2} f_2(y) f_1(x) dy dx$$

or

$$P(\text{Triangle with } R_1 \leq R_2) = \int_0^{1/2} \int_{1/2}^{x+1/2} \frac{1}{1} \times \frac{1}{1} dy dx = \frac{1}{8}$$

Assuming next that $R_2 \leq R_1$, the three sides of the triangle being formed have lengths $R_2$, $R_1 - R_2$ and $1 - R_1$, and since $R_1$ and $R_2$ are identical distributions and independent, we may simply interchange the rolls of $R_1$ and $R_2$ and use the result of the previous calculation. Thus, the symmetry in the problem leads to the area shown below



The horizontal axis is $R_1$, the vertical axis is $R_2$ and the right triangle is shown in bold

and

$$P(\text{Triangle with } R_2 \leq R_1) = \frac{1}{8}$$

Thus we find that

$$P = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

or $P = 25\%$. ■

*Example #12: Another Pipe-Cleaner Problem*

Consider the same pipe cleaner of the previous example. A point is chosen at random on this wire and then a second point is also chosen at random *but only to the right of* the first point. The wire is then bent at these two points to see if it is possible to form a triangle. Determine the probability that a triangle can be formed.

To solve this, let the two points be located at $R_1$ and $R_2$ with $R_2$ to the right of $R_1$. We are given that $R_1 \sim U[0,1)$, but this time we have $R_2 \sim U[R_1, 1)$, so that their pdfs are $f_1(x) = 1$ for $0 \leq x < 1$ and $f_2(y) = 1(1-x)$ for $x \leq y < 1$. The three sides of the "possible" triangle being formed have lengths $R_1$, $R_2 - R_1$ and $1 - R_2$, and for these to form the sides of a triangle, we must have

$$R_1 \leq (R_2 - R_1) + (1 - R_2) \quad \text{and} \quad R_2 - R_1 \leq (R_1) + (1 - R_2)$$

and

$$1 - R_2 \leq (R_1) + (R_2 - R_1)$$

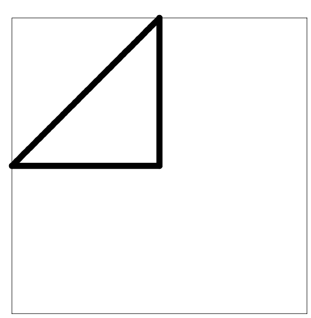
which all reduce to

$$R_1 \leq 1/2 \qquad , \qquad R_2 - R_1 \leq 1/2 \qquad , \qquad 1/2 \leq R_2.$$

Thus we want to compute

$$P(\text{Triangle with } R_1 \leq R_2) = P((R_1 \leq 1/2) \cap (R_2 \leq R_1 + 1/2) \cap (1/2 \leq R_2))$$

which is a region in the unit square that looks like a right triangle with side lengths 1/2 and 1/2, as shown below,



The horizontal axis is $R_1$, the vertical axis is $R_2$ and the right triangle is shown in bold but because $R_1$ and $R_2$ are not independent and not both uniform in the unit square here, we may not use simple areas to compute the probability. Instead, we have

$$P(\text{Triangle with } R_1 \leq R_2) = \int_0^{1/2} \int_{1/2}^{x+1/2} f_2(y|x) f_1(x) dy dx$$

or

$$P(\text{Triangle with } R_1 \leq R_2) = \int_0^{1/2} \int_{1/2}^{x+1/2} \frac{1}{1-x} \times \frac{1}{1} dy dx$$

which reduces to

$$P(\text{Triangle with } R_1 \leq R_2) = \ln(2) - \frac{1}{2}$$

or $P \simeq 19.3\%$. The fact that this is smaller than 25% from the previous example makes sense since we are being more restrictive in the choice of $R_2$ here. Note that in computing $f_2(y|x)$, we have

$$f_2(y|x) = \frac{1}{1-x}$$

since $Y|X \sim U[x,1)$. ■

*Example #13: Characterization of the Normal Distribution*

Let $X$ and $Y$ denote the horizontal and vertical miss distance when a bullet is fired at a target, and assume that (1) $X$ and $Y$ are independent continuous random variables having differentiable density functions and (2) the joint density $f(x,y) = f_X(x) f_Y(y)$ of $X$ and $Y$ depends on $(x,y)$ only through $x^2 + y^2$. This second assumption states that the probability of the bullet landing on any point of the $xy$ plane depends only on the distance of the point from the target and not on its angle of orientation. An equivalent way of phrasing this assumption is to say that the joint density function is rotational invariant. It is a rather interesting fact that only 1 and 2 imply that $X$ and $Y$ *must be* normally distributed random variables. To prove this, note first that the assumptions 1 and 2 yields the relation

