Chapter 7

Sums of Independent Random Variables

7.1 Sums of Discrete Random Variables

In this chapter we turn to the important question of determining the distribution of a sum of independent random variables in terms of the distributions of the individual constituents. In this section we consider only sums of discrete random variables, reserving the case of continuous random variables for the next section.

We consider here only random variables whose values are integers. Their distribution functions are then defined on these integers. We shall find it convenient to assume here that these distribution functions are defined for *all* integers, by defining them to be 0 where they are not otherwise defined.

Convolutions

Suppose X and Y are two independent discrete random variables with distribution functions $m_1(x)$ and $m_2(x)$. Let $Z = X + Y$. We would like to determine the distribution function $m_3(x)$ of Z. To do this, it is enough to determine the probability that Z takes on the value z, where z is an arbitrary integer. Suppose that $X = k$, where k is some integer. Then $Z = z$ if and only if $Y = z - k$. So the event $Z = z$ is the union of the pairwise disjoint events

$$(X = k) \text{ and } (Y = z - k) ,$$

where k runs over the integers. Since these events are pairwise disjoint, we have

$$P(Z = z) = \sum_{k=-\infty}^{\infty} P(X = k) \cdot P(Y = z - k) .$$

Thus, we have found the distribution function of the random variable Z. This leads to the following definition.

Definition 7.1 Let X and Y be two independent integer-valued random variables, with distribution functions $m_1(x)$ and $m_2(x)$ respectively. Then the *convolution* of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 * m_2$ given by

$$m_3(j) = \sum_k m_1(k) \cdot m_2(j - k) ,$$

for $j = \ldots, -2, -1, 0, 1, 2, \ldots$. The function $m_3(x)$ is the distribution function of the random variable $Z = X + Y$. $\square$

It is easy to see that the convolution operation is commutative, and it is straightforward to show that it is also associative.

Now let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent random variables of an independent trials process with common distribution function m defined on the integers. Then the distribution function of S_1 is m. We can write

$$S_n = S_{n-1} + X_n .$$

Thus, since we know the distribution function of X_n is m, we can find the distribution function of S_n by induction.

Example 7.1 A die is rolled twice. Let X_1 and X_2 be the outcomes, and let $S_2 = X_1 + X_2$ be the sum of these outcomes. Then X_1 and X_2 have the common distribution function:

$$m = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

The distribution function of S_2 is then the convolution of this distribution with itself. Thus,

$$
\begin{aligned}
P(S_2 = 2) &= m(1)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} , \\
P(S_2 = 3) &= m(1)m(2) + m(2)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{36} , \\
P(S_2 = 4) &= m(1)m(3) + m(2)m(2) + m(3)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{3}{36} .
\end{aligned}
$$

Continuing in this way we would find $P(S_2 = 5) = 4/36$, $P(S_2 = 6) = 5/36$, $P(S_2 = 7) = 6/36$, $P(S_2 = 8) = 5/36$, $P(S_2 = 9) = 4/36$, $P(S_2 = 10) = 3/36$, $P(S_2 = 11) = 2/36$, and $P(S_2 = 12) = 1/36$.

The distribution for S_3 would then be the convolution of the distribution for S_2 with the distribution for X_3. Thus

$$P(S_3 = 3) = P(S_2 = 2)P(X_3 = 1)$$

$$
\begin{aligned}
&= \quad \frac{1}{36} \cdot \frac{1}{6} = \frac{1}{216} \;, \\
P(S_3 = 4) &= \quad P(S_2 = 3)P(X_3 = 1) + P(S_2 = 2)P(X_3 = 2) \\
&= \quad \frac{2}{36} \cdot \frac{1}{6} + \frac{1}{36} \cdot \frac{1}{6} = \frac{3}{216} \;,
\end{aligned}
$$

and so forth.

This is clearly a tedious job, and a program should be written to carry out this calculation. To do this we first write a program to form the convolution of two densities p and q and return the density r. We can then write a program to find the density for the sum S_n of n independent random variables with a common density p, at least in the case that the random variables have a finite number of possible values.

Running this program for the example of rolling a die n times for $n = 10, 20, 30$ results in the distributions shown in Figure 7.1. We see that, as in the case of Bernoulli trials, the distributions become bell-shaped. We shall discuss in Chapter 9 a very general theorem called the *Central Limit Theorem* that will explain this phenomenon. □

Example 7.2 A well-known method for evaluating a bridge hand is: an ace is assigned a value of 4, a king 3, a queen 2, and a jack 1. All other cards are assigned a value of 0. The *point count* of the hand is then the sum of the values of the cards in the hand. (It is actually more complicated than this, taking into account voids in suits, and so forth, but we consider here this simplified form of the point count.) If a card is dealt at random to a player, then the point count for this card has distribution

$$
p_X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 36/52 & 4/52 & 4/52 & 4/52 & 4/52 \end{pmatrix} .
$$

Let us regard the total hand of 13 cards as 13 independent trials with this common distribution. (Again this is not quite correct because we assume here that we are always choosing a card from a full deck.) Then the distribution for the point count C for the hand can be found from the program **NFoldConvolution** by using the distribution for a single card and choosing $n = 13$. A player with a point count of 13 or more is said to have an *opening bid.* The probability of having an opening bid is then

$$
P(C \geq 13) \;.
$$

Since we have the distribution of C, it is easy to compute this probability. Doing this we find that

$$
P(C \geq 13) = .2845 \;,
$$

so that about one in four hands should be an opening bid according to this simplified model. A more realistic discussion of this problem can be found in Epstein, *The Theory of Gambling and Statistical Logic.*[1] □

