Continuous Random Variables

P. Sam Johnson

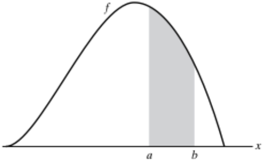
Introduction

There exist random variables whose set of possible values is uncountable. Two examples are the time that a train arrives at a specified stop and the lifetime of a transistor.

Let X be such a random variable. We say that X is a *continuous* random variable if there exists a nonnegative function f, defined for all real $x \in (-\infty, \infty)$, having the property that, for any set B of real numbers,

$$P\{X \in B\} = \int_B f(x)\, dx. \tag{1}$$

The function f is called the *probability density function* of the random variable X.



$P(a \leq X \leq b)$ = area of shaded region

Introduction

In words, Equation (1) states that the probability that X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, f must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)\, dx.$$

All probability statements about X can be answered in terms of f. For instance, from Equation (1), letting $B = [a, b]$, we obtain

$$P\{a \leq X \leq b\} = \int_{a}^{b} f(x)\, dx. \tag{2}$$

If we let $a = b$ in Equation (2), we get

$$P\{X = a\} = \int_a^a f(x)\, dx = 0.$$

In words, this equation states that the probability that a continuous random variable will assume any fixed value is zero. Hence, for a continuous random variable,

$$P\{X < a\} = P\{X \le a\} = F(a) = \int_{-\infty}^a f(x)\, dx.$$

Example 1.

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \textit{otherwise.} \end{cases}$$

(a) *What is the value of C?*

(b) *Find $P\{X > 1\}$.*

Solution :

(a) Since f is a probability density function, we must have $\int_{-\infty}^{\infty} f(x)dx = 1$, implying that $C\int_0^2 (4x - 2x^2)dx = 1$. hence $C = \frac{3}{8}$.

(b) $P\{X > 1\} = \int_1^{\infty} f(x)dx = \frac{3}{8}\int_1^2 (4x - 2x^2)dx = \frac{1}{2}$.

Example 2.

The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

What is the probability that

(a) *a computer will function between* 50 *and* 150 *hours before breaking down?*

(b) *it will function for fewer than* 100 *hours?*

Solution

(a) Since

$$1 = \int_{-\infty}^{\infty} f(x)dx = \lambda \int_{0}^{\infty} e^{-x/100} dx$$

we obtain

$$1 = -\lambda(100)e^{-x/100}\Big|_0^{\infty} = 100\lambda \quad \text{or} \quad \lambda = \frac{1}{100}.$$

Hence, the probability that a computer will function between 50 and 150 hours before breaking down is given by

$$P\{50 < X < 150\} = \int_{50}^{150} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\Big|_{50}^{150}$$
$$= e^{-1/2} - e^{-3/2} \approx .384.$$

(b) Similarly,

$$P\{X < 100\} = \int_{0}^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\Big|_{0}^{100} = 1 - e^{-1} \approx .633.$$

In other words, approximately 63.3 percent of the time, a computer will fail before registering 100 hours of use.

Example 3.

The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100 \\ \frac{100}{x^2} & x > 100. \end{cases}$$

What is the probability that exactly 2 *of* 5 *such tubes in a radio set will have to be replaced within the first* 150 *hours of operation? Assume that the events* E_i*,* $i = 1, 2, 3, 4, 5$*, that the* i*th such tube will have to be replaced within this time are independent.*

Solution: From the statement of the problem, we have

$$P(E_i) = \int_0^{150} f(x)dx = 100 \int_{100}^{150} x^{-2} dx = \frac{1}{3}.$$

Hence, from the independence of the events E_i, it follows that the desired probability is

$$\binom{5}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = \frac{80}{243}.$$

Relationship between the cumulative distribution F and the probability density f

The relationship between the cumulative distribution F and the probability density f is expressed by

$$F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)\, dx.$$

Differentiating both sides of the preceding equation yields

$$\frac{d}{da} F(a) = f(a).$$

That is, the density is the derivative of the cumulative distribution function. A somewhat more intuitive interpretation of the density function may be obtained from Equation (2) as follows:

$$P\left\{a - \frac{\varepsilon}{2} \leq X \leq a + \frac{\varepsilon}{2}\right\} = \int_{a-\varepsilon/2}^{a+\varepsilon/2} f(x)\, dx \approx \varepsilon f(a)$$

when ε is small and when $f(\cdot)$ is continuous at $x = a$. In other words, the probability that X will be contained in an interval of length ε around the point a is approximately $\varepsilon f(a)$. From this result we see that $f(a)$ is a measure of how likely it is that the random variable will be near a.

Example 4.

If X is continuous with distribution function F_X and density function f_X, find the density function of $Y = 2X$.

Solution: We will determine f_Y in two ways. The first way is to derive, and then differentiate, the distribution function of Y:

$$F_Y(a) = P\{Y \leq a\} = P\{2X \leq a\} = P\{X \leq a/2\} = F_X(a/2).$$

Differentiation gives

$$f_Y(a) = \frac{1}{2} f_X(a/2).$$

Another way to determine f_Y is to note that

$$\begin{aligned} \varepsilon f_Y(a) &\approx P\{a - \frac{\varepsilon}{2} \leq Y \leq a + \frac{\varepsilon}{2}\} = P\{a - \frac{\varepsilon}{2} \leq 2X \leq a + \frac{\varepsilon}{2}\} \\ &= P\{\frac{a}{2} - \frac{\varepsilon}{4} \leq X \leq \frac{a}{2} + \frac{\varepsilon}{4}\} \approx \frac{\varepsilon}{2} f_X(a/2). \end{aligned}$$

Dividing through by ε gives the same result as before.

Expectation and Variance of Continuous Random Variables

If X is a continuous random variable having probability density function $f(x)$, then, because

$$f(x)dx \approx P\{x \leq X \leq x + dx\} \text{ for } dx \text{ small}$$

it is easy to see that the analogous definition is to define the expected value of X by

$$E[X] = \int_{-\infty}^{\infty} xf(x)\, dx.$$

Example 5.

Find $E[X]$ when the density function of X is

$$f(x) = \begin{cases} 2x & \text{if} \quad 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Solution:

$$E[X] = \int_0^1 xf(x)dx = \int_0^1 2x^2 dx = \frac{2}{3}.$$

Example 6.

The density function of X is given by $f(x) = \begin{cases} 1 & \text{if } \; 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$ *Find* $E[e^X]$.

Solution: Let $Y = e^X$. We start by determining F_Y, the probability distribution function of Y. Now, for $1 \leq x \leq e$,

$$\begin{aligned} F_Y(x) &= P\{Y \leq x\} = P\{e^X \leq x\} = P\{X \leq \log(x)\} \\ &= \int_0^{\log(x)} f(y)dy = \log(x). \end{aligned}$$

By differentiating $F_Y(x)$, we can conclude that the probability density function of Y is given by

$$f_Y(x) = \frac{1}{x}, \quad 1 \leq x \leq e.$$

Hence,

$$E[e^X] = E[Y] = \int_{-\infty}^{\infty} x f_Y(x)dx = \int_1^e dx = e - 1.$$

Proposition 7.

If X is a continuous random variable with probability density function $f(x)$, then, for any real-valued function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\ dx.$$

An application of Proposition (7) to Example (6) yields

$$\begin{aligned} E[e^X] &= \int_0^1 e^x dx \quad \text{since } f(x) = 1,\ 0 < x < 1 \\ &= e - 1 \end{aligned}$$

which is in accord with the result obtained in that example.

The proof of Proposition (7) is more involved than that of its discrete random variable analog. We will present such a proof under the provision that the random variable $g(X)$ is nonnegative.

Lemma

We will need the following lemma, which is of independent interest.

Lemma 8.

For a nonnegative random variable Y,

$$E[Y] = \int_0^\infty P\{Y > y\}\, dy.$$

Proof: We present a proof when Y is a continuous random variable with probability density function f_Y. We have

$$\int_0^\infty P\{Y > y\} dy = \int_0^\infty \int_y^\infty f_Y(x)\, dx dy$$

where we have used the fact that $P\{Y > y\} = \int_y^\infty f_Y(x)\, dx$. Interchanging the order of integration in the preceding equation yields

$$\begin{aligned} \int_0^\infty P\{Y > y\} dy &= \int_0^\infty \left(\int_0^x dy \right) f_Y(x)\, dx \\ &= \int_0^\infty x f_Y(x)\, dx \\ &= E[Y]. \end{aligned}$$

Proof of Proposition (7)

From Lemma (8), for any function g for which $g(x) \geq 0$,

$$
\begin{aligned}
E[g(X)] &= \int_0^\infty P\{g(X) > y\}\ dy \\
&= \int_0^\infty \int_{x:g(x)>y} f(x)\ dx\, dy \\
&= \int_{x:g(x)>0} \int_0^{g(x)} dy\ f(x)\ dx \\
&= \int_{x:g(x)>0} g(x) f(x)\ dx
\end{aligned}
$$

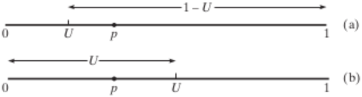
which completes the proof.