$$f(x,y) = f_X(x) f_Y(y) \qquad \text{and} \qquad f(x,y) = g(x^2 + y^2),$$

respectively, so that

$$f_X(x) f_Y(y) = g(x^2 + y^2)$$

for some differentiable function $g$. Differentiating this equation with resect to $x$ leads to

$$f'_X(x) f_Y(y) = 2xg'(x^2 + y^2)$$

and so

$$\frac{f'_X(x) f_Y(y)}{f_X(x) f_Y(y)} = \frac{2xg'(x^2 + y^2)}{g(x^2 + y^2)} = \frac{f'_X(x)}{f_X(x)}$$

or

$$\frac{f'_X(x)}{2x f_X(x)} = \frac{g'(x^2 + y^2)}{g(x^2 + y^2)}.$$

Setting $y^2 = R^2 - x^2$, we see that

$$\frac{f'_X(x)}{2x f_X(x)} = \frac{g'(R^2)}{g(R^2)} = \text{a constant} = -c$$

since $R$ is a constant, and so

$$\frac{f'_X(x)}{f_X(x)} = -2cx \qquad \text{or} \qquad \frac{df_X(x)}{f_X(x)} = -2cxdx$$

or

$$\ln(f_X(x)) = -cx^2 + c_1 \qquad \text{or} \qquad f_X(x) = Ae^{-cx^2}.$$

Since

$$1 = \int_{-\infty}^{+\infty} f_X(x)dx = A \int_{-\infty}^{+\infty} e^{-cx^2} dx = A \int_{-\infty}^{+\infty} e^{-cx^2} dx = 2A \int_{0}^{+\infty} e^{-cx^2} dx.$$

Setting $z = cx^2$, we have $dz = 2cxdx$, we have

$$1 = 2A \int_0^{+\infty} e^{-z} \frac{dz}{2c(z/c)^{1/2}} = \frac{A}{\sqrt{c}} \int_0^{+\infty} z^{-1/2} e^{-z} dz = \frac{A}{\sqrt{c}} \sqrt{\pi}$$

which says that $A = (c/\pi)^{1/2}$, which further says that

$$f_X(x) = \sqrt{\frac{c}{\pi}} e^{-cx^2}.$$

Setting

$$c = \frac{1}{2\sigma^2} \qquad \text{we have} \qquad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(x/\sigma)^2}$$

for $-\infty < x < +\infty$, which shows that $X \sim N(0, \sigma^2)$. Using the same argument for $Y$, we have

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\bar{\sigma}} e^{-\frac{1}{2}(y/\bar{\sigma})^2}$$

for $-\infty < y < +\infty$. Finally, using assumption 2, we could easily show that $\bar{\sigma} = \sigma$ and the student should do this. Thus we find that

$$f(x, y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(x/\sigma)^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(y/\sigma)^2}$$

or

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}(x^2+y^2)/\sigma^2}$$

for $-\infty < x, y < +\infty$. Using this we may now compute the probability that the point $(X, Y)$ is within a distance $R$ from the origin, *i.e.*, $P(D \leq R)$ where

$$D = \sqrt{X^2 + Y^2}$$

as

$$P(D \leq R) = P(D^2 \leq R^2) = P(X^2 + Y^2 \leq R^2)$$

which leads to

$$P(D \leq R) = \int_{-R}^{+R} \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} f(x, y) dy dx = \int_{-R}^{+R} \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}(x^2+y^2)/\sigma^2} dy dx.$$

Using polar coordinates $x = r\cos(\theta)$ and $y = r\sin(\theta)$, we have $x^2 + y^2 = r^2$ and $dA = dydx = rdrd\theta$, and this becomes

$$P(D \leq R) = \int_0^{2\pi} \int_0^R \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}(r^2/\sigma^2)} r dr d\theta = \frac{1}{\sigma^2} \int_0^R e^{-\frac{1}{2}(r^2/\sigma^2)} r dr$$

or

$$P(D \leq R) = -e^{-\frac{1}{2}(r^2/\sigma^2)}|_0^R = 1 - e^{-\frac{1}{2}(R/\sigma)^2}$$

which is a nice simple answer in the end. ■

## 3. The Minimum of $n$ Independent Random Variables

Suppose that $X_1$, $X_2$, $X_3$, …, $X_n$ are independent random variables with cdfs $F_1(x)$, $F_2(x)$, $F_3(x)$, …, $F_n(x)$, respectively, and suppose that $Y$ is a random variable defined by

$$Y = \min\{X_1, X_2, X_3, \ldots, X_n\}. \tag{15a}$$

Then the cdf of $Y$ (denoted by $G(y)$) can be computed using

$$\begin{aligned} G(y) &= P(Y \leq y) = 1 - P(Y > y) = 1 - P(X_1 > y, X_2 > y, \ldots, X_n > y) \\ &= 1 - P(X_1 > y)P(X_2 > y) \cdots P(X_n > y) = 1 - \prod_{k=1}^{n} P(X_k > y). \end{aligned}$$

