Section 2.4

Cumulative Distribution Function (CDF)

Definition 2.11 Cumulative Distribution Function (CDF)

The cumulative distribution function *(CDF) of random variable* X *is*

$$F_X(x) = P[X \leq x].$$

Theorem 2.2

For any discrete random variable X with range $S_X = \{x_1, x_2, \ldots\}$ satisfying $x_1 \le x_2 \le \ldots$,

(a) $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.

(b) For all $x' \ge x$, $F_X(x') \ge F_X(x)$.

(c) For $x_i \in S_X$ and ϵ, an arbitrarily small positive number,

$$F_X(x_i) - F_X(x_i - \epsilon) = P_X(x_i).$$

(d) $F_X(x) = F_X(x_i)$ for all x such that $x_i \le x < x_{i+1}$.

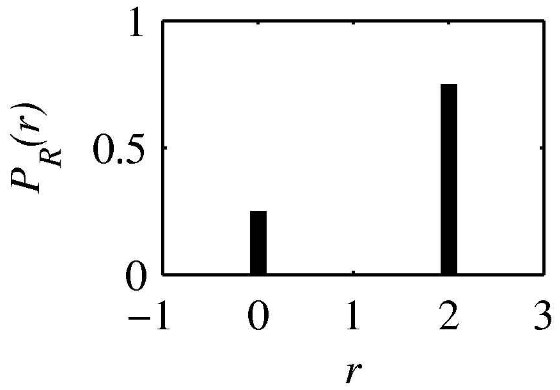
Theorem 2.3

For all $b \geq a$,

$$F_X(b) - F_X(a) = P[a < X \leq b].$$

Example 2.23 Problem

In Example 2.6, we found that random variable R has PMF

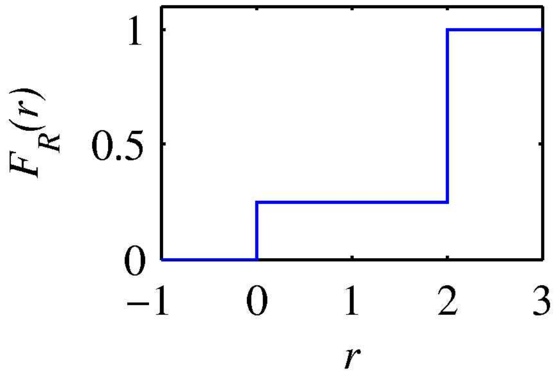


$$P_R(r) = \begin{cases} 1/4 & r = 0, \\ 3/4 & r = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Find and sketch the CDF of random variable R.

Example 2.23 Solution

From the PMF $P_R(r)$, random variable R has CDF



$$F_R(r) = P[R \le r] = \begin{cases} 0 & r < 0, \\ 1/4 & 0 \le r < 2, \\ 1 & r \ge 2. \end{cases}$$

Keep in mind that at the discontinuities $r = 0$ and $r = 2$, the values of $F_R(r)$ are the upper values: $F_R(0) = 1/4$, and $F_R(2) = 1$. Math texts call this the *right hand limit* of $F_R(r)$.

Section 2.5

Averages

Definition 2.12 Mode

A mode *of random variable* X *is a number* x_{mod} *satisfying* $P_X(x_{\text{mod}}) \geq P_X(x)$ *for all* x.

Definition 2.13 Median

A median, x_{med}, *of random variable* X *is a number that satisfies*

$$P\left[X < x_{\text{med}}\right] = P\left[X > x_{\text{med}}\right]$$

Definition 2.14 Expected Value

The expected value *of* X *is*

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

This is also called as "mean."

Example 2.25 Problem

For one quiz, 10 students have the following grades (on a scale of 0 to 10):

$$9, 5, 10, 8, 4, 7, 5, 5, 8, 7$$

Find the mean, the median, and the mode.

Example 2.25 Solution

The sum of the ten grades is 68. The mean value is 68/10 = 6.8. The median is 7 since there are four scores below 7 and four scores above 7. The mode is 5 since that score occurs more often than any other. It occurs three times.

Theorem 2.4

The Bernoulli (p) random variable X has expected value $E[X] = p$.