---

[1]R. A. Epstein, *The Theory of Gambling and Statistical Logic,* rev. ed. (New York: Academic Press, 1977).
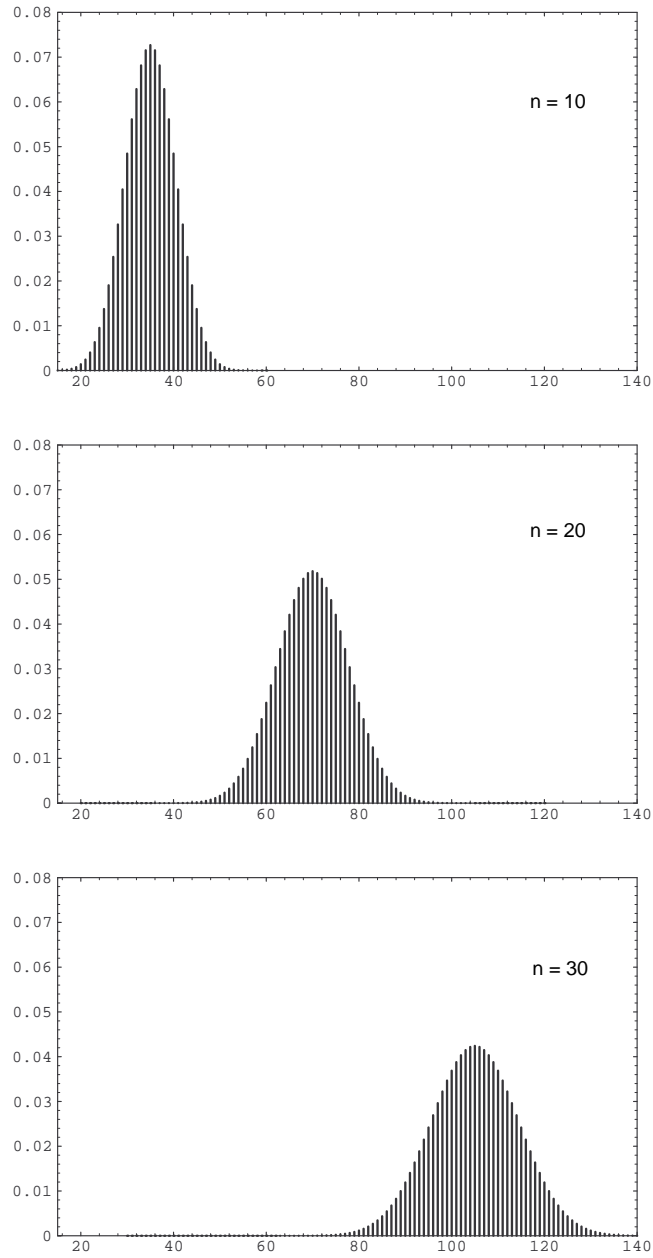
Figure 7.1: Density of S_n for rolling a die n times.

For certain special distributions it is possible to find an expression for the distribution that results from convoluting the distribution with itself n times.

The convolution of two binomial distributions, one with parameters m and p and the other with parameters n and p, is a binomial distribution with parameters $(m+n)$ and p. This fact follows easily from a consideration of the experiment which consists of first tossing a coin m times, and then tossing it n more times.

The convolution of k geometric distributions with common parameter p is a negative binomial distribution with parameters p and k. This can be seen by considering the experiment which consists of tossing a coin until the kth head appears.

Exercises

1 A die is rolled three times. Find the probability that the sum of the outcomes is

(a) greater than 9.

(b) an odd number.

2 The price of a stock on a given trading day changes according to the distribution

$$p_X = \begin{pmatrix} -1 & 0 & 1 & 2 \\ 1/4 & 1/2 & 1/8 & 1/8 \end{pmatrix}.$$

Find the distribution for the change in stock price after two (independent) trading days.

3 Let X_1 and X_2 be independent random variables with common distribution

$$p_X = \begin{pmatrix} 0 & 1 & 2 \\ 1/8 & 3/8 & 1/2 \end{pmatrix}.$$

Find the distribution of the sum $X_1 + X_2$.

4 In one play of a certain game you win an amount X with distribution

$$p_X = \begin{pmatrix} 1 & 2 & 3 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

Using the program **NFoldConvolution** find the distribution for your total winnings after ten (independent) plays. Plot this distribution.

5 Consider the following two experiments: the first has outcome X taking on the values 0, 1, and 2 with equal probabilities; the second results in an (independent) outcome Y taking on the value 3 with probability 1/4 and 4 with probability 3/4. Find the distribution of

(a) $Y + X$.

(b) $Y - X$.

6 People arrive at a queue according to the following scheme: During each minute of time either 0 or 1 person arrives. The probability that 1 person arrives is p and that no person arrives is $q = 1 - p$. Let C_r be the number of customers arriving in the first r minutes. Consider a Bernoulli trials process with a success if a person arrives in a unit time and failure if no person arrives in a unit time. Let T_r be the number of failures before the rth success.

(a) What is the distribution for T_r?

(b) What is the distribution for C_r?

(c) Find the mean and variance for the number of customers arriving in the first r minutes.

7 (a) A die is rolled three times with outcomes X_1, X_2, and X_3. Let Y_3 be the maximum of the values obtained. Show that

$$P(Y_3 \leq j) = P(X_1 \leq j)^3 \ .$$

Use this to find the distribution of Y_3. Does Y_3 have a bell-shaped distribution?