But

$$P(X_k > y) = 1 - P(X_k \leq y) = 1 - F_k(y)$$

and so

$$G(y) = 1 - \prod_{k=1}^{n} (1 - F_k(y)). \tag{15b}$$

*Example #14: The Minimum of $n$ Exponential Random Variables*

Suppose now that $X_1$, $X_2$, $X_3$, $\ldots$, $X_n$ are all independent and exponential with parameters: $\lambda_1$, $\lambda_2$, $\lambda_3$, $\ldots$, $\lambda_n$, respectively, then

$$F_k(y) = 1 - e^{-\lambda_k y} \qquad \text{and so} \qquad 1 - F_k(y) = e^{-\lambda_k y}.$$

Then, if

$$Y = \min\{X_1, X_2, X_3, \ldots, X_n\} \tag{16a}$$

we find that

$$G(y) = 1 - \prod_{k=1}^{n} (1 - F_k(y)) = 1 - \prod_{k=1}^{n} e^{-\lambda_k y}$$

or simply

$$G(y) = 1 - e^{-(\lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_n)y} = 1 - e^{-\lambda y}$$

showing that $Y$ is *also exponential* with parameter

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_n \tag{16b}$$

for $y \geq 0$. ■

*Example #15: The Minimum of $n$ Geometric Random Variables*

Suppose now that $X_1$, $X_2$, $X_3$, …, $X_n$ are all independent and geometric with parameters: $p_1$, $p_2$, $p_3$, …, $p_n$, respectively, then

$$F_k(y) = 1 - (1 - p_k)^y \qquad \text{and so} \qquad 1 - F_k(y) = (1 - p_k)^y$$

and so

$$Y = \min\{X_1, X_2, X_3, \ldots, X_n\} \tag{17a}$$

leads to

$$G(y) = 1 - \prod_{k=1}^{n}(1 - F_k(y)) = 1 - \prod_{k=1}^{n}(1 - p_k)^y$$

or simply

$$G(y) = 1 - (1 - p)^y \qquad \text{with} \qquad (1 - p) = \prod_{k=1}^{n}(1 - p_k)$$

showing that $Y$ is also geometric with parameter

$$p = 1 - \prod_{k=1}^{n}(1 - p_k) \tag{17b}$$

for $y = 1, 2, 3, \ldots$. ■

*Example #16: The Minimum of $n$ Uniform Random Variables*

Suppose now that $X_1$, $X_2$, $X_3$, …, $X_n$ are all independent and continuously uniform in the interval $[a, b)$, then

$$F_k(y) = \frac{y - a}{b - a} \qquad \text{and} \qquad 1 - F_k(y) = \frac{b - y}{b - a}$$

and

$$Y = \min\{X_1, X_2, X_3, \ldots, X_n\} \tag{18a}$$

has cdf given by

$$G(y) = 1 - \prod_{k=1}^{n}(1 - F_k(y)) = 1 - \prod_{k=1}^{n}\left(\frac{b - y}{b - a}\right)$$

or

$$G(y) = 1 - \left(\frac{b-y}{b-a}\right)^n \tag{18b}$$

for $a \leq y \leq b$. ■

## 4. The Maximum of $n$ Independent Random Variables

Suppose that $X_1$, $X_2$, $X_3$, …, $X_n$ are independent random variables with cdfs: $F_1(x)$, $F_2(x)$, $F_3(x)$, …, $F_n(x)$, respectively, and suppose that $Z$ is a random variable defined by

$$Z = \max\{X_1, X_2, X_3, \ldots, X_n\}, \tag{19a}$$

then the cdf of $Z$ (denoted by $H(z)$) can be computed using

$$H(z) = \Pr(Z \leq z) = \Pr(X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z)$$

resulting in

$$H(z) = \prod_{k=1}^{n} \Pr(X_k \leq z) = \prod_{k=1}^{n} F_k(z). \tag{19b}$$

*Example #17: The Maximum of $n$ Uniform Random Variables*

Suppose now that $X_1$, $X_2$, $X_3$, …, $X_n$ are all independent and continuously uniform in the interval $[a, b)$, then

$$F_k(y) = \frac{y-a}{b-a} \qquad \text{and} \qquad 1 - F_k(y) = \frac{b-y}{b-a}$$