Proof: Theorem 2.4

$$E[X] = 0 \cdot P_X(0) + 1 P_X(1) = 0(1 - p) + 1(p) = p.$$

Example 2.26 Problem

Random variable R in Example 2.6 has PMF

$$P_R(r) = \begin{cases} 1/4 & r = 0, \\ 3/4 & r = 2, \\ 0 & \text{otherwise.} \end{cases}$$

What is $E[R]$?

Example 2.26 Solution

$$E[R] = \mu_R = 0 \cdot P_R(0) + 2P_R(2) = 0(1/4) + 2(3/4) = 3/2.$$

Theorem 2.5

The geometric (p) random variable X has expected value $E[X] = 1/p$.

Proof: Theorem 2.5

Let $q = 1 - p$. The PMF of X becomes

$$P_X(x) = \begin{cases} pq^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

The expected value $E[X]$ is the infinite sum

$$E[X] = \sum_{x=1}^{\infty} x P_X(x) = \sum_{x=1}^{\infty} x p q^{x-1}.$$

Applying the identity of Math Fact B.7, we have

$$E[X] = p \sum_{x=1}^{\infty} x q^{x-1} = \frac{p}{q} \sum_{x=1}^{\infty} x q^{x} = \frac{p}{q} \frac{q}{1-q^2} = \frac{p}{p^2} = \frac{1}{p}.$$

Math Fact B.7: If $|q| < 1$, $\sum_{i=1}^{\infty} i q^i = \frac{q}{(1-q)^2}$. 오타: $E[X] = \frac{p}{q}\frac{q}{1-q^2} \Rightarrow \frac{p}{q}\frac{q}{(1-q)^2}$

Theorem 2.6

The Poisson (α) random variable in Definition 2.10 has expected value $E[X] = \alpha$.

Proof: Theorem 2.6

$$E[X] = \sum_{x=0}^{\infty} x P_X(x) = \sum_{x=0}^{\infty} x \frac{\alpha^x}{x!} e^{-\alpha}.$$

We observe that $x/x! = 1/(x-1)!$ and also that the $x = 0$ term in the sum is zero. In addition, we substitute $\alpha^x = \alpha \cdot \alpha^{x-1}$ to factor α from the sum to obtain

$$E[X] = \alpha \sum_{x=1}^{\infty} \frac{\alpha^{x-1}}{(x-1)!} e^{-\alpha}.$$

Next we substitute $l = x - 1$, with the result

$$E[X] = \alpha \underbrace{\sum_{l=0}^{\infty} \frac{\alpha^l}{l!} e^{-\alpha}}_{1} = \alpha.$$

We can conclude that the marked sum equals 1 either by invoking the identity $e^{\alpha} = \sum_{l=0}^{\infty} \alpha^l / l!$ or by applying Theorem 2.1(b) to the fact that the marked sum is the sum of the Poisson PMF over all values in the range of the random variable.

Section 2.6

Functions of a Random Variable

Definition 2.15 Derived Random Variable

Each sample value y *of a* derived random variable Y *is a mathematical function* $g(x)$ *of a sample value* x *of another random variable* X. *We adopt the notation* $Y = g(X)$ *to describe the relationship of the two random variables.*

Example 2.27 Problem

The random variable X is the number of pages in a facsimile transmission. Based on experience, you have a probability model $P_X(x)$ for the number of pages in each fax you send. The phone company offers you a new charging plan for faxes: \$0.10 for the first page, \$0.09 for the second page, etc., down to \$0.06 for the fifth page. For all faxes between 6 and 10 pages, the phone company will charge \$0.50 per fax. (It will not accept faxes longer than ten pages.) Find a function $Y = g(X)$ for the charge in cents for sending one fax.

Example 2.27 Solution

The following function corresponds to the new charging plan.

$$Y = g(X) = \begin{cases} 10.5X - 0.5X^2 & 1 \leq X \leq 5 \\ 50 & 6 \leq X \leq 10 \end{cases}$$