(b) Now let Y_n be the maximum value when n dice are rolled. Find the distribution of Y_n. Is this distribution bell-shaped for large values of n?

8 A baseball player is to play in the World Series. Based upon his season play, you estimate that if he comes to bat four times in a game the number of hits he will get has a distribution

$$p_X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ .4 & .2 & .2 & .1 & .1 \end{pmatrix}.$$

Assume that the player comes to bat four times in each game of the series.

(a) Let X denote the number of hits that he gets in a series. Using the program **NFoldConvolution**, find the distribution of X for each of the possible series lengths: four-game, five-game, six-game, seven-game.

(b) Using one of the distribution found in part (a), find the probability that his batting average exceeds .400 in a four-game series. (The batting average is the number of hits divided by the number of times at bat.)

(c) Given the distribution p_X, what is his long-term batting average?

9 Prove that you cannot load two dice in such a way that the probabilities for any sum from 2 to 12 are the same. (Be sure to consider the case where one or more sides turn up with probability zero.)

10 (Lévy[2]) Assume that n is an integer, not prime. Show that you can find two distributions a and b on the nonnegative integers such that the convolution of

[2]See M. Krasner and B. Ranulae, "Sur une Proprieté des Polynomes de la Division du Circle"; and the following note by J. Hadamard, in *C. R. Acad. Sci.,* vol. 204 (1937), pp. 397–399.

a and b is the equiprobable distribution on the set 0, 1, 2, ..., $n-1$. If n is prime this is not possible, but the proof is not so easy. (Assume that neither a nor b is concentrated at 0.)

11 Assume that you are playing craps with dice that are loaded in the following way: faces two, three, four, and five all come up with the same probability $(1/6) + r$. Faces one and six come up with probability $(1/6) - 2r$, with $0 < r < .02$. Write a computer program to find the probability of winning at craps with these dice, and using your program find which values of r make craps a favorable game for the player with these dice.

7.2 Sums of Continuous Random Variables

In this section we consider the continuous version of the problem posed in the previous section: How are sums of independent random variables distributed?

Convolutions

Definition 7.2 Let X and Y be two continuous random variables with density functions $f(x)$ and $g(y)$, respectively. Assume that both $f(x)$ and $g(y)$ are defined for all real numbers. Then the *convolution* $f * g$ of f and g is the function given by

$$
\begin{aligned}
(f * g)(z) &= \int_{-\infty}^{+\infty} f(z - y)g(y)\, dy \\
&= \int_{-\infty}^{+\infty} g(z - x)f(x)\, dx \ .
\end{aligned}
$$

$\square$

This definition is analogous to the definition, given in Section 7.1, of the convolution of two distribution functions. Thus it should not be surprising that if X and Y are independent, then the density of their sum is the convolution of their densities. This fact is stated as a theorem below, and its proof is left as an exercise (see Exercise 1).

Theorem 7.1 Let X and Y be two independent random variables with density functions $f_X(x)$ and $f_Y(y)$ defined for all x. Then the sum $Z = X + Y$ is a random variable with density function $f_Z(z)$, where f_Z is the convolution of f_X and f_Y. $\square$

To get a better understanding of this important result, we will look at some examples.

Sum of Two Independent Uniform Random Variables

Example 7.3 Suppose we choose independently two numbers at random from the interval $[0, 1]$ with uniform probability density. What is the density of their sum?

Let X and Y be random variables describing our choices and $Z = X + Y$ their sum. Then we have

$$f_X(x) = f_Y(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise;} \end{cases}$$

and the density function for the sum is given by

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y) \, dy \ .$$

Since $f_Y(y) = 1$ if $0 \le y \le 1$ and 0 otherwise, this becomes

$$f_Z(z) = \int_0^1 f_X(z - y) \, dy \ .$$

Now the integrand is 0 unless $0 \le z - y \le 1$ (i.e., unless $z - 1 \le y \le z$) and then it is 1. So if $0 \le z \le 1$, we have

$$f_Z(z) = \int_0^z dy = z \ ,$$

while if $1 < z \le 2$, we have

$$f_Z(z) = \int_{z-1}^1 dy = 2 - z \ ,$$

and if $z < 0$ or $z > 2$ we have $f_Z(z) = 0$ (see Figure 7.2). Hence,

$$f_Z(z) = \begin{cases} z, & \text{if } 0 \le z \le 1, \\ 2 - z, & \text{if } 1 < z \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Note that this result agrees with that of Example 2.4.   □

Sum of Two Independent Exponential Random Variables

Example 7.4 Suppose we choose two numbers at random from the interval $[0, \infty)$ with an *exponential* density with parameter λ. What is the density of their sum?