and

$$Z = \max\{X_1, X_2, X_3, \ldots, X_n\} \tag{20a}$$

has cdf given by

$$H(z) = \prod_{k=1}^{n} F_k(z) = \prod_{k=1}^{n} \left(\frac{z-a}{b-a}\right)$$

or

$$H(z) = \left(\frac{z-a}{b-a}\right)^n \tag{20b}$$

for $a \leq z \leq b$. ■

## 5. The Sum of Two Independent Discrete Random Variables

Suppose that $X$ and $Y$ are both independent discrete random variables with pmfs $p_X(x)$ and $p_Y(y)$ and suppose that $Z = X + Y$. Then the probability that $Z = z$ is given by

$$p_Z(z) = P(Z = z) = P(Y + X = z) = \sum_x P(Y + x = z | X = x) P(X = x)$$

which can be expressed as

$$p_Z(z) = \sum_x P(Y = z - x) p_X(x)$$

or simply

$$p_Z(z) = \sum_x p_X(x) p_Y(z - x) \tag{21}$$

where the sum in over all possible values of $x$ in which both $p_X(x)$ and $p_Y(z - x)$ are non-zero. This is known as a *convolution sum.*

*Example #18: The Sum of Two Geometric Distributions*

Suppose that $X$ is geometric with parameter $0 \leq p \leq 1$ so that

$$p_X(x) = \begin{cases} 0, & \text{for} \quad x = 0, -1, -2, ... \\ p(1-p)^{x-1}, & \text{for} \quad x = 1, 2, 3, ... \end{cases} \tag{22a}$$

and $Y$ is geometric with parameter $0 \leq q \leq 1$ so that

$$p_Y(y) = \begin{cases} 0, & \text{for} \quad y = 0, -1, -2, ... \\ q(1-q)^{y-1}, & \text{for} \quad y = 1, 2, 3, ... \end{cases} . \tag{22b}$$

Assuming first that $p \neq q$, we note that if $Z = X + Y$, we have

$$p_Z(z) = \sum_x p_X(x) p_Y(z - x) = \sum_{x=1}^{z-1} p_X(x) p_Y(z - x)$$

for $z = 2, 3, 4, ...,$ since $p_X(x)$ and $p_Y(z-x)$ are both non-zero only for $x = 1, 2, 3, ..., z-1$. This says that

$$p_Z(z) = \sum_{x=1}^{z-1} p(1-p)^{x-1} q(1-q)^{z-x-1} = pq(1-q)^{z-2} \sum_{x=1}^{z-1} \left(\frac{1-p}{1-q}\right)^{x-1}$$

or

$$p_Z(z) = pq(1-q)^{z-2} \sum_{x=0}^{z-2} \left(\frac{1-p}{1-q}\right)^{x} = pq(1-q)^{z-2} \left(\frac{1-((1-p)/(1-q))^{z-1}}{1-(1-p)/(1-q)}\right)$$

which reduces to

$$p_Z(z) = pq\left(\frac{(1-q)^{z-1}-(1-p)^{z-1}}{p-q}\right) \tag{22c}$$

for $z = 2, 3, 4, ...,$ and $p \neq q$, and $p_Z(z) = 0$ for $z = 1, 0, -1, -2, ....$ As a check, we note that

$$\begin{aligned}
\sum_{z=2}^{\infty} p_Z(z) &= \sum_{z=2}^{\infty} pq\left(\frac{(1-q)^{z-1}-(1-p)^{z-1}}{p-q}\right) \\
&= \frac{pq}{p-q}\left(\sum_{z=2}^{\infty}(1-q)^{z-1} - \sum_{z=2}^{\infty}(1-p)^{z-1}\right) \\
&= \frac{pq}{p-q}\left(\sum_{z=1}^{\infty}(1-q)^{z} - \sum_{z=1}^{\infty}(1-p)^{z}\right) \\
&= \frac{pq}{p-q}\left(\frac{1-q}{1-(1-q)} - \frac{1-p}{1-(1-p)}\right)
\end{aligned}$$

or

$$\sum_{z=2}^{\infty} p_Z(z) = 1$$

as it must. Assuming next that $p = q$, we note that if $Z = X + Y$, we have

$$p_Z(z) = \sum_{x} p_X(x)p_Y(z-x) = \sum_{x=1}^{z-1} p_X(x)p_Y(z-x)$$

for $z = 2, 3, 4, ...,$ since $p_X(x)$ and $p_Y(z-x)$ are both non-zero only for $x = 1, 2, 3, ..., z-1$. This says that