Example 9.

A stick of length 1 is split at a point U that is uniformly distributed over $(0, 1)$. Determine the expected length of the piece that contains the point p, $0 \leq p \leq 1$.

Solution: Let $L_p(U)$ denote the length of the substick that contains the point p, and note that

$$L_p(U) = \begin{cases} 1 - U & U < p \\ U & U > p. \end{cases}$$



Substick containing point p: (a) $U < p$; **(b)** $U > p$.

Hence, from Proposition (7),

$$E[L_p(U)] = \int_0^1 L_p(u)du = \int_0^p (1-u)du + \int_p^1 udu = \frac{1}{2} - \frac{(1-p)^2}{2} + \frac{1}{2} - \frac{p^2}{2} = \frac{1}{2} + p(1-p).$$

Since $p(1-p)$ is maximized when $p = \frac{1}{2}$, it is interesting to note that the expected length of the substick containing the point p is maximized when p is the midpoint of the original stick.

Example 10.

Suppose that if you are s minutes early for an appointment, then you incur the cost cs, and if you are s minutes late, then you incur the cost ks. Suppose also that the travel time from where you presently are to the location of your appointment is a continuous random variable having probability density function f. Determine the time at which you should depart if you want to minimize your expected cost.

Solution: Let X denote the travel time. If you leave t minutes before your appointment, then your cost, call it $C_t(X)$, is given by

$$C_t(X) = \begin{cases} c(t - X) & \text{if} \quad X \leq t \\ k(X - t) & \text{if} \quad X \geq t. \end{cases}$$

Solution (Contd...)

Therefore,

$$\begin{aligned} E[C_t(X)] &= \int_0^\infty C_t(x)f(x)dx \\ &= \int_0^t c(t-x)f(x)dx + \int_t^\infty k(x-t)f(x)dx \\ &= ct\int_0^t f(x)dx - c\int_0^t xf(x)dx + k\int_t^\infty xf(x)dx - kt\int_t^\infty f(x)dx. \end{aligned}$$

The value of t that minimizes $E[C_t(X)]$ can now be obtained by calculus. Differentiation yields

$$\begin{aligned} \frac{d}{dt}E[C_t(X)] &= ctf(t) + cF(t) - ctf(t) - ktf(t) + ktf(t) - k[1-F(t)] \\ &= (k+c)F(t) - k. \end{aligned}$$

Equating the rightmost side to zero shows that the minimal expected cost is obtained when you leave t^* minutes before your appointment, where t^* satisfies

$$F(t^*) = \frac{k}{k+c}.$$

Corollary 11.

If a and b are constants, then

$$E[aX + b] = aE[X] + b.$$

The proof of Corollary (19) for a continuous random variable X is the same as the one given for a discrete random variable. The only modification is that the sum is replaced by an integral and the probability mass function by a probability density function.
The variance of a continuous random variable is defined exactly as it is for a discrete random variable, namely, if X is a random variable with expected value μ, then the variance of X is defined (for any type of random variable) by

$$Var(X) = E[(X - \mu)^2].$$

The alternative formula,

$$Var(X) = E[X^2] - (E[X])^2$$

is established in a manner similar to its counterpart in the discrete case.

Example 12.

Find $Var(X)$ for X as given in Example (5).

Solution: We first compute $E[X^2]$:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^1 2x^3 dx = \frac{1}{2}.$$

Hence, since $E[X] = \frac{2}{3}$, we obtain

$$Var(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}.$$

It can be shown that, for constants a and b, $Var(aX + b) = a^2 Var(X)$. The proof mimics the one given for discrete random variables.
There are several important classes of continuous random variables that appear frequently in applications of probability.

We shall now discuss some of them.

The Uniform Random Variable

A random variable is said to be *uniformly* distributed over the interval $(0, 1)$ if its probability density function is given by

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Note that Equation (3) is a density function, since $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 dx = 1$. Because $f(x) > 0$ only when $x \in (0, 1)$, it follows that X must assume a value in interval $(0, 1)$. Also, since $f(x)$ is constant for $x \in (0, 1)$, X is just as likely to be near any value in $(0, 1)$ as it is to be near any other value.

To verify this statement, note that, for any $0 < a < b < 1$,

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx = b - a.$$

In other words, the probability that X is in any particular subinterval of $(0, 1)$ equals the length of that subinterval.

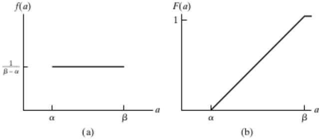
In general, we say that X is a uniform random variable on the interval (α, β) if the probability density function of X is given by

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if} \quad \alpha < x < \beta \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Since $F(a) = \int_{-\infty}^{a} f(x)dx$, it follows from Equation (4) that the distribution function of a uniform random variable on the interval (α, β) is given by

$$F(a) = \begin{cases} 0 & a \leq \alpha \\ \frac{a-\alpha}{\beta-\alpha} & \alpha < a < \beta \\ 1 & a \geq \beta. \end{cases}$$

The following figure presents a graph of $f(a)$ and $F(a)$.



Graph of (a) f(a) **and (b)** F(a) **for a uniform** (α, β) **random variable.**

Example 13.

Let X be uniformly distributed over (α, β). Find (a) $E[X]$ and (b) $Var(X)$.

Solution: (a)

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx$$
$$= \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} = \frac{\beta + \alpha}{2}.$$

In words, the expected value of a random variable that is uniformly distributed over some interval is equal to the midpoint of that interval.

Solution (Contd...)

To find $Var(X)$, we first calculate $E[X^2]$.

$$E[X^2] = \int_{\alpha}^{\beta} \frac{1}{\beta - \alpha} x^2 dx = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}.$$

Hence,

$$Var(X) = \frac{\beta^2 + \alpha\beta + \alpha^2}{3} - \frac{(\alpha + \beta)^2}{4} = \frac{(\beta - \alpha)^2}{12}.$$

Therefore, the variance of a random variable that is uniformly distributed over some interval is the square of the length of that interval divided by 12.

Example 14.

If X is uniformly distributed over $(0, 10)$, calculate the probability that (a) $X < 3$, (b) $X > 6$, and (c) $3 < X < 8$.

Solution:

(a) $P\{X < 3\} = \int_0^3 \frac{1}{10} dx = \frac{3}{10}$

(b) $P\{X > 6\} = \int_6^{10} \frac{1}{10} dx = \frac{4}{10}$

(c) $P\{3 < X < 8\} = \int_3^8 \frac{1}{10} dx = \frac{1}{2}$.

Example 15.

Buses arrive at a specified stop at 15*-minute intervals starting at* 7 *A.M. That is, they arrive at* 7*,* 7 : 15*,* 7 : 30*,* 7 : 45*, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between* 7 *and* 7 : 30*, find the probability that he waits*

(a) *less than* 5 *minutes for a bus;*

(b) *more than* 10 *minutes for a bus.*

Solution: Let X denote the number of minutes past 7 that the passenger arrives at the stop. Since X is a uniform random variable over the interval $(0, 30)$, it follows that the passenger will have to wait less than 5 minutes if (and only if) he arrives between 7 : 10 and 7 : 15 or between 7 : 25 and 7 : 30. Hence, the desired probability for part (a) is

$$P\{10 < X < 15\} + P\{25 < X < 30\} = \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx = \frac{1}{3}.$$

Similarly, he would have to wait more than 10 minutes if he arrives between 7 and 7 : 05 or between 7 : 15 and 7 : 20, so the probability for part (b) is

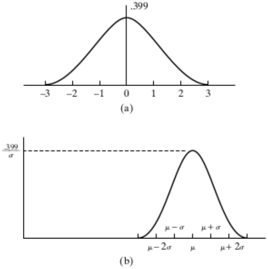
$$P\{0 < X < 5\} + P\{15 < X < 20\} = \frac{1}{3}.$$

Normal Random Variables

We say that X is a normal random variable, or simply that X is normally distributed, with parameters μ and σ^2 if the density of X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty.$$

This density function is a bell-shaped curve that is symmetric about μ.