Let X, Y, and $Z = X + Y$ denote the relevant random variables, and f_X, f_Y, and f_Z their densities. Then

$$f_X(x) = f_Y(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \ge 0, \\ 0, & \text{otherwise;} \end{cases}$$
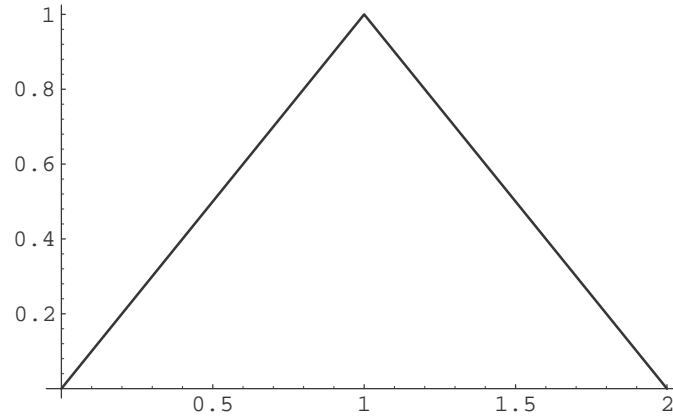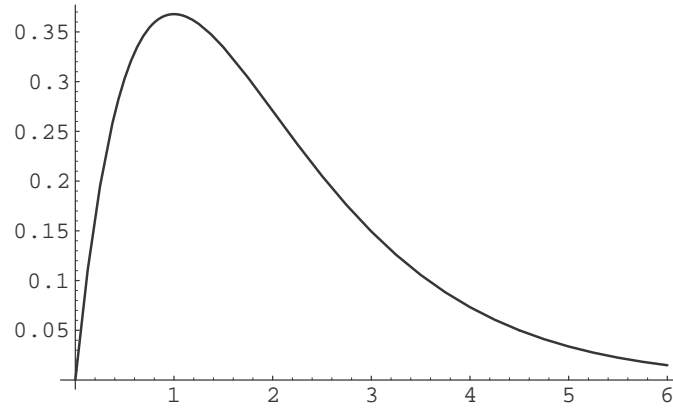
Figure 7.2: Convolution of two uniform densities.

Figure 7.3: Convolution of two exponential densities with $\lambda = 1$.

and so, if $z > 0$,

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) \, dy \\
&= \int_0^z \lambda e^{-\lambda(z-y)} \lambda e^{-\lambda y} \, dy \\
&= \int_0^z \lambda^2 e^{-\lambda z} \, dy \\
&= \lambda^2 z e^{-\lambda z},
\end{aligned}
$$

while if $z < 0$, $f_Z(z) = 0$ (see Figure 7.3). Hence,

$$
f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z}, & \text{if } z \geq 0, \\ 0, & \text{otherwise.} \end{cases}
$$

$\square$

Sum of Two Independent Normal Random Variables

Example 7.5 It is an interesting and important fact that the convolution of two normal densities with means μ_1 and μ_2 and variances σ_1 and σ_2 is again a normal density, with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$. We will show this in the special case that both random variables are standard normal. The general case can be done in the same way, but the calculation is messier. Another way to show the general result is given in Example 10.17.

Suppose X and Y are two independent random variables, each with the standard *normal* density (see Example 5.8). We have

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

and so

$$
\begin{aligned}
f_Z(z) &= f_X * f_Y(z) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-(z-y)^2/2} e^{-y^2/2} \, dy \\
&= \frac{1}{2\pi} e^{-z^2/4} \int_{-\infty}^{+\infty} e^{-(y-z/2)^2} \, dy \\
&= \frac{1}{2\pi} e^{-z^2/4} \sqrt{\pi} \left[\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(y-z/2)^2} \, dy \right] \ .
\end{aligned}
$$

The expression in the brackets equals 1, since it is the integral of the normal density function with $\mu = 0$ and $\sigma = \sqrt{2}$. So, we have

$$f_Z(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4} \ .$$

$\square$

Sum of Two Independent Cauchy Random Variables

Example 7.6 Choose two numbers at random from the interval $(-\infty, +\infty)$ with the Cauchy density with parameter $a = 1$ (see Example 5.10). Then

$$f_X(x) = f_Y(x) = \frac{1}{\pi(1+x^2)} \ ,$$

and $Z = X + Y$ has density

$$f_Z(z) = \frac{1}{\pi^2} \int_{-\infty}^{+\infty} \frac{1}{1+(z-y)^2} \frac{1}{1+y^2} \, dy \ .$$

This integral requires some effort, and we give here only the result (see Section 10.3, or Dwass[3]):

$$f_Z(z) = \frac{2}{\pi(4 + z^2)} \ .$$

Now, suppose that we ask for the density function of the *average*

$$A = (1/2)(X + Y)$$

of X and Y. Then $A = (1/2)Z$. Exercise 5.2.19 shows that if U and V are two continuous random variables with density functions $f_U(x)$ and $f_V(x)$, respectively, and if $V = aU$, then

$$f_V(x) = \left(\frac{1}{a}\right) f_U\left(\frac{x}{a}\right) \ .$$

Thus, we have

$$f_A(z) = 2f_Z(2z) = \frac{1}{\pi(1 + z^2)} \ .$$

Hence, the density function for the average of two random variables, each having a Cauchy density, is again a random variable with a Cauchy density; this remarkable property is a peculiarity of the Cauchy density. One consequence of this is if the error in a certain measurement process had a Cauchy density and you averaged a number of measurements, the average could not be expected to be any more accurate than any one of your individual measurements! $\qquad \square$

Rayleigh Density

Example 7.7 Suppose X and Y are two independent standard normal random variables. Now suppose we locate a point P in the xy-plane with coordinates (X, Y) and ask: What is the density of the square of the distance of P from the origin? (We have already simulated this problem in Example 5.9.) Here, with the preceding notation, we have

$$f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \ .$$

Moreover, if X^2 denotes the square of X, then (see Theorem 5.1 and the discussion following)

$$f_{X^2}(r) = \begin{cases} \frac{1}{2\sqrt{r}}(f_X(\sqrt{r}) + f_X(-\sqrt{r})) & \text{if } r > 0, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \frac{1}{\sqrt{2\pi r}}(e^{-r/2}) & \text{if } r > 0, \\ 0 & \text{otherwise.} \end{cases}$$

---

[3]M. Dwass, "On the Convolution of Cauchy Distributions," *American Mathematical Monthly*, vol. 92, no. 1, (1985), pp. 55–57; see also R. Nelson, letters to the Editor, ibid., p. 679.