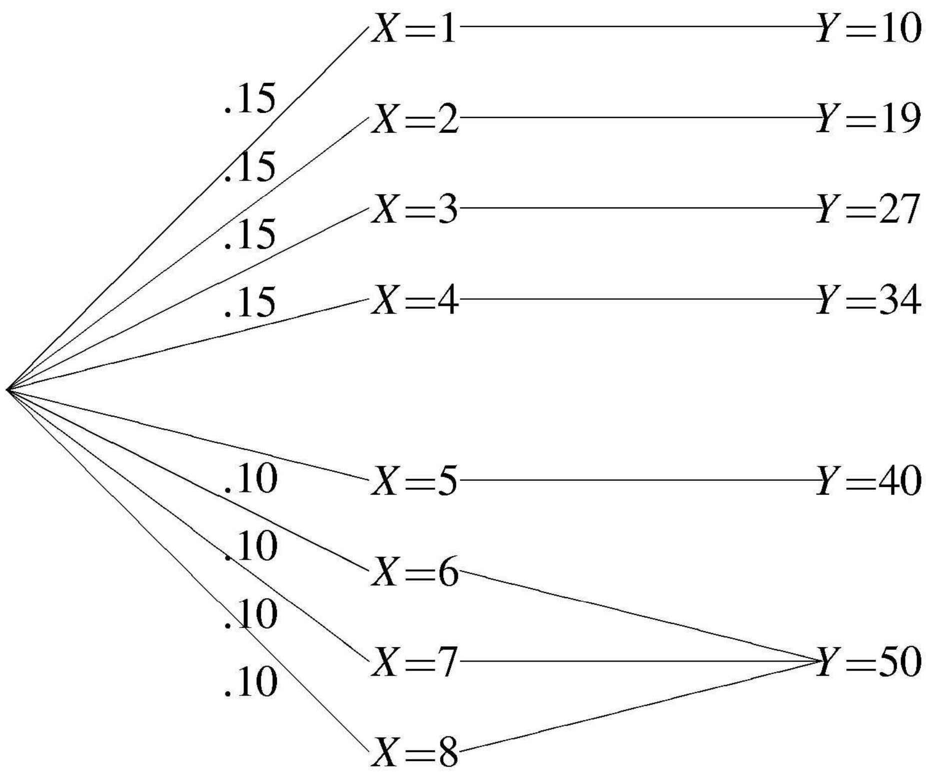
You would like a probability model $P_Y(y)$ for your phone bill under the new charging plan. You can analyze this model to decide whether to accept the new plan.

Theorem 2.9

For a discrete random variable X, the PMF of $Y = g(X)$ is

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x).$$

Figure 2.1



The derived random variable $Y = g(X)$ for Example 2.29.

Example 2.28 Problem

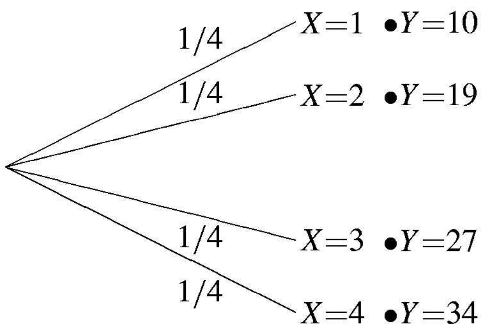
In Example 2.27, suppose all your faxes contain 1, 2, 3, or 4 pages with equal probability. Find the PMF and expected value of Y, the charge for a fax.

Example 2.28 Solution

From the problem statement, the number of pages X has PMF

$$P_X(x) = \begin{cases} 1/4 & x = 1, 2, 3, 4, \\ 0 & \text{otherwise.} \end{cases}$$

The charge for the fax, Y, has range $S_Y = \{10, 19, 27, 34\}$ corresponding to $S_X = \{1, 2, 3, 4\}$. The experiment can be described by the following tree. Here each value of Y results in a unique value of X. Hence, we can use Equation (2.66) to find $P_Y(y)$.

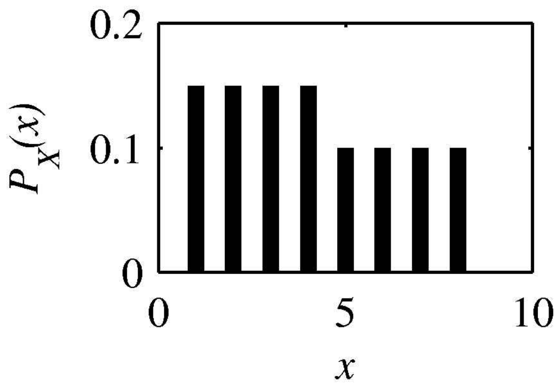


$$P_Y(y) = \begin{cases} 1/4 & y = 10, 19, 27, 34, \\ 0 & \text{otherwise.} \end{cases}$$

The expected fax bill is $E[Y] = (1/4)(10 + 19 + 27 + 34) = 22.5$ cents.