Normal density function: (a) $\mu = 0$, $\sigma = 1$; (b) arbitrary μ, σ^2.

Normal Random Variables

The normal distribution was introduced by the French mathematician Abraham DeMoivre in 1733, who used it to approximate probabilities associated with binomial random variables when the binomial parameter n is large. This result was later extended by Laplace and others and is now encompassed in a probability theorem known as the central limit theorem.

The central limit theorem, one of the two most important results in probability theory, gives a theoretical base to the often noted empirical observation that, in practice, many random phenomena obey, at least approximately, a normal probability distribution. Some examples of random phenomena obeying this behavior are the height of a man, the velocity in any direction of a molecule in gas, and the error made in measuring a physical quantity.

Normal Random Variables

To prove that $f(x)$ is indeed a probability density function, we need to show that

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx = 1.$$

Making the substitution $y = (x-\mu)/\sigma$, we see that

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}\,dy.$$

Hence, we must show that

$$\int_{-\infty}^{\infty} e^{-y^2/2}\,dy = \sqrt{2\pi}.$$

Toward this end, let $I = \int_{-\infty}^{\infty} e^{-y^2/2}\,dy$. Then

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-y^2/2}\,dy \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(y^2+x^2)/2}\,dy\,dx. \end{aligned}$$

Normal Random Variables

We now evaluate the double integral by means of a change of variables to polar coordinates. (That is, let $x = r\cos\theta$, $y = r\sin\theta$, and $dydx = rd\theta dr$.) Thus,

$$\begin{aligned} I^2 &= \int_0^\infty \int_0^{2\pi} e^{-r^2/2} r d\theta dr \\ &= 2\pi \int_0^\infty r e^{-r^2/2} dr \\ &= -2\pi e^{-r^2/2}\Big|_0^\infty = 2\pi. \end{aligned}$$

Hence, $I = \sqrt{2\pi}$, and the result is proved.
An important fact about normal random variables is that if X is normally distributed with parameters μ and σ^2, then $Y = aX + b$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$. To prove this statement, suppose that $a > 0$. (The proof when $a < 0$ is similar.) Let F_Y denote the cumulative distribution function of Y. Then

$$\begin{aligned} F_Y(x) &= P\{Y \leq x\} = P\{aX + b \leq x\} \\ &= P\{X \leq \frac{x-b}{a}\} = F_X\left(\frac{x-b}{a}\right) \end{aligned}$$

where F_X is the cumulative distribution function of X.

Normal Random Variables

By differentiation, the density function of Y is then

$$
\begin{aligned}
f_Y(x) &= \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\
&= \frac{1}{a\sigma\sqrt{2\pi}} \exp\{-\left(\frac{x-b}{a} - \mu\right)^2 / 2\sigma^2\} \\
&= \frac{1}{a\sigma\sqrt{2\pi}} \exp\{-(x-b-a\mu)^2/2(a\sigma)^2\}
\end{aligned}
$$

which shows that Y is normal with parameters $a\mu + b$ and $a^2\sigma^2$.
An important implication of the preceding result is that if X is normally distributed with parameters μ and σ^2, then $Z = (X - \mu)/\sigma$ is normally distributed with parameters 0 and 1. Such a random variable is said to be a *standard*, or a *unit*, normal random variable.

We now show that the parameters μ and σ^2 of a normal random variable represent, respectively, its expected value and variance.

Example 16.

Find $E[X]$ and $Var(X)$ when X is a normal random variable with parameters μ and σ^2.

Solution: Let us start by finding the mean and variance of the standard normal random variable $Z = (X - \mu)/\sigma$. We have

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} x f_Z(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx \\ &= -\frac{1}{\sqrt{2\pi}} \; e^{-x^2/2} \Big|_{-\infty}^{\infty} \\ &= 0. \end{aligned}$$

Thus,

$$\begin{aligned} Var(Z) &= E[Z^2] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx. \end{aligned}$$

Solution (Contd...)

Integration by parts (with $u = x$ and $dv = xe^{-x^2/2}$) now gives

$$
\begin{aligned}
Var(Z) &= \frac{1}{\sqrt{2\pi}}\left(-xe^{-x^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2}dx\right) \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx \\
&= 1.
\end{aligned}
$$

Because $X = \mu + \sigma Z$, the preceding yields the results

$$E[X] = \mu + \sigma E[Z] = \mu$$

and

$$Var(X) = \sigma^2 Var(Z) = \sigma^2.$$

It is customary to denote the cumulative distribution function of a standard normal random variable by $\Phi(x)$. That is,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}dy.$$

Solution (Contd...)

The values of $\Phi(x)$ for nonnegative x are given in the following table. For negative values of x, $\Phi(x)$ can be obtained from the relationship

$$\Phi(-x) = 1 - \Phi(x), \quad -\infty < x < \infty. \tag{5}$$

The proof of Equation (5), which follows from the symmetry of the standard normal density. This equation states that if Z is a standard normal random variable, then

$$P\{Z \leq -x\} = P\{Z > x\}, \quad -\infty < x < \infty.$$

Since $Z = (X - \mu)/\sigma$ is a standard normal random variable whenever X is normally distributed with parameters μ and σ^2, it follows that the distribution function of X can be expressed as

$$F_X(a) = P\{X \leq a\} = P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Table for Standard Normal Curve

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Example 17.

If X is a normal random variable with parameters $\mu = 3$ and $\sigma^2 = 9$, find (a) $P\{2 < X < 5\}$; (b) $P\{X > 0\}$; (c) $P\{|X - 3| > 6\}$.

Solution: (a)

$$\begin{aligned} P\{2 < X < 5\} &= P\left\{\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right\} \\ &= P\left\{-\frac{1}{3} < Z < \frac{2}{3}\right\} \\ &= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) \\ &= \Phi\left(\frac{2}{3}\right) - \left[1 - \Phi\left(\frac{1}{3}\right)\right] \approx .3779. \end{aligned}$$

Solution (Contd...)

(b)

$$P\{X > 0\} = P\left\{\frac{X-3}{3} > \frac{0-3}{3}\right\} = P\{Z > -1\}$$
$$= 1 - \Phi(-1)$$
$$= \Phi(1)$$
$$\approx .8413.$$

(c)

$$P\{|X-3| > 6\} = P\{X > 9\} + P\{X < -3\}$$
$$= P\left\{\frac{X-3}{3} > \frac{9-3}{3}\right\} + P\left\{\frac{X-3}{3} < \frac{-3-3}{3}\right\}$$
$$= P\{Z > 2\} + P\{Z < -2\}$$
$$= 1 - \Phi(2) + \Phi(-2)$$
$$= 2[1 - \Phi(2)]$$
$$\approx .0456.$$

Example 18.

An examination is frequently regarded as being good (in the sense of determining a valid grade spread for those taking it) if the test scores of those taking the examination can be approximated by a normal density function. (In other words, a graph of the frequency of grade scores should have approximately the bell-shaped form of the normal density.) The instructor often uses the test scores to estimate the normal parameters μ and σ^2 and then assigns the letter grade A to those whose test score is greater than $\mu + \sigma$, B to those whose score is between μ and $\mu + \sigma$, C to those whose score is between $\mu - \sigma$ and μ, D to those whose score is between $\mu - 2\sigma$ and $\mu - \sigma$, and F to those getting a score below $\mu - 2\sigma$. (This strategy is sometimes referred to as grading "on the curve.")

Solution :

$$P\{X > \mu + \sigma\} = P\left\{\frac{X - \mu}{\sigma} > 1\right\} = 1 - \Phi(1) \approx .1587.$$

$$P\{\mu < X < \mu + \sigma\} = P\left\{0 < \frac{X - \mu}{\sigma} < 1\right\} = \Phi(1) - \Phi(0) \approx .3413.$$

$$P\{\mu - \sigma < X < \mu\} = P\left\{-1 < \frac{X - \mu}{\sigma} < 0\right\}$$

$$= \Phi(0) - \Phi(-1) \approx .3413.$$

Example (Contd...)

$$P\{\mu - 2\sigma < X < \mu - \sigma\} = P\{-2 < \frac{X-\mu}{\sigma} < -1\}$$
$$= \Phi(2) - \Phi(1) \approx .1359.$$
$$P\{X < \mu - 2\sigma\} = P\left\{\frac{X-\mu}{\sigma} < -2\right\} = \Phi(-2) \approx .0228.$$

It follows that approximately 16 percent of the class will receive an A grade on the examination, 34 percent a B grade, 34 percent a C grade, and 14 percent a D grade; 2 percent will fail.

Example 19.