This is a gamma density with $\lambda = 1/2$, $\beta = 1/2$ (see Example 7.4). Now let $R^2 = X^2 + Y^2$. Then

$$
\begin{aligned}
f_{R^2}(r) &= \int_{-\infty}^{+\infty} f_{X^2}(r-s) f_{Y^2}(s) \, ds \\
&= \frac{1}{4\pi} \int_{-\infty}^{+\infty} e^{-(r-s)/2} \frac{r-s}{2}^{-1/2} e^{-s} \frac{s}{2}^{-1/2} \, ds \; , \\
&= \begin{cases} \frac{1}{2} e^{-r^2/2}, & \text{if } r \geq 0, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

Hence, R^2 has a gamma density with $\lambda = 1/2$, $\beta = 1$. We can interpret this result as giving the density for the square of the distance of P from the center of a target if its coordinates are normally distributed.

The density of the random variable R is obtained from that of R^2 in the usual way (see Theorem 5.1), and we find

$$
f_R(r) = \begin{cases} \frac{1}{2} e^{-r^2/2} \cdot 2r = r e^{-r^2/2}, & \text{if } r \geq 0, \\ 0, & \text{otherwise.} \end{cases}
$$

Physicists will recognize this as a Rayleigh density. Our result here agrees with our simulation in Example 5.9.   □

Chi-Squared Density

More generally, the same method shows that the sum of the squares of n independent normally distributed random variables with mean 0 and standard deviation 1 has a gamma density with $\lambda = 1/2$ and $\beta = n/2$. Such a density is called a *chi-squared density* with n degrees of freedom. This density was introduced in Chapter 4.3. In Example 5.10, we used this density to test the hypothesis that two traits were independent.

Another important use of the chi-squared density is in comparing experimental data with a theoretical discrete distribution, to see whether the data supports the theoretical model. More specifically, suppose that we have an experiment with a finite set of outcomes. If the set of outcomes is countable, we group them into finitely many sets of outcomes. We propose a theoretical distribution which we think will model the experiment well. We obtain some data by repeating the experiment a number of times. Now we wish to check how well the theoretical distribution fits the data.

Let X be the random variable which represents a theoretical outcome in the model of the experiment, and let $m(x)$ be the distribution function of X. In a manner similar to what was done in Example 5.10, we calculate the value of the expression

$$
V = \sum_x \frac{(o_x - n \cdot m(x))^2}{n \cdot m(x)} \; ,
$$

where the sum runs over all possible outcomes x, n is the number of data points, and o_x denotes the number of outcomes of type x observed in the data. Then

Outcome	Observed Frequency
1	15
2	8
3	7
4	5
5	7
6	18

Table 7.1: Observed data.

for moderate or large values of n, the quantity V is approximately chi-squared distributed, with $\nu-1$ degrees of freedom, where ν represents the number of possible outcomes. The proof of this is beyond the scope of this book, but we will illustrate the reasonableness of this statement in the next example. If the value of V is very large, when compared with the appropriate chi-squared density function, then we would tend to reject the hypothesis that the model is an appropriate one for the experiment at hand. We now give an example of this procedure.

Example 7.8 Suppose we are given a single die. We wish to test the hypothesis that the die is fair. Thus, our theoretical distribution is the uniform distribution on the integers between 1 and 6. So, if we roll the die n times, the expected number of data points of each type is $n/6$. Thus, if o_i denotes the actual number of data points of type i, for $1 \leq i \leq 6$, then the expression

$$V = \sum_{i=1}^{6} \frac{(o_i - n/6)^2}{n/6}$$

is approximately chi-squared distributed with 5 degrees of freedom.

Now suppose that we actually roll the die 60 times and obtain the data in Table 7.1. If we calculate V for this data, we obtain the value 13.6. The graph of the chi-squared density with 5 degrees of freedom is shown in Figure 7.4. One sees that values as large as 13.6 are rarely taken on by V if the die is fair, so we would reject the hypothesis that the die is fair. (When using this test, a statistician will reject the hypothesis if the data gives a value of V which is larger than 95% of the values one would expect to obtain if the hypothesis is true.)

In Figure 7.5, we show the results of rolling a die 60 times, then calculating V, and then repeating this experiment 1000 times. The program that performs these calculations is called **DieTest**. We have superimposed the chi-squared density with 5 degrees of freedom; one can see that the data values fit the curve fairly well, which supports the statement that the chi-squared density is the correct one to use. □