Example 2.29 Problem

Suppose the probability model for the number of pages X of a fax in Example 2.28 is



$$P_X(x) = \begin{cases} 0.15 & x = 1, 2, 3, 4 \\ 0.1 & x = 5, 6, 7, 8 \\ 0 & \text{otherwise} \end{cases}$$

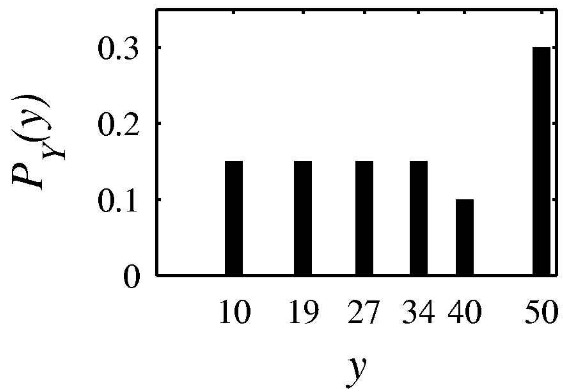
For the pricing plan given in Example 2.27, what is the PMF and expected value of Y, the cost of a fax?

Example 2.29 Solution

Now we have three values of X, specifically $(6,7,8)$, transformed by $g(\cdot)$ into $Y = 50$. For this situation we need the more general view of the PMF of Y, given by Theorem 2.9. In particular, $y_6 = 50$, and we have to add the probabilities of the outcomes $X = 6$, $X = 7$, and $X = 8$ to find $P_Y(50)$. That is,

$$P_Y(50) = P_X(6) + P_X(7) + P_X(8) = 0.30.$$

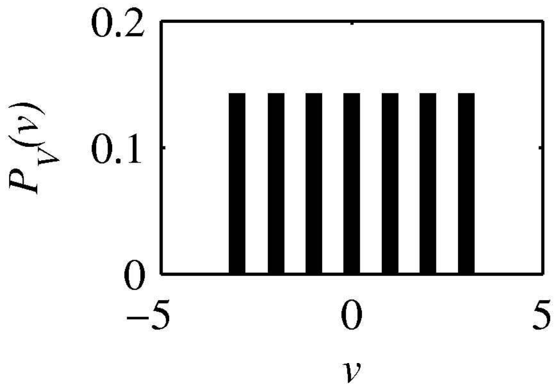
The steps in the procedure are illustrated in the diagram of Figure 2.1. Applying Theorem 2.9, we have



$$P_Y(y) = \begin{cases} 0.15 & y = 10, 19, 27, 34, \\ 0.10 & y = 40, \\ 0.30 & y = 50, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.30 Problem

The amplitude V (volts) of a sinusoidal signal is a random variable with PMF

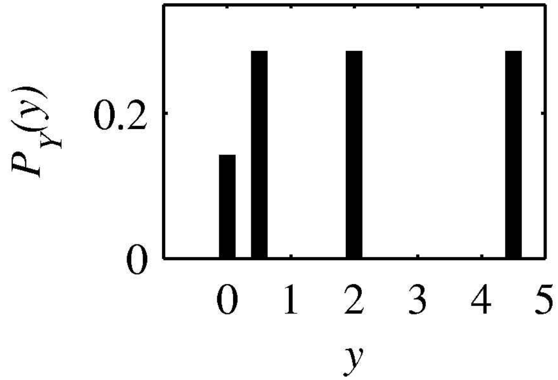


$$P_V(v) = \begin{cases} 1/7 & v = -3, -2, \ldots, 3 \\ 0 & \text{otherwise} \end{cases}$$

Let $Y = V^2/2$ watts denote the average power of the transmitted signal. Find $P_Y(y)$.

Example 2.30 Solution

The possible values of Y are $S_Y = \{0, 0.5, 2, 4.5\}$. Since $Y = y$ when $V = \sqrt{2y}$ or $V = -\sqrt{2y}$, we see that $P_Y(0) = P_V(0) = 1/7$. For $y = 1/2, 2, 9/2$, $P_Y(y) = P_V(\sqrt{2y}) + P_V(-\sqrt{2y}) = 2/7$. Therefore,



$$P_Y(y) = \begin{cases} 1/7 & y = 0, \\ 2/7 & y = 1/2, 2, 9/2, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 2.6

Monitor three phone calls and observe whether each one is a voice call or a data call. The random variable N is the number of voice calls. Assume N has PMF

$$P_N(n) = \begin{cases} 0.1 & n = 0, \\ 0.3 & n = 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases}$$

Voice calls cost 25 cents each and data calls cost 40 cents each. T cents is the cost of the three telephone calls monitored in the experiment.