An expert witness in a paternity suit testifies that the length (in days) of human gestation is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short gestation indicated by the testimony?

Solution: Let X denote the length of the gestation, and assume that the defendant is the father. Then the probability that the birth could occur within the indicated period is

$$\begin{aligned} P\{X > 290 \text{ or } X < 240\} &= P\{X > 290\} + P\{X < 240\} \\ &= P\left\{\frac{X - 270}{10} > 2\right\} + P\left\{\frac{X - 270}{10} < -3\right\} \\ &= 1 - \Phi(2) + 1 - \Phi(3) \\ &\approx .0241. \end{aligned}$$

Example 20.

Suppose that a binary message, either 0 *or* 1*, must be transmitted by wire from location* A *to location* B*. However, the data sent over the wire are subject to a channel noise disturbance, so, to reduce the possibility of error, the value* 2 *is sent over the wire when the message is* 1 *and the value* -2 *is sent when the message is* 0*. If* x*,* $x = \pm 2$*, is the value sent at location* A*, then* R*, the value received at location* B*, is given by* $R = x + N$*, where* N *is the channel noise disturbance. When the message is received at location* B*, the receiver decodes it according to the following rule:*
If $R \geq .5$*, then* 1 *is concluded.*
If $R < .5$*, then* 0 *is concluded.*

Example (contd...)

Because the channel noise is often normally distributed, we will determine the error probabilities when N is a standard normal random variable. Two types of errors can occur: One is that the message 1 can be incorrectly determined to be 0, and the other is that 0 can be incorrectly determined to be 1. The first type of error will occur if the message is 1 and $2 + N < .5$, whereas the second will occur if the message is 0 and $-2 + N \geq .5$. Hence,

$$\begin{aligned} P\{\text{error}|\text{message is } 1\} &= P\{N < -1.5\} \\ &= 1 - \Phi(1.5) \approx .0668 \end{aligned}$$

and

$$\begin{aligned} P\{\text{error}|\text{message is } 0\} &= P\{N \geq 2.5\} \\ &= 1 - \Phi(2.5) \approx .0062. \end{aligned}$$

The Normal Approximation to The Binomial Distribution

An important result in probability theory known as the DeMoivre-Laplace limit theorem states that when n is large, a binomial random variable with parameters n and p will have approximately the same distribution as a normal random variable with the same mean and variance as the binomial.

This result was proved originally for the special case of $p = \frac{1}{2}$ by DeMoivre in 1733 and was then extended to general p by Laplace in 1812. It formally states that if we "standardize" the binomial by first subtracting its mean np and then dividing the result by its standard deviation $\sqrt{np(1-p)}$, then the distribution function of this standardized random variable (which has mean 0 and variance 1) will converge to the standard normal distribution function as $n \to \infty$.

The DeMoivre-Laplace Limit Theorem

Theorem 21.

If S_n denotes the number of successes that occur when n independent trials, each resulting in a success with probability p, are performed, then, for any $a < b$,

$$P\left\{a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right\} \to \Phi(b) - \Phi(a)$$

as $n \to \infty$.

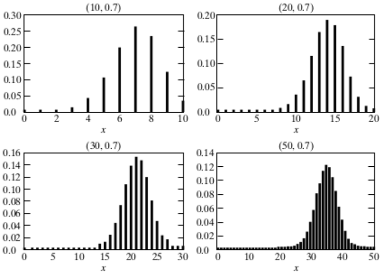
Because the preceding theorem is only a special case of the central limit theorem, we shall not present a proof.

The DeMoivre-Laplace Limit Theorem

Note that we now have two possible approximations to binomial probabilities: the Poisson approximation, which is good when n is large and p is small, and the normal approximation, which can be shown to be quite good when $np(1-p)$ is large.

The normal approximation will, in general, be quite good for values of n satisfying $np(1-p) \geq 10$.

Normal Approximation



The probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger.

Example 22.

Let X be the number of times that a fair coin that is flipped 40 times lands on heads. Find the probability that $X = 20$. Use the normal approximation and then compare it with the exact solution.

Solution : To employ the normal approximation, note that because the binomial is a discrete integer-valued random variable, whereas the normal is a continuous random variable, it is best to write $P\{X = i\}$ as $P\{i - 1/2 < X < i + 1/2\}$ before applying the normal approximation (this is called the *continuity correction*). Doing so gives

$$\begin{aligned} P\{X = 20\} &= P\{19.5 \leq X < 20.5\} \\ &= P\left\{\frac{19.5 - 20}{\sqrt{10}} < \frac{X - 20}{\sqrt{10}} < \frac{20.5 - 20}{\sqrt{10}}\right\} \\ &\approx P\left\{-.16 < \frac{X - 20}{\sqrt{10}} < .16\right\} \\ &\approx \Phi(.16) - \Phi(-.16) \approx .1272. \end{aligned}$$

The exact result is

$$P\{X = 20\} = \binom{40}{20}\left(\frac{1}{2}\right)^{40} \approx .1254.$$

Example 23.

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

Solution: If X denotes the number of students that attend, then X is a binomial random variable with parameters $n = 450$ and $p = .3$. Using the continuity correction, we see that the normal approximation yields

$$\begin{aligned} P\{X \geq 150.5\} &= P\left\{\frac{X - (450)(.3)}{\sqrt{450(.3)(.7)}} \geq \frac{150.5 - (450)(.3)}{\sqrt{450(.3)(.7)}}\right\} \\ &\approx 1 - \Phi(1.59) \\ &\approx .0559. \end{aligned}$$

Hence, less than 6 percent of the time do more than 150 of the first 450 accepted actually attend.

Example 24.

To determine the effectiveness of a certain diet in reducing the amount of cholesterol in the bloodstream, 100 people are put on the diet. After they have been on the diet for a sufficient length of time, their cholesterol count will be taken. The nutritionist running this experiment has decided to endorse the diet if at least 65 percent of the people have a lower cholesterol count after going on the diet. What is the probability that the nutritionist endorses the new diet if, in fact, it has no effect on the cholesterol level?

Solution. Let us assume that if the diet has no effect on the cholesterol count, then, strictly by chance, each person's count will be lower than it was before the diet with probability $\frac{1}{2}$. Hence, if X is the number of people whose count is lowered, then the probability that the nutritionist will endorse the diet when it actually has no effect on the cholesterol count is

$$\begin{aligned}\sum_{i=65}^{100} \binom{100}{i} \left(\frac{1}{2}\right)^{100} &= P\{X \geq 64.5\} \\ &= P\left\{ \frac{x - (100)(\frac{1}{2})}{\sqrt{100(\frac{1}{2})(\frac{1}{2})}} \geq 2.9 \right\} \\ &\approx 1 - \Phi(2.9) \\ &\approx .0019.\end{aligned}$$

Example 25.

Fifty-two percent of the residents of New York City are in favor of outlawing cigarette smoking in publicly owned areas. Approximate the probability that more than 50 *percent of a random sample of n people from New York are in favor of this prohibition when*

(a) $n = 11$

(b) $n = 101$

(c) $n = 1001$

How large would n have to be to make this probability exceed .95*?*

Solution. Let N denote the number of residents of New York City. To answer the preceding question, we must first understand that a random sample of size n is a sample such that the n people were chosen in such a manner that each of the $\binom{N}{n}$ subsets of n people had the same chance of being the chosen subset. Consequently, S_n, the number of people in the sample who are in favor of the smoking prohibition, is a hypergeometric random variable. That is, S_n has the same distribution as the number of white balls obtained when n balls are chosen from an urn of N balls, of which $.52N$ are white. But because N and $.52N$ are both large in comparison with the sample size n, it follows from the binomial approximation to the hypergeometric that the distribution of S_n is closely approximated by a binomial distribution with parameters n and $p = .52$.

Example (contd...)

The normal approximation to the binomial distribution then shows that

$$\begin{aligned} P\{S_n > .5n\} &= P\left\{ \frac{S_n - .52n}{\sqrt{n(.52)(.48)}} > \frac{.5n - .52n}{\sqrt{n(.52)(.48)}} \right\} \\ &= P\left\{ \frac{S_n - .52n}{\sqrt{n(.52)(.48)}} > -.04\sqrt{n} \right\} \\ &\approx \Phi(.04\sqrt{n}). \end{aligned}$$

Thus,

$$P\{S_n > .5n\} \approx \begin{cases} \Phi(.1328) = .5528, & \text{if } n = 11 \\ \Phi(.4020) = .6562, & \text{if } n = 101 \\ \Phi(1.2665) = .8973, & \text{if } n = 1001. \end{cases}$$

In order for this probability to be at least .95, we would need $\Phi(.04\sqrt{n}) > .95$. Because $\Phi(x)$ is an increasing function and $\Phi(1.645) = .95$, this means that

$$.04\sqrt{n} > 1.645 \quad \text{or} \qquad n \geq 1691.266$$

That is, the sample size would have to be at least 1692.