So far we have looked at several important special cases for which the convolution integral can be evaluated explicitly. In general, the convolution of two continuous densities cannot be evaluated explicitly, and we must resort to numerical methods. Fortunately, these prove to be remarkably effective, at least for bounded densities.
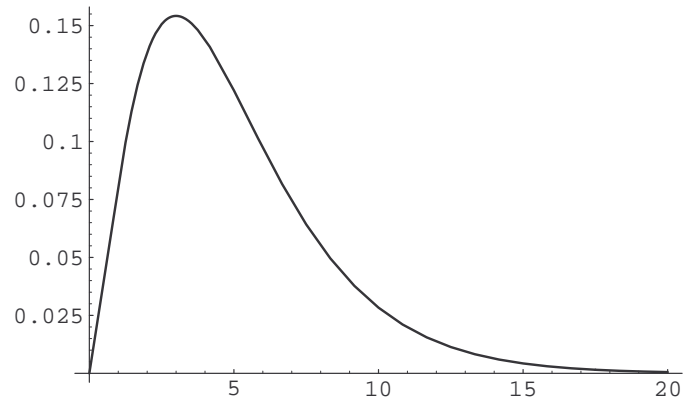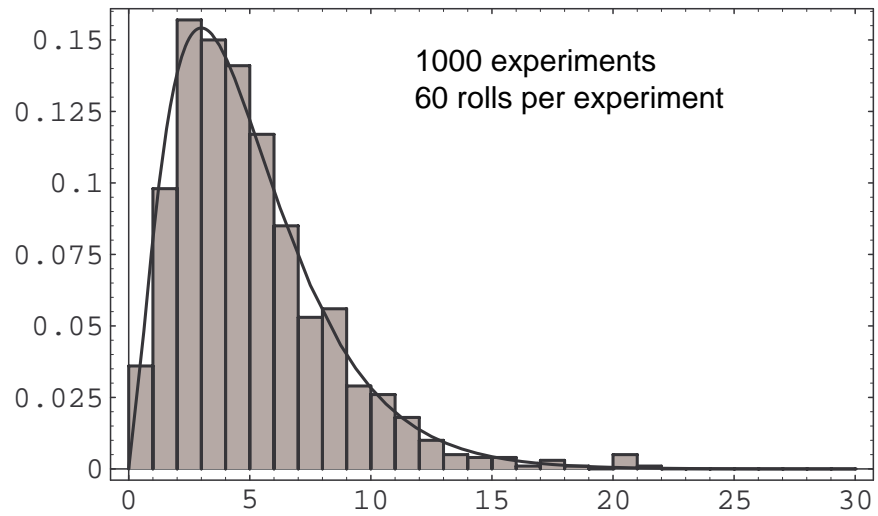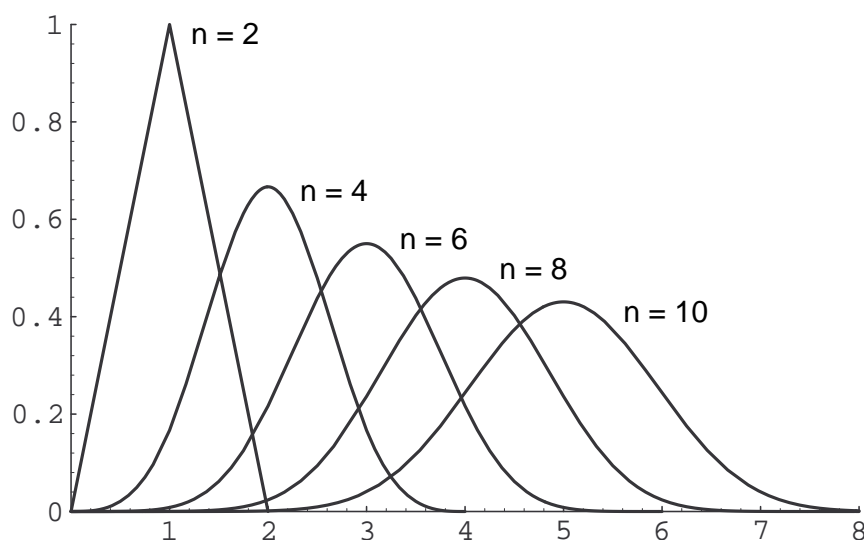
Figure 7.4: Chi-squared density with 5 degrees of freedom.

Figure 7.5: Rolling a fair die.

Figure 7.6: Convolution of n uniform densities.

Independent Trials

We now consider briefly the distribution of the sum of n independent random variables, all having the same density function. If X_1, X_2, ..., X_n are these random variables and $S_n = X_1 + X_2 + \cdots + X_n$ is their sum, then we will have

$$f_{S_n}(x) = (f_{X_1} * f_{X_2} * \cdots * f_{X_n})(x) \ ,$$

where the right-hand side is an n-fold convolution. It is possible to calculate this density for general values of n in certain simple cases.

Example 7.9 Suppose the X_i are uniformly distributed on the interval $[0, 1]$. Then

$$f_{X_i}(x) = \begin{cases} 1, & \text{if } 0 \leq x \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

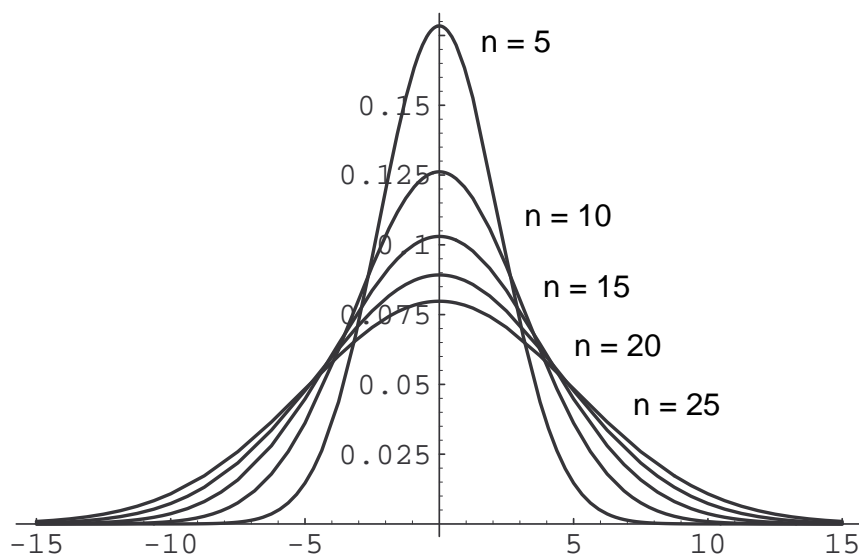and $f_{S_n}(x)$ is given by the formula[4]