$$p_Z(z) = \sum_{x=1}^{z-1} p(1-p)^{x-1} p(1-p)^{z-x-1} = p^2(1-p)^{z-2} \sum_{x=1}^{z-1} 1$$

or

$$p_Z(z) = (z-1)p^2(1-p)^{z-2} \tag{22d}$$

for $z = 2, 3, 4, ...,$ and $p = q$, and $p_Z(z) = 0$ for $z = 1, 0, -1, -2, ....$ As a check, we note that

$$\sum_{z=2}^{\infty} p_Z(z) = \sum_{z=2}^{\infty}(z-1)p^2(1-p)^{z-2} = p^2 \sum_{z=1}^{\infty} z(1-p)^{z-1} = 1$$

as it must. One final point is to note that

$$\lim_{q \to p}\left(pq\left(\frac{(1-q)^{z-1} - (1-p)^{z-1}}{p-q}\right)\right) = p^2(z-1)(1-p)^{z-2}$$

showing that Equation (22d) is consistent with Equation (22c). ■

*Example #19: The Sum of Two Poisson Distributions*

Suppose that $X$ and $Y$ are two Poisson random variables with parameters $\alpha$ and $\beta$, respectively, so that

$$p_X(x) = \frac{e^{-\alpha}\alpha^x}{x!} \qquad \text{and} \qquad p_Y(y) = \frac{e^{-\beta}\beta^y}{y!} \tag{23a}$$

for $x = 0, 1, 2, ...,$ and $y = 0, 1, 2, 3, ...,$ and suppose that $Z = X + Y$, then the pmf of $Z$ can be computed as

$$\begin{aligned} p_Z(z) &= \sum_{x=0}^{z} p_X(x)p_Y(z-x) = \sum_{x=0}^{z} \frac{e^{-\alpha}\alpha^x}{x!}\frac{e^{-\beta}\beta^{z-x}}{(z-x)!} = \sum_{x=0}^{z} \frac{e^{-(\alpha+\beta)}\alpha^x\beta^{z-x}}{x!(z-x)!} \\ &= \frac{e^{-(\alpha+\beta)}}{z!}\sum_{x=0}^{z} z!\frac{\alpha^x\beta^{z-x}}{x!(z-x)!} = \frac{e^{-(\alpha+\beta)}}{z!}\sum_{x=0}^{z}\binom{z}{x}\alpha^x\beta^{z-x} \end{aligned}$$

and so we have

$$p_Z(z) = \frac{e^{-(\alpha+\beta)}(\alpha+\beta)^z}{z!} \tag{23b}$$

for $z = 0, 1, 2, 3, ...,$ which shows that $Z = X + Y$ is also Poisson with parameter $\alpha + \beta$. In general, if $X_1$, $X_2$, $X_3$, …, $X_n$ are all Poisson with parameters $\alpha_1$, $\alpha_2$, $\alpha_3$, …, $\alpha_n$, then

$$X = X_1 + X_2 + X_3 + \cdots + X_n \tag{23c}$$

is Poisson with parameter

$$\alpha = \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n. \tag{23d}$$

## 6. The Sum of Two Independent Continuous Random Variables

Suppose that $X$ and $Y$ are both independent continuous random variables with pdfs $f_X(x)$ and $f_Y(y)$ and suppose that $Z = X + Y$. Then the probability that $Z \leq z$ is given by

$$F_Z(z) = P(Z \leq z) = P(Y + X \leq z) = \int P(Y + x \leq z | X = x) P(X = x) dx$$

where the integration in over all possible values of $x$. This can be expressed as

$$F_Z(z) = \int P(Y \leq z - x | X = x) f_X(x) dx$$

or simply

$$F_Z(z) = \int P(Y \leq z - x) f_X(x) dx = \int F_Y(z - x) f_X(x) dx.$$

By taking the derivative with respect to $z$, we find that

$$f_Z(z) = \int f_X(x) f_Y(z - x) dx \tag{24}$$

where the integration is over all values of $x$ in which both $f_X(x)$ and $f_Y(z - x)$ are non-zero. This is known as a *convolution integral.*

*Example #20: The Sum of Two Exponential Distributions*

Suppose that $X$ is exponential with parameter $0 < \lambda$ so that

$$f_X(x) = \begin{cases} 0, & \text{for} \quad x < 0 \\ \\ \lambda e^{-\lambda x}, & \text{for} \quad x \geq 0 \end{cases} \tag{25a}$$

and $Y$ is exponential with parameter $0 < \mu$ so that

$$f_Y(y) = \begin{cases} 0, & \text{for} \quad y < 0 \\ \\ \mu e^{-\mu y}, & \text{for} \quad y \geq 0 \end{cases}. \tag{25b}$$