Exponential Random Variables

A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

is said to be an exponential random variable (or, more simply, is said to be exponentially distributed) with parameter λ. The cumulative distribution function $F(a)$ of an exponential random variable is given by

$$F(a) = P\{X \leq a\} = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}, \quad a \geq 0.$$

Note that $F(\infty) = \int_0^\infty \lambda e^{-\lambda x} dx = 1$, as, of course, it must. The parameter λ will now be shown to equal the reciprocal of the expected value.

Example 26.

Let X be an exponential random variable with parameter λ. Calculate (a)$E[X]$ and (b)$Var(X)$.

Solution. (a) Since the density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

we obtain, for $n > 0$,

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx.$$

Example (contd...)

Integrating by parts (with $\lambda e^{-\lambda x} = dv$ and $u = x^n$) yields

$$E[X^n] = x^n e^{-\lambda x}|_0^\infty + \int_0^\infty e^{-\lambda x} n x^{n-1} dx$$
$$= 0 + \frac{n}{\lambda}\int_0^\infty \lambda e^{-\lambda x} x^{n-1} dx = \frac{n}{\lambda} E[X^{n-1}].$$

Letting $n = 1$ and then $n = 2$ gives

$$E[X] = \frac{1}{\lambda}, \qquad E[X^2] = \frac{2}{\lambda}E[X] = \frac{2}{\lambda^2}.$$

(b) Hence

$$Var(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Thus, the mean of the exponential is the reciprocal of its parameter λ, and the variance is the mean squared.

Exponential Distribution

In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs.

For instance, the amount of time (starting from now) until an earthquake occurs, or until a new war breaks out, or until a telephone call you receive turns out to be a wrong number are all random variables that tend in practice to have exponential distributions.

Example 27.

Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait

1. *more than* 10 *minutes;*
2. *between* 10 *and* 20 *minutes.*

Solution. Let X denote the length of the call made by the person in the booth. Then the desired probabilities are

(a) $P\{X > 10\} = 1 - F(10) = e^{-1} \approx .368.$

(b) $P\{10 < X < 20\} = F(20) - F(10) = e^{-1} - e^{-2} \approx .233.$

Memoryless non-negative random variable

We say that a nonnegative random variable X is memoryless if

$$P\{X > s + t | X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0 \tag{6}$$

If we think of X as being the lifetime of some instrument, Equation 6 states that the probability that the instrument survives for at least $s + t$ hours, given that it has survived t hours, is the same as the initial probability that it survives for at least s hours.

In other words, if the instrument is alive at age t, the distribution of the remaining amount of time that it survives is the same as the original lifetime distribution. (That is, it is as if the instrument does not "remember" that it has already been in use for a time t.)

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

or

$$P\{X > s + t\} = P\{X > s\}P\{X > t\} \tag{7}$$

Since Equation (7) is satisfied when X is exponentially distributed (for $e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t}$), it follows that exponentially distributed random variables are memoryless.

Memoryless non-negative random variable

Example 28.

Consider a post office that is staffed by two clerks. Suppose that when Mr. Smith enters the system, he discovers that Ms. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Ms. Jones or Mr. Brown leaves. If the amount of time that a clerk spends with a customer is exponentially distributed with parameter λ, what is the probability that, of the three customers, Mr. Smith is the last to leave the post office?

Solution : The answer is obtained by reasoning as follows: Consider the time at which Mr. Smith first finds a free clerk. At this point, either Ms. Jones or Mr. Brown would have just left, and the other one would still be in service.

However, because the exponential is memoryless, it follows that the additional amount of time that this other person (either Ms. Jones or Mr. Brown) would still have to spend in the post office is exponentially distributed with parameter λ. That is, it is the same as if service for that person were just starting at this point. Hence, by symmetry, the probability that the remaining person finishes before Smith leaves must equal $\frac{1}{2}$.

Example

It turns out that not only is the exponential distribution memoryless, but it is also the unique distribution possessing this property. To see this, suppose that X is memoryless and let $\overline{F}(x) = P\{X > x\}$. Then, $\overline{F}(s+t) = \overline{F}(s)\overline{F}(t)$. That is, $\overline{F}()$ satisfies the functional equation $g(s+t) = g(s)g(t)$. However, it turns out that the only right continuous solution of this functional equation is[1]

$$g(x) = e^{-\lambda x} \tag{8}$$

and, since a distribution function is always right continuous, we must have $\overline{F}(x) = e^{-\lambda x}$ or $F(x) = P\{X \leq x\} = 1 - e^{-\lambda x}$ which shows that X is exponentially distributed.

[1] One can prove Equation (8) as follows: If $g(s+t) = g(s)g(t)$, then

$$g\left(\frac{2}{n}\right) = g\left(\frac{1}{n} + \frac{1}{n}\right) = g^2\left(\frac{1}{n}\right)$$

and repeating this yields $g(m/n) = g^m(1/n)$. Also,

$$g(1) = g\left(\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}\right) = g^n\left(\frac{1}{n}\right) \quad \text{or} \quad g\left(\frac{1}{n}\right) = (g(1))^{1/n}$$

Hence, $g(m/n) = (g(1))^{m/n}$, which, since g is right continuous, implies that $g(x) = (g(1))^x$. Because $g(1) = \left(g\left(\frac{1}{2}\right)\right)^2 \geq 0$, we obtain $g(x) = e^{-\lambda x}$, where $\lambda = -\log(g(1))$.

Laplace Distribution

A variation of the exponential distribution is the distribution of a random variable that is equally likely to be either positive or negative and whose absolute value is exponentially distributed with parameter λ, $\lambda \geq 0$. Such a random variable is said to have a *Laplace* distribution,[2] and its density is given by

$$f(x) = \frac{1}{2}\lambda e^{-\lambda|x|} \qquad -\infty < x < \infty.$$

Its distribution function is given by

$$F(x) = \begin{cases} \frac{1}{2}\int\limits_{-\infty}^{x} \lambda e^{\lambda x} dx & x < 0 \\ \frac{1}{2}\int\limits_{-\infty}^{0} \lambda e^{\lambda x} dx + \frac{1}{2}\int\limits_{0}^{x} \lambda e^{-\lambda x} dx & x > 0 \end{cases}$$

$$= \begin{cases} \frac{1}{2}e^{\lambda x} & x < 0 \\ 1 - \frac{1}{2}e^{-\lambda x} & x > 0. \end{cases}$$

[2] It also is sometimes called the double exponential random variable.

Example 29.

Consider again Example (20), which supposes that a binary message is to be transmitted from A to B, with the value 2 being sent when the message is 1 and -2 when it is 0. However, suppose now that, rather than being a standard normal random variable, the channel noise N is a Laplacian random variable with parameter $\lambda = 1$. Suppose again that if R is the value received at location B, then the message is decoded as follows:
If $R \geq .5$, then 1 is concluded.
If $R < .5$, then 0 is concluded.

In this case, where the noise is Laplacian with parameter $\lambda = 1$, the two types of errors will have probabilities given by

$$P\{\text{error}|\text{message 1 is sent}\} = P\{N < -1.5\} = \frac{1}{2}e^{-1.5} \approx .1116.$$

$$P\{\text{error}|\text{message 0 is sent}\} = P\{N \geq 2.5\} = \frac{1}{2}e^{-2.5} \approx .041.$$

On comparing this with the results of Example (20), we see that the error probabilities are higher when the noise is Laplacian with $\lambda = 1$ than when it is a standard normal variable.

Hazard Rate Functions

Consider a positive continuous random variable X that we interpret as being the lifetime of some item. Let X have distribution function F and density f. The *hazard rate* (sometimes called the *failure rate*) function $\lambda(t)$ of F is defined by

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)}, \quad \text{where } \overline{F} = 1 - F.$$

To interpret $\lambda(t)$, suppose that the item has survived for a time t and we desire the probability that it will not survive for an additional time dt. That is, consider $P\{X \in (t, t + dt) | X > t\}$. Now,

$$\begin{aligned} P\{X \in (t, t + dt) | X > t\} &= \frac{P\{X \in (t, t + dt), X > t\}}{P\{X > t\}} \\ &= \frac{P\{X \in (t, t + dt)\}}{P\{X > t\}} \\ &\approx \frac{f(t)}{\overline{F}(t)} dt. \end{aligned}$$

Thus, $\lambda(t)$ represents the conditional probability intensity that a t-unit-old item will fail.