$$f_{S_n}(x) = \begin{cases} \frac{1}{(n-1)!} \sum_{0 \leq j \leq x} (-1)^j \binom{n}{j} (x-j)^{n-1}, & \text{if } 0 < x < n, \\ 0, & \text{otherwise.} \end{cases}$$

The density $f_{S_n}(x)$ for $n = 2$, 4, 6, 8, 10 is shown in Figure 7.6.

If the X_i are distributed normally, with mean 0 and variance 1, then (cf. Example 7.5)

$$f_{X_i}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \ ,$$

---

[4]J. B. Uspensky, *Introduction to Mathematical Probability* (New York: McGraw-Hill, 1937), p. 277.

Figure 7.7: Convolution of n standard normal densities.

and
$$f_{S_n}(x) = \frac{1}{\sqrt{2\pi n}} e^{-x^2/2n} \ .$$

Here the density f_{S_n} for $n = 5, 10, 15, 20, 25$ is shown in Figure 7.7.

If the X_i are all exponentially distributed, with mean $1/\lambda$, then
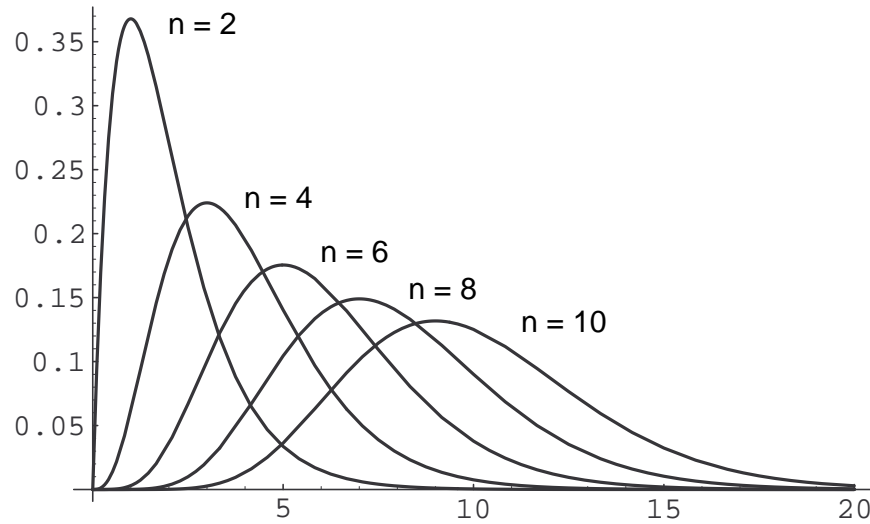
$$f_{X_i}(x) = \lambda e^{-\lambda x} \ ,$$

and

$$f_{S_n}(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!} \ .$$

In this case the density f_{S_n} for $n = 2, 4, 6, 8, 10$ is shown in Figure 7.8.   □

Exercises

1 Let X and Y be independent real-valued random variables with density func-
tions $f_X(x)$ and $f_Y(y)$, respectively. Show that the density function of the
sum $X + Y$ is the convolution of the functions $f_X(x)$ and $f_Y(y)$. *Hint*: Let $\bar{X}$
be the joint random variable (X, Y). Then the joint density function of $\bar{X}$ is
$f_X(x)f_Y(y)$, since X and Y are independent. Now compute the probability
that $X + Y \leq z$, by integrating the joint density function over the appropriate
region in the plane. This gives the cumulative distribution function of Z. Now
differentiate this function with respect to z to obtain the density function of
z.

2 Let X and Y be independent random variables defined on the space Ω, with
density functions f_X and f_Y, respectively. Suppose that $Z = X + Y$. Find
the density f_Z of Z if

Figure 7.8: Convolution of n exponential densities with $\lambda = 1$.

(a)
$$f_X(x) = f_Y(x) = \begin{cases} 1/2, & \text{if } -1 \leq x \leq +1, \\ 0, & \text{otherwise.} \end{cases}$$

(b)
$$f_X(x) = f_Y(x) = \begin{cases} 1/2, & \text{if } 3 \leq x \leq 5, \\ 0, & \text{otherwise.} \end{cases}$$

(c)
$$f_X(x) = \begin{cases} 1/2, & \text{if } -1 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_Y(x) = \begin{cases} 1/2, & \text{if } 3 \leq x \leq 5, \\ 0, & \text{otherwise.} \end{cases}$$

(d) What can you say about the set $E = \{ z : f_Z(z) > 0 \}$ in each case?

3 Suppose again that $Z = X + Y$. Find f_Z if

(a)
$$f_X(x) = f_Y(x) = \begin{cases} x/2, & \text{if } 0 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(b)
$$f_X(x) = f_Y(x) = \begin{cases} (1/2)(x - 3), & \text{if } 3 < x < 5, \\ 0, & \text{otherwise.} \end{cases}$$

(c)
$$f_X(x) = \begin{cases} 1/2, & \text{if } 0 < x < 2, \\ 0, & \text{otherwise,} \end{cases}$$

$$f_Y(x) = \begin{cases} x/2, & \text{if } 0 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(d) What can you say about the set $E = \{\, z : f_Z(z) > 0 \,\}$ in each case?