Assuming first that $\lambda \neq \mu$, we note that if $Z = X + Y$, we have

$$f_Z(z) = \int_0^\infty f_X(x) f_Y(z - x) dx = \int_0^z \lambda e^{-\lambda x} \mu e^{-\mu(z-x)} dx$$

for $z \geq 0$, since $f_X(x)$ and $f_Y(z - x)$ are both non-zero only for $0 \leq x \leq z$. This says that

$$f_Z(z) = \lambda\mu e^{-\mu z} \int_0^z e^{-\lambda x + \mu x} dx = \lambda\mu e^{-\mu z} \left( \frac{1 - e^{-z(\lambda-\mu)}}{\lambda - \mu} \right)$$

or

$$f_Z(z) = \lambda\mu \left( \frac{e^{-\mu z} - e^{-\lambda z}}{\lambda - \mu} \right) \tag{25c}$$

for $z \geq 0$ and $\lambda \neq \mu$, and $f_Z(z) = 0$ for $z < 0$. As a check, we note that

$$\int_0^\infty p_Z(z) dz = \int_0^\infty \lambda\mu \left( \frac{e^{-\mu z} - e^{-\lambda z}}{\lambda - \mu} \right) dz = 1$$

as it must. Assuming next that $\lambda = \mu$, we note that if $Z = X + Y$, we have

$$f_Z(z) = \int_0^\infty f_X(x) f_Y(z - x) dx = \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx$$

for $z \geq 0$, since $f_X(x)$ and $f_Y(z - x)$ are both non-zero only for $0 \leq x \leq z$. This says that

$$f_Z(z) = \lambda^2 z e^{-\lambda z} \tag{25d}$$

for $z \geq 0$ and $\lambda = \mu$ and $f_Z(z) = 0$ for $z < 0$. As a check, we note that

$$\int_0^\infty f_Z(z) dz = \int_0^\infty \lambda^2 z e^{-\lambda z} dz = 1$$

as it must. One final point is to note that

$$\lim_{\mu \to \lambda} \left( \lambda\mu \left( \frac{e^{-\mu z} - e^{-\lambda z}}{\lambda - \mu} \right) \right) = \lambda^2 z e^{-\lambda z}$$

showing that Equation (25d) is consistent with Equation (25c). ■

*Example #21: The Sum of Two Normal Distributions*

Suppose that $X$ is normal with parameters $\mu_1$ and $\sigma_1^2$ so that

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) \tag{26a}$$

and $Y$ is normal with parameters $\mu_2$ and $\sigma_2^2$ so that

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right). \tag{26b}$$

Then if $Z = X + Y$, we have

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{z-x-\mu_2}{\sigma_2}\right)^2\right) dx. \end{aligned}$$

Setting $w = (x-\mu_1)/\sigma_1$, we have $x = \mu_1 + w\sigma_1$ and $dw = dx/\sigma_1$, and so

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi\sigma_2} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}w^2\right) \exp\left(-\frac{1}{2}\left(\frac{z-\mu_1-w\sigma_1-\mu_2}{\sigma_2}\right)^2\right) dw \\ &= \frac{1}{2\pi\sigma_2} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}w^2\right) \exp\left(-\frac{1}{2}\left(\frac{z-w\sigma_1-(\mu_1+\mu_2)}{\sigma_2}\right)^2\right) dw \end{aligned}$$

or

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)}{\sigma_2}\right)^2\right) \\ &\quad \times \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}w^2\left(1+\frac{\sigma_1^2}{\sigma_2^2}\right) + w\sigma_1\left(\frac{z-(\mu_1+\mu_2)}{\sigma_2^2}\right)\right) dw. \end{aligned}$$

Using the fact that

$$\int_{-\infty}^{+\infty} e^{-aw^2+bw} dw = e^{b^2/4a}\sqrt{\frac{\pi}{a}} \tag{26c}$$

for $a > 0$, we see that

$$f_Z(z) = \frac{1}{2\pi\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)}{\sigma_2}\right)^2\right) \exp\frac{\left(\sigma_1\left(\frac{z-(\mu_1+\mu_2)}{\sigma_2^2}\right)\right)^2}{2\left(1+\frac{\sigma_1^2}{\sigma_2^2}\right)} \sqrt{\frac{2\pi}{1+\frac{\sigma_1^2}{\sigma_2^2}}}$$

which reduces to

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left(-\frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2\right) \tag{26d}$$

showing that $Z = X + Y$ is also normal with parameters $\mu_1 + \mu_2$ and $\sigma_1^2 + \sigma_2^2$. ■