Hazard Rate Functions

Suppose now that the lifetime distribution is exponential. Then, by the memoryless property, it follows that the distribution of remaining life for a t-year-old item is the same as that for a new item. Hence, $\lambda(t)$ should be constant. In fact, this checks out, since

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda.$$

Thus, the failure rate function for the exponential distribution is constant. The parameter λ is often referred to as the *rate* of the distribution.

It turns out that the failure rate function $\lambda(t)$ uniquely determines the distribution F. To prove this, note that, by definition,

$$\lambda(t) = \frac{\frac{d}{dt}F(t)}{1 - F(t)}.$$

Integrating both sides yields

$$\log(1 - F(t)) = -\int_0^t \lambda(t)dt + k$$

or

$$1 - F(t) = e^k \exp\left\{-\int_0^t \lambda(t)dt\right\}.$$

Hazard Rate Functions

Letting $t = 0$ shows that $k = 0$; thus,

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(t)dt\right\}.$$

Hence, a distribution function of a positive continuous random variable can be specified by giving its hazard rate function. For instance, if a random variable has a linear hazard rate function-that is, if

$$\lambda(t) = a + bt$$

then its distribution function is given by

$$F(t) = 1 - e^{-at-bt^2/2}$$

and differentiation yields its density, namely,

$$f(t) = (a + bt)e^{-(at+bt^2/2)} \quad t \geq 0.$$

When $a = 0$, the preceding equation is known as the *Rayleigh density function*.

The Gamma Distribution

A random variable is said to have a gamma distribution with parameters $(\alpha, \lambda), \lambda > 0, \alpha > 0$, if its density function is given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\Gamma(\alpha)$, called the gamma function, is defined as

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy.$$

Integration of $\Gamma(\alpha)$ by parts yields

$$\begin{aligned} \Gamma(\alpha) &= -e^{-y} y^{\alpha-1}|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2} dy \\ &= (\alpha-1)\int_0^\infty e^{-y} y^{\alpha-2} dy \\ &= (\alpha-1)\Gamma(\alpha-1). \end{aligned} \tag{9}$$

For integral values of α, say, $\alpha = n$, we obtain, by applying Equation 9 repeatedly,

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= \cdots \\ &= (n-1)(n-2)\cdots 3\cdot 2\Gamma(1)\end{aligned}$$

Since $\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1$, it follows that, for integral values of n,

$$\Gamma(n) = (n-1)!$$

Other Continuous Distributions

When α is a positive integer, say, $\alpha = n$, the gamma distribution with parameters (α, λ) often arises, in practice as the distribution of the amount of time one has to wait until a total of n events has occurred.

The amount of time one has to wait until a total of n events has occurred will be a gamma random variable with parameters (n, λ).

To prove this, let T_n denote the time at which the nth event occurs, and note that T_n is less than or equal to t if and only if the number of events that have occurred by time t is at least n. That is, with $N(t)$ equal to the number of events in $[0, t]$,

$$P\{T_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} P\{N(t) = j\} = \sum_{j=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^j}{j!}$$

where the final identity follows because the number of events in $[0, t]$ has a Poisson distribution with parameter λt. Differentiation of the preceding now yields the density function of T_n :

$$f(t) = \sum_{j=n}^{\infty} \frac{e^{-\lambda t} j(\lambda t)^{j-1}\lambda}{j!} - \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!}$$

$$= \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^{j-1}}{j-1!} - \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!}$$

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

Hence, T_n has the gamma distribution with parameters (n, λ). (This distribution is often referred to in the literature as the n-Erlang distribution.) Note that when $n = 1$, this distribution reduces to the exponential distribution.

The gamma distribution with $\lambda = \frac{1}{2}$ and $\alpha = n/2, n$ a positive integer, is called the X_n^2 (read "chi-squared") distribution with n degrees of freedom. The chi-squared distribution often arises in practice as the distribution of the error involved in attempting to hit a target in n-dimensional space when each coordinate error is normally distributed.

Example 30.

Let X be a gamma random variable with parameters α and λ. Calculate (a) $E[X]$ and (b)$Var(X)$.

Solution: (a)

$$\begin{aligned} E[X] &= \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda x e^{-\lambda x} (\lambda x)^{\alpha-1} dx \\ &= \frac{1}{\lambda \Gamma(\alpha)} \int_0^\infty \lambda e^{-\lambda x} (\lambda x)^{\alpha} dx = \frac{\Gamma(\alpha+1)}{\lambda \Gamma(\alpha)} = \frac{\alpha}{\lambda} \text{ by Equation (9).} \end{aligned}$$

By first calculating $E[X^2]$, we can show that

$$Var(X) = \frac{\alpha}{\lambda^2}.$$

The Weibull Distribution

The Weibull distribution is widely used in engineering practice due to its versatility. It was originally proposed for the interpretation of fatigue data, but now its use has been extended to many other engineering problems. In particular, it is widely used in the field of life phenomena as the distribution of the lifetime of some object, especially when the "weakest link" model is appropriate for the object.

That is, consider an object consisting of many parts, and suppose that the object experiences death (failure) when any of its parts fail. It has been shown (both theoretically and empirically) that under these conditions a Weibull distribution provides a close approximation to the distribution of the lifetime of the item.

The Weibull Distribution

The Weibull distribution function has the form

$$F(x) = \begin{cases} 0 & x \leq v \\ 1 - \exp\left\{ - \left(\frac{x-v}{\alpha}\right)^{\beta} \right\} & x > v. \end{cases} \tag{10}$$

A random variable whose cumulative distribution function is given by Equation (10) is said to be a Weibull random variable with parameters v, α, and β. Differentiation yields the density:

$$F(x) = \begin{cases} 0 & x \leq v \\ \frac{\beta}{\alpha}\left(\frac{x-v}{\alpha}\right)^{\beta-1} \exp\left\{ - \left(\frac{x-v}{\alpha}\right)^{\beta} \right\} & x > v. \end{cases}$$

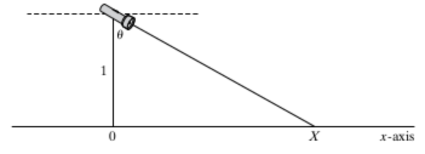
The Cauchy Distribution

A random variable is said to have a Cauchy distribution with parameter $\theta, -\infty < \theta < \infty$, if its density is given by

$$f(x) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2} \quad -\infty < x < \infty.$$

Example 31.

Suppose that a narrow-beam flashlight is spun around its center, which is located a unit distance from the x-axis. Consider the point X at which the beam intersects the x-axis when the flashlight has stopped spinning. (If the beam is not pointing toward the x-axis, repeat the experiment.)



Cauchy Distribution

As indicated in the above figure, the point X is determined by the angle θ between the flashlight and the y-axis, which, from the physical situation, appears to be uniformly distributed between $-\pi/2$ and $\pi/2$. The distribution function of X is thus given by

$$\begin{aligned} F(x) &= P\{X \leq x\} \\ &= P\{\tan\theta \leq x\} \\ &= P\{\theta \leq \tan^{-1} x\} \\ &= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x \end{aligned}$$

where the last equality follows since θ, being uniform over $(-\pi/2, \pi/2)$, has distribution

$$P\{\theta \leq a\} = \frac{a - (-\pi/2)}{\pi} = \frac{1}{2} + \frac{a}{\pi}, \quad -\frac{\pi}{2} < a < \frac{\pi}{2}.$$

Hence, the density function of X is given by

$$f(x) = \frac{d}{dx}F(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

and we see that X has the Cauchy distribution.[1]

[1] That $\frac{d}{dx}(\tan^{-1} x) = 1/(1+x^2)$ can be seen as follows: If $y = \tan^{-1} x$, then $\tan y = x$, so

$$1 = \frac{d}{dx}(\tan y) = \frac{d}{dy}(\tan y)\frac{dy}{dx} = \frac{d}{dy}\left(\frac{\sin y}{\cos y}\right)\frac{dy}{dx} = \left(\frac{\cos^2 y + \sin^2 y}{\cos^2 y}\right)\frac{dy}{dx}$$

or

$$\frac{dy}{dx} = \frac{\cos^2 y}{\sin^2 y + \cos^2 y} = \frac{1}{\tan^2 y + 1} = \frac{1}{x^2+1}.$$

The Beta Distribution

A random variable is said to have a beta distribution if its density is given by

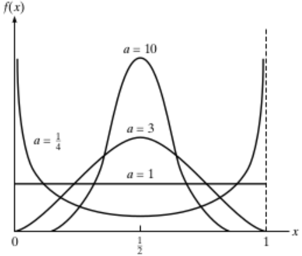
$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

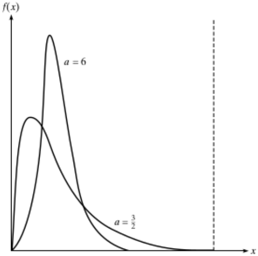
where

$$B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx.$$

The beta distribution can be used to model a random phenomenon whose set of possible values is some finite interval $[c, d]$ - which, by letting c denote the origin and taking $d - c$ as a unit measurement, can be transformed into the interval $[0, 1]$.