4 Let X, Y, and Z be independent random variables with

$$f_X(x) = f_Y(x) = f_Z(x) = \begin{cases} 1, & \text{if } 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that $W = X + Y + Z$. Find f_W directly, and compare your answer with that given by the formula in Example 7.9. *Hint*: See Example 7.3.

5 Suppose that X and Y are independent and $Z = X + Y$. Find f_Z if

(a)
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_Y(x) = \begin{cases} \mu e^{-\mu x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(b)
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_Y(x) = \begin{cases} 1, & \text{if } 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

6 Suppose again that $Z = X + Y$. Find f_Z if

$$f_X(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_Y(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(x-\mu_2)^2/2\sigma_2^2} \ .$$

*7 Suppose that $R^2 = X^2 + Y^2$. Find f_{R^2} and f_R if

$$f_X(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_Y(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(x-\mu_2)^2/2\sigma_2^2} \ .$$

8 Suppose that $R^2 = X^2 + Y^2$. Find f_{R^2} and f_R if

$$f_X(x) = f_Y(x) = \begin{cases} 1/2, & \text{if } -1 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

9 Assume that the service time for a customer at a bank is exponentially distributed with mean service time 2 minutes. Let X be the total service time for 10 customers. Estimate the probability that $X > 22$ minutes.

10 Let X_1, X_2, ..., X_n be n independent random variables each of which has an exponential density with mean μ. Let M be the *minimum* value of the X_j. Show that the density for M is exponential with mean μ/n. *Hint*: Use cumulative distribution functions.

11 A company buys 100 lightbulbs, each of which has an exponential lifetime of 1000 hours. What is the expected time for the first of these bulbs to burn out? (See Exercise 10.)

12 An insurance company assumes that the time between claims from each of its homeowners' policies is exponentially distributed with mean μ. It would like to estimate μ by averaging the times for a number of policies, but this is not very practical since the time between claims is about 30 years. At Galambos'[5] suggestion the company puts its customers in groups of 50 and observes the time of the first claim within each group. Show that this provides a practical way to estimate the value of μ.

13 Particles are subject to collisions that cause them to split into two parts with each part a fraction of the parent. Suppose that this fraction is uniformly distributed between 0 and 1. Following a single particle through several splittings we obtain a fraction of the original particle $Z_n = X_1 \cdot X_2 \cdot \ldots \cdot X_n$ where each X_j is uniformly distributed between 0 and 1. Show that the density for the random variable Z_n is

$$f_n(z) = \frac{1}{(n-1)!}(-\log z)^{n-1}.$$

Hint: Show that $Y_k = -\log X_k$ is exponentially distributed. Use this to find the density function for $S_n = Y_1 + Y_2 + \cdots + Y_n$, and from this the cumulative distribution and density of $Z_n = e^{-S_n}$.

14 Assume that X_1 and X_2 are independent random variables, each having an exponential density with parameter λ. Show that $Z = X_1 - X_2$ has density

$$f_Z(z) = (1/2)\lambda e^{-\lambda|z|} .$$

15 Suppose we want to test a coin for fairness. We flip the coin n times and record the number of times X_0 that the coin turns up tails and the number of times $X_1 = n - X_0$ that the coin turns up heads. Now we set

$$Z = \sum_{i=0}^{1} \frac{(X_i - n/2)^2}{n/2} .$$

Then for a fair coin Z has approximately a chi-squared distribution with $2 - 1 = 1$ degree of freedom. Verify this by computer simulation first for a fair coin ($p = 1/2$) and then for a biased coin ($p = 1/3$).

---

[5] J. Galambos, *Introductory Probability Theory* (New York: Marcel Dekker, 1984), p. 159.

16 Verify your answers in Exercise 2(a) by computer simulation: Choose X and Y from $[-1, 1]$ with uniform density and calculate $Z = X + Y$. Repeat this experiment 500 times, recording the outcomes in a bar graph on $[-2, 2]$ with 40 bars. Does the density f_Z calculated in Exercise 2(a) describe the shape of your bar graph? Try this for Exercises 2(b) and Exercise 2(c), too.

17 Verify your answers to Exercise 3 by computer simulation.

18 Verify your answer to Exercise 4 by computer simulation.

19 The *support* of a function $f(x)$ is defined to be the set

$$\{x \ : \ f(x) > 0\} .$$

Suppose that X and Y are two continuous random variables with density functions $f_X(x)$ and $f_Y(y)$, respectively, and suppose that the supports of these density functions are the intervals $[a, b]$ and $[c, d]$, respectively. Find the support of the density function of the random variable $X + Y$.

20 Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables, all having a common density function f_X with support $[a, b]$ (see Exercise 19). Let $S_n = X_1 + X_2 + \cdots + X_n$, with density function f_{S_n}. Show that the support of f_{S_n} is the interval $[na, nb]$. *Hint*: Write $f_{S_n} = f_{S_{n-1}} * f_X$. Now use Exercise 19 to establish the desired result by induction.

21 Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables, all having a common density function f_X. Let $A = S_n/n$ be their average. Find f_A if

(a) $f_X(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ (normal density).

(b) $f_X(x) = e^{-x}$ (exponential density).

   Hint: Write $f_A(x)$ in terms of $f_{S_n}(x)$.