*Example #22: The Sum of Two Identical Uniform Distributions*

Suppose that $X \sim U[a, b)$ and $Y \sim U[a, b)$ so that if $Z = X + Y$, then

$$R_Z = \{x | 2a \leq x < 2b\}.$$

Then

$$f_X(x) = \begin{cases} 1/(b-a), & \text{for} \quad a \leq x < b \\ 0, & \text{for} \quad \text{all other values of } x \end{cases} \tag{27a}$$

and

$$f_Y(y) = \begin{cases} 1/(b-a), & \text{for} \quad a \leq y < b \\ 0, & \text{for} \quad \text{all other values of } y \end{cases} \tag{27b}$$

and

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)dx = \int_a^b \frac{1}{b-a} f_Y(z-x)dx$$

or

$$f_Z(z) = \frac{1}{b-a}\int_a^b f_Y(z-x)dx.$$

Now $f_Y(z-x) \neq 0$ only for $a \leq z-x < b$, which says that

$$z - b < x \leq z - a$$

and so

$$f_Z(z) = \frac{1}{b-a}\int_{\max(a,z-b)}^{\min(b,z-a)} \frac{1}{b-a}dx = \frac{1}{(b-a)^2}\int_{\max(a,z-b)}^{\min(b,z-a)} dx$$

or

$$f_Z(z) = \frac{\min(b,z-a) - \max(a,z-b)}{(b-a)^2}. \tag{27c}$$

For $2a \le z \le a+b$, we have $z-a \le b$ and $z-b \le a$ so that

$$\min(b,z-a) = z-a \qquad \text{and} \qquad \max(a,z-b) = a$$

and then

$$f_Z(z) = \frac{z-a-a}{(b-a)^2} = \frac{z-2a}{(b-a)^2}.$$

But for $a+b \le z \le 2b$, we have $z-a \ge b$ and $z-b \ge a$ so that

$$\min(b,z-a) = b \qquad \text{and} \qquad \max(a,z-b) = z-b$$

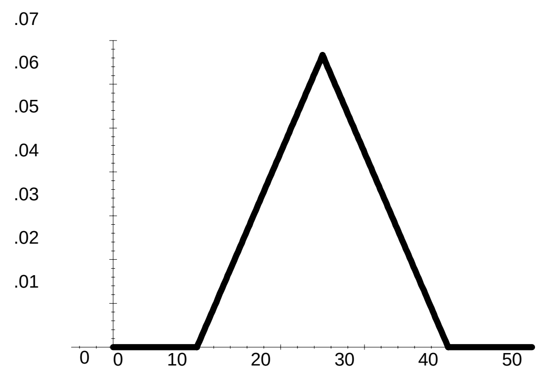
and then

$$f_Z(z) = \frac{b-(z-b)}{(b-a)^2} = \frac{2b-z}{(b-a)^2}.$$

Thus we find that

$$f_Z(z) = \frac{1}{(b-a)^2} \times \begin{cases} z-2a, & \text{for} \quad 2a \le z \le a+b \\ \\ 2b-z, & \text{for} \quad a+b \le z \le 2b \end{cases} \tag{27d}$$

which is a symmetric triangular distribution. A plot using $a = 5$ and $b = 20$ is shown below.



Plot of $f_Z(z)$ versus $z$ for $a = 5$ and $b = 20$

and it shows a symmetric triangular distribution. ■

## 7. The Erlang Distribution

We have seen in the previous chapter that the sum of $k$ independent and identical exponential distributions (all with parameter $\lambda$)

$$X = X_1 + X_2 + X_3 + \cdots + X_k, \tag{28a}$$

is an Erlang distribution with parameters $\theta = \lambda/k$ and $k$. The pdf is given by

$$f_X(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \tag{28b}$$

for $x \geq 0$ and $f_X(x) = 0$ for $x < 0$. We had also seen that the cdf of the Erlang distribution is given by

$$F_X(x) = 1 - e^{-\lambda x} \sum_{i=0}^{k-1} \frac{(\lambda x)^i}{i!} \tag{28c}$$

for $0 < x$, and $F(x) = 0$ for $x \leq 0$.

*Example #23: Being In The Dark*

A box contains four identical light bulbs. Each light bulb has a lifetime distribution that is exponential with a mean of 150 hours. Person $A$ takes two of these light bulbs into a room and turns them both on at the same time and leaves them turned on. Person $B$ takes the other two light bulbs into a different room and turns on only one of the light bulbs and then turns on the other light bulb only when and if the first light bulb burns out. (a) Compute the probability that person $A$ will be in the dark at the end of one week (168 hours). (b) Compute the probability that person $B$ will be in the dark at the end of one week (168 hours).