When $a = b$, the beta density is symmetric about $\frac{1}{2}$, giving more and more weight to regions about $\frac{1}{2}$ as the common value a increases. (See figure on the left.) When $b > a$, the density is skewed to the left (in the sense that smaller values become more likely); and it is skewed to the right when $a > b$. (See figure on the right.)





The Beta Distribution

The relationship

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{11}$$

can be shown to exist between

$$B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$$

and the gamma function.

Upon using Equation (9) along with the identity (11), it is an easy matter to show that if X is a beta random variable with parameters a and b, then

$$E[X] = \frac{a}{a+b}, \quad Var(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

The Distribution of a Function of a Random Variable

Often, we know the probability distribution of a random variable and are interested in determining the distribution of some function of it.

For instance, suppose that we know the distribution of X and want to find the distribution of $g(X)$. To do so, it is necessary to express the event that $g(X) \leq y$ in terms of X being in some set. We illustrate with the following examples.

Example 32.

Let X be uniformly distributed over $(0,1)$. We obtain the distribution of the random variable Y, defined by $Y = X^n$, as follows: For $0 \le y \le 1$,

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} \\ &= P\{X^n \le y\} \\ &= p\{X \le y^{1/n}\} \\ &= F_X(y^{1/n}) \\ &= y^{1/n}. \end{aligned}$$

For instance, the density function of Y is given by

$$f_Y(y) = \begin{cases} \frac{1}{n} y^{1/n-1} & 0 \le y \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

Example 33.

If X is a continuous random variable with probability density f_X, then the distribution of $Y = X^2$ is obtained as follows: For $y \geq 0$,

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} \\ &= P\{X^2 \leq y\} \\ &= P\{P - \sqrt{y} \leq X \leq \sqrt{y}\} \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Differentiation yields

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})].$$

Example 34.

If X has a probability density f_X, then $Y = |X|$ has a density function that is obtained as follows: For $y \geq 0$,

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} \\ &= P\{|X| \leq y\} \\ &= P\{-y \leq X \leq y\} \\ &= F_X(y) - F_X(-y). \end{aligned}$$

Hence, on differentiation, we obtain

$$f_Y(y) = f_X(y) + f_X(-y), \quad y \geq 0.$$

Theorem

The method employed in Examples (32) through (34) can be used to prove the following theorem.

Theorem 35.

Let X be a continuous random variable having probability density function f_X. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x. Then the random variable Y defined by $Y = g(X)$ has a probability density function given by

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)]\left|\frac{d}{dy}g^{-1}(y)\right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

where $g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

We shall prove Theorem (35) when $g(x)$ is an increasing function.

Proof

Suppose that $y = g(x)$ for some x. Then, with $Y = g(X)$,

$$\begin{aligned} F_Y(y) &= P\{g(X) \leq y\} \\ &= P\{X \leq g^{-1}(y)\} \\ &= F_X(g^{-1}(y)). \end{aligned}$$

Differentiation gives

$$f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)$$

which agrees with Theorem (35), since $g^{-1}(y)$ is nondecreasing, so its derivative is nonnegative.

When $y \neq g(x)$ for any x, then $F_Y(y)$ is either 0 or 1, and in either case $f_Y(y) = 0$.

Example 36.

Let X be a continuous nonnegative random variable with density function f, and let $Y = X^n$. Find f_Y, the probability density function of Y.

Solution. If $g(x) = x^n$, then $g^{-1}(y) = y^{1/n}$ and $\frac{d}{dy}\{g^{-1}(y)\} = \frac{1}{n}y^{1/n-1}$.

Hence, from Theorem (35), we obtain

$$f_Y(y) = \frac{1}{n}y^{1/n-1}f(y^{1/n}).$$

For $n = 2$, this gives

$$f_Y(y) = \frac{a}{2\sqrt{y}}f(\sqrt{y})$$

which (since $X \geq 0$) is in agreement with the result of Example (33).

Exercise 37.

Let X be a random variable with probability density function

$$f(x) = \begin{cases} c(1-x^2) & -1 < x < 1 \\ 0 & \textit{otherwise.} \end{cases}$$

1. *What is the value of c?*
2. *What is the cumulative distribution function of X?*

Solution :

1. $c \int_{-1}^{1} (1-x^2)dx = 1 \implies c = 3/4.$
2. $F(x) = \frac{3}{4} \int_{-1}^{x} (1-x^2)dx = \frac{3}{4}\left(x - \frac{x^3}{3} + \frac{2}{3}\right), \ -1 < x < 1.$

Exercise 38.

A system consisting of one original unit plus a spare can function for a random amount of time X. If the density of X is given (in units of months) by

$$f(x) = \begin{cases} Cxe^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

what is the probability that the system functions for at least 5 months?

Solution : $\int_0^\infty xe^{-x/2}dx = -2xe^{-x/2} - 4e^{-x/2}$. Hence,

$$c\int_0^\infty xe^{-x/2}dx = 1 \implies c = 1/4.$$

$$P\{X > 5\} = \frac{1}{4}\int_5^\infty xe^{-x/2}dx = \frac{1}{4}[10e^{-5/2} + 4e^{-5/2}] = \frac{14}{4}e^{-5/2}.$$

Exercise 39.

The probability density function of X, the lifetime of a certain type of electronic device (measured in hours), is given by

$$f(x) = \begin{cases} \frac{10}{x^2} & x > 10 \\ 0 & x \leq 10. \end{cases}$$

(a) *Find* $P\{X > 20\}$.

(b) *What is the cumulative distribution function of* X*?*

(c) *What is the probability that, of 6 such types of devices, at least 3 will function for at least 15 hours? What assumptions are you making?*

Solution :

(a) $\int_{20}^{\infty} \frac{10}{x^2} dx = \frac{-10}{x} \int_{20}^{\infty} = 1/2.$

(b) $F(y) = \int_{10}^{y} \frac{10}{x^2} dx = 1 - \frac{10}{y}$, $y > 10$. $F(y) = 0$ for $y < 10$.

(c) $\sum_{i=3}^{6} \binom{6}{i} \left(\frac{2}{3}\right)^i \left(\frac{1}{3}\right)^{6-i}$ since $\overline{F}(15) = \frac{10}{15}$. Assuming independence of the events that the devices exceed 15 hours.

Exercise 40.

Compute $E[X]$ *if* X *has a density function given by*

(a) $f(x) = \begin{cases} \frac{1}{4} xe{-}x/2 & x > 0 \\ 0 & otherwise \end{cases};$

(b) $f(x) = \begin{cases} c(1 - x^2) & -1 < x, 1 \\ 0 & otherwise \end{cases};$

(c) $f(x) = \begin{cases} \frac{5}{x^2} & x > 5 \\ 0 & x \leq 5 \end{cases}.$

Solution :

(a) $E[X] = \frac{1}{4} \int_0^\infty x^2 e^{-x/2} dx = 2 \int_0^\infty y^2 e^{-y} dx = 2\Gamma(3) = 4.$

(b) By symmetry of $f(x)$ about $x = 0$, $E[X] = 0$.

(c) $E[X] = \int_5^\infty \frac{5}{x} dx = \infty.$

Exercise 41.

The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & 0 \leq x \leq 1 \\ 0 & \textit{otherwise.} \end{cases}$$

If $E[X] = \frac{3}{5}$, find a and b.

Solution : $\int_0^1 (a + bx^2)dx = 1$ or $a + \frac{b}{3} = 1$.

$\int_0^1 x(a + bx^2)dx = \frac{3}{5}$ or $\frac{a}{2} + \frac{b}{4} = 3/5$.

Hence, $a = \frac{3}{5}$, $b = \frac{6}{5}$.

Exercise 42.

Trains headed for destination A arrive at the train station at 15-minute intervals starting at 7 A.M., whereas trains headed for destination B arrive at 15-minute intervals starting at $7:05$ A.M.