(1) Express T as a function of N. (2) Find $P_T(t)$ and $E[T]$.

Section 2.7

Expected Value of a Derived Random Variable

Theorem 2.10

Given a random variable X with PMF $P_X(x)$ and the derived random variable $Y = g(X)$, the expected value of Y is

$$E[Y] = \mu_Y = \sum_{x \in S_X} g(x) P_X(x)$$

Proof: Theorem 2.10

From the definition of $E[Y]$ and Theorem 2.9, we can write

$$E[Y] = \sum_{y \in S_Y} y P_Y(y) = \sum_{y \in S_Y} y \sum_{x:g(x)=y} P_X(x) = \sum_{y \in S_Y} \sum_{x:g(x)=y} g(x) P_X(x),$$

where the last double summation follows because $g(x) = y$ for each x in the inner sum. Since $g(x)$ transforms each possible outcome $x \in S_X$ to a value $y \in S_Y$, the preceding double summation can be written as a single sum over over all possible values $x \in S_X$. That is,

$$E[Y] = \sum_{x \in S_X} g(x) P_X(x)$$

Example 2.31 Problem

In Example 2.28,

$$P_X(x) = \begin{cases} 1/4 & x = 1, 2, 3, 4, \\ 0 & \text{otherwise}, \end{cases}$$

and

$$Y = g(X) = \begin{cases} 10.5X - 0.5X^2 & 1 \le X \le 5, \\ 50 & 6 \le X \le 10. \end{cases}$$

What is $E[Y]$?

Example 2.31 Solution

Applying Theorem 2.10 we have

$$
\begin{aligned}
E[Y] &= \sum_{x=1}^{4} P_X(x)\, g(x) \\
&= (1/4)[(10.5)(1) - (0.5)(1)^2] + (1/4)[(10.5)(2) - (0.5)(2)^2] \\
&\quad + (1/4)[(10.5)(3) - (0.5)(3)^2] + (1/4)[(10.5)(4) - (0.5)(4)^2] \\
&= (1/4)[10 + 19 + 27 + 34] = 22.5 \text{ cents.}
\end{aligned}
$$

Theorem 2.11

For any random variable X,

$$E\left[X - \mu_X\right] = 0.$$

Proof: Theorem 2.11

Defining $g(X) = X - \mu_X$ and applying Theorem 2.10 yields

$$E[g(X)] = \sum_{x \in S_X} (x - \mu_X) P_X(x) = \sum_{x \in S_X} x P_X(x) - \mu_X \sum_{x \in S_X} P_X(x).$$

The first term on the right side is μ_X by definition. In the second term, $\sum_{x \in S_X} P_X(x) = 1$, so both terms on the right side are μ_X and the difference is zero.

Theorem 2.12

For any random variable X,

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

Example 2.32 Problem

Recall that in Examples 2.6 and 2.26, we found that R has PMF

$$P_R(r) = \begin{cases} 1/4 & r = 0, \\ 3/4 & r = 2, \\ 0 & \text{otherwise,} \end{cases}$$

and expected value $E[R] = 3/2$. What is the expected value of $V = g(R) = 4R + 7$?

Example 2.32 Solution

From Theorem 2.12,

$$E[V] = E[g(R)] = 4E[R] + 7 = 4(3/2) + 7 = 13.$$

We can verify this result by applying Theorem 2.10. Using the PMF $P_R(r)$ given in Example 2.6, we can write

$$E[V] = g(0)P_R(0) + g(2)P_R(2) = 7(1/4) + 15(3/4) = 13.$$

Example 2.33 Problem

In Example 2.32, let $W = h(R) = R^2$. What is $E[W]$?

Example 2.33 Solution

Theorem 2.10 gives

$$E[W] = \sum h(r) P_R(r) = (1/4)0^2 + (3/4)2^2 = 3.$$

Note that this is not the same as $h(E[W]) = (3/2)^2$.