To solve part (a), let $X_1$, $X_2$, $X_3$, and $X_4$ be the random variables for the lifetime (in hours) of the four light bulbs so that

$$F_1(x) = F_2(x) = F_3(x) = F_4(x) = 1 - e^{-\lambda x}$$

with $\lambda = 1/150$. Since person $A$ keeps both light bulbs on at the same time, this person will be in the dark after one week ($24 \times 7 = 168$ hours) only when

$$Y = \max(X_1, X_2) \leq 168.$$

It should be clear that

$$F_{\max}(y) = P(Y \leq y) = P(\max(X_1, X_2) \leq y) = P((X_1 \leq y) \cap (X_2 \leq y)).$$

Since $X_1$ and $X_2$ are independent, we have

$$F_{\max}(y) = P((X_1 \leq y) \cap (X_2 \leq y)) = P(X_1 \leq y)P(X_2 \leq y)$$

resulting in

$$F_{\max}(y) = F_1(y)F_2(y) = (1 - e^{-y/150})^2$$

and so

$$P(Y \leq 168) = (1 - e^{-168/150})^2 \simeq 0.4539.$$

To solve (b), we note that since person $B$ keeps the light bulbs on only one at the same time, this person will be in the dark after one week ($24 \times 7 = 168$ hours) only when

$$Z = X_3 + X_4 \leq 168.$$

From our notes in class we know that $Z$ will be an Erlang distribution with parameters $k = 2$, and $\theta = 1/300$. Then

$$F_{\text{sum}}(z) = 1 - \sum_{i=0}^{2-1} e^{-2(1/300)z} \frac{(2(1/300)z)^i}{i!}$$

which reduces to

$$F_{\text{sum}}(z) = 1 - \left(1 + \frac{z}{150}\right) e^{-z/150}.$$

Then

$$P(Z \leq 168) = 1 - \left(1 + \frac{168}{150}\right) e^{-168/150},$$

or $P(Z \leq 168) \simeq 0.3083$. ■

**8. Sum of Squared of Independent Standard Normals is Chi-Squared**

We shall show in the next chapter that if $Z_1$, $Z_2$, $Z_3$, ..., $Z_n$ are all independent and from $N(0, 1)$, then the random variable

$$X = Z_1^2 + Z_2^2 + Z_3^2 + \cdots + Z_n^2$$

has a chi-squared distribution with $n$ degrees of freedom.

*The Chi-Squared Distribution With $\nu$ Degrees of Freedom ($\chi^2_\nu$)*

The chi-squared distribution with $\nu$ degrees of freedom (denoted by $\chi^2_\nu$) is a special case of a *gamma distribution* with $\beta = \nu/2$ and $\theta = 1/\nu$. This then has a pdf given by

$$f_\nu(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}$$

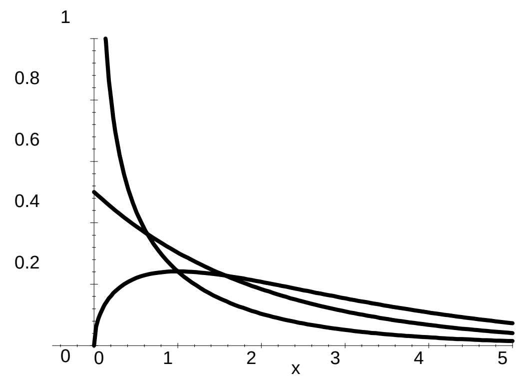
for $0 \leq x$ and $f(x) = 0$ for $x < 0$. The cdf of this distribution is given by

$$F_\nu(x) = \int_0^x f(z)dz = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x z^{\nu/2-1} e^{-z/2} dz.$$

Recall that

$$\Gamma(\nu/2) = \int_0^\infty z^{\nu/2-1} e^{-z} dz$$

is the *gamma function* evaluated at $\nu/2$ and is given by $(\nu/2 - 1)!$ when $\nu/2$ is an integer. The *mean* and *variance* of the $\chi^2_\nu$ distribution are given by $E(X) = \nu$ and $V(X) = 2\nu$, respectively. Typical plots of $f_\nu(x)$ are shown in the figure below.




Plots of $f_\nu(x)$ for $\nu = 1$ and 2 and 3
showing: $f_1(0) > f_2(0) > f_3(0)$

Note that the mode of the $\chi^2_\nu$ distribution is given by $\nu - 2$ for $2 \le \nu$. Note also that for $\nu = 2$, we find that

$$f_2(x) = \frac{1}{2^{2/2}\Gamma(2/2)} x^{2/2-1} e^{-x/2} = \frac{1}{2} e^{-x/2}$$

which is an *exponential distribution* with mean 2.