(a) *If a certain passenger arrives at the station at a time uniformly distributed between 7 and 8 A.M. and then gets on the first train that arrives, what proportion of time does he or she go to destination A?*

(b) *What if the passenger arrives at a time uniformly distributed between $7:10$ and $8:10$ A.M.?*

Solution :

(a)

$$\begin{aligned} P\{\text{goes to } A\} &= P\{5 < X < 15 \text{ or } 20 < X < 30 \text{ or } 35 < X < 45 \text{ or } 50 < X < 60\}. \\ &= 2/3 \text{ since } X \text{ is uniform } (0, 60). \end{aligned}$$

(b) same answer as in (a).

Exercise 43.

If X is a normal random variable with parameters $\mu = 10$ and $\sigma^2 = 36$, compute

(a) $P\{X > 5\}$;

(b) $P\{4 < X < 16\}$;

(c) $P\{X < 8\}$;

(d) $P\{X < 20\}$;

(e) $P\{X > 16\}$.

Solution :

1. $\Phi(.8333) = .7977$
2. $2\Phi(1) - 1 = .6827$
3. $1 - \Phi(.3333) = .3695$
4. $\Phi(1.6667) = .9522$
5. $1 - \Phi(1) = .1587$

Exercise 44.

If 65 percent of the population of a large community is in favor of a proposed rise in school taxes, approximate the probability that a random sample of 100 people will contain

(a) *at least 50 who are in favor of the proposition;*

(b) *between 60 and 70 inclusive who are in favor;*

(c) *fewer than 75 in favor.*

Solution : Let X denote the number in favor. Then X is binomial with mean 65 and standard deviation $\sqrt{65(.35)} \approx 4.77$. Also let Z be a standard normal random variable.

(a) $P\{X \geq 50\} = P\{X \geq 49.5\} = P\{X - 65\}/4.77 \geq -15.5/4.77$
$\approx P\{Z \geq -3.25\} \approx .9994$

(b) $P\{59.5 \leq X \leq 70.5\} \approx P\{-5.5/4.77 \leq Z \leq 5.5/4.77\}$
$= 2P\{Z \leq 1.15\} - 1 \approx .75$

(c) $P\{X \leq 74.5\} \approx P\{Z \leq 9.5/4.77\} \approx .977$

Exercise

Exercise 45.

Suppose that the height, in inches, of a 25-year-old man is a normal random variable with parameters $\mu = 71$ and $\sigma^2 = 6.25$. What percentage of 25-year-old men are over 6 feet, 2 inches tall? What percentage of men in the 6-footer club are over 6 feet, 5 inches?

Exercise 46.

The lifetimes of interactive computer chips produced by a certain semiconductor manufacturer are normally distributed with parameters $\mu = 1.4 \times 10^6$ *hours and* $\sigma = 3 \times 10^5$ *hours. What is the approximate probability that a batch of 100 chips will contain at least* 20 *whose lifetimes are less than* 1.8×10^6?

Solution : With C denoting the life of a chip, and ϕ the standard normal distribution function we have

$$\begin{aligned} P\{C < 1.8 \times 10^6\} &= \phi\left(\frac{1.8 \times 10^6 - 1.4 \times 10^6}{3 \times 10^5}\right) \\ &= \phi(1.33) \\ &= .9082. \end{aligned}$$

Thus, if N is the number of the chips whose life is less than 1.8×10^6 then N is a binomial random variable with parameters $(100, .9082)$. Hence,

$$P\{N > 19.5\} \approx 1 - \phi\left(\frac{19.5 - 90.82}{90.82(.0918)}\right) = 1 - \phi(-24.7) \approx 1.$$

Exercise 47.

In $10,000$ *independent tosses of a coin, the coin landed on heads* 5800 *times. Is it reasonable to assume that the coin is not fair? Explain.*

Solution :

$$P\{X > 5,799.5\} = P\left\{Z > \frac{799.5}{\sqrt{2,500}}\right\}$$
$$= P\{Z > 15.99\} = \text{negligible.}$$

Exercise 48.

The time (in hours) required to repair a machine is an exponentially distributed random variable with parameter $\lambda = \frac{1}{2}$. What is

(a) *the probability that a repair time exceeds* 2 *hours?*

(b) *the conditional probability that a repair takes at least* 10 *hours, given that its duration exceeds* 9 *hours?*

Solution :

(a) e^{-1}

(b) $e^{-1/2}$

Exercise 49.

Suppose that the life distribution of an item has the hazard rate function $\lambda(t) = t^3, t > 0$. What is the probability that

(a) *the item survives to age 2?*

(b) *the item's lifetime is between .4 and 1.4?*

(c) *a 1-year-old item will survive to age 2?*

Solution :

(a) $1 - F(2) = \exp\left[-\int_0^2 t^3 dt\right] = e^{-4}$

(b) $\exp[-(.4)^4/4] - \exp[-(1.4)^4/4]$

(c) $\exp\left[-\int_1^2 t^3 dt\right] = e^{-15/4}$

Exercise 50.

If X is uniformly distributed over $(-1, 1)$, find

(a) $P\{|X| > \frac{1}{2}\}$;

(b) *the density function of the random variable $|X|$.*

Solution :

(a) $P\{|X| > 1/2\} = P\{X > 1/2\} + P\{X < -1/2\} = 1/2$

(b) $P\{|X| \leq a\} = P\{-a \leq X \leq a\} = a$, $0 < a < 1$.
Therefore, $f_{|X|}(a) = 1$, $0 < a < 1$.
That is, $|X|$ is uniform on $(0, 1)$.

Exercise 51.

If Y is uniformly distributed over $(0, 5)$, what is the probability that the roots of the equation $4x^2 + 4xY + Y + 2 = 0$ are both real?

Solution : For both roots to be real the discriminant $(4Y)^2 - 44(Y + 2)$ must be ≥ 0. That is, we need that $Y^2 \geq Y + 2$.

Now in the interval $0 < Y < 5$.

$$Y^2 \geq Y + 2 \iff Y \geq 2 \text{ and so}$$
$$P\{Y^2 \geq Y + 2\} = P\{Y \geq 2\} = 3/5.$$

Exercise 52.

If X is an exponential random variable with parameter $\lambda = 1$, compute the probability density function of the random variable Y defined by $Y = log X$.

Solution :

$$F_Y(y) = P\{\log X \leq y\}$$
$$= P\{X \leq e^y\} = F_X(e^y)$$

$$f_Y(y) = f_X(e^y)e^y = e^y e^{-e^y}$$

Exercise 53.

If X is uniformly distributed over $(0, 1)$, find the density function of $Y = e^X$.

Solution :

$$\begin{aligned} F_Y(y) &= P\{e^X \leq y\} \\ &= F_X(\log y). \end{aligned}$$

$$f_Y(y) = f_X(\log y)\frac{1}{y} = \frac{1}{y}, \quad 1 < y < e.$$

Exercise 54.

The standard deviation of X, denoted $SD(X)$, is given by

$$SD(X) = \sqrt{Var(X).}$$

Find $SD(aX + b)$ if X has variance σ^2.

Solution :

$$SD(aX + b) = \sqrt{Var(aX + b)} = \sqrt{a^2\sigma^2} = |a|\sigma.$$

Exercise 55.

Let X be a random variable that takes on values between 0 and c. That is, $P\{0 \leq X \leq c\} = 1$. Show that

$$Var(X) \leq \frac{c^2}{4}.$$

Solution : Since $0 \leq X \leq c$, it follows that $X^2 \leq cX$. Hence,

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \leq E[cX - (E[X])^2 \\ &= cE[X] - (E[X])^2 = E[X](c - E[X]) \\ &= c^2[\alpha(1-\alpha)] \quad \text{where } \alpha = E[X]/c \leq c^2/4 \end{aligned}$$

where the last inequality first uses the hypothesis that $P\{0 \leq X \leq c\} = 1$ to calculate that $0 \leq \alpha \leq 1$ and then uses calculus to show that $\underset{0 \leq \alpha \leq 1}{\text{maximum}}\ \alpha(1-\alpha) = 1/4$.

Exercise 56.

If X is a beta random variable with parameters a and b, show that

$$E[X] = \frac{a}{a+b},$$
$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

Solution : $(X - a)/(b - a)$.

References

1. Meyer, *Introductory Probability and Statistical Applications*, 2nd Edition, Oxford & IBH Publishing Company, India.
2. Sheldon Ross, *First Course in Probability*, Sixth Edition, Pearson Publisher, India.
3. Dimitri P. Bertsekas and John N. Tsitsiklis, *Introduction to Probability*, Athena Scientific, Belmont, Massachusetts.