Chapter 2
Discrete Random Variables

Chien-Kang Huang (黃乾綱)
臺大工科海洋系副教授

Outline

2.1 Definitions

Review of Probability Model

- It begins with a physical model of an experiment.
  - An experiment consists of a procedure and observations.
- The set of all possible observations, S, is the sample space of the experiment.
- S is the beginning of the mathematical probability model.
- In addition to S, the mathematical model includes a rule for assigning numbers between 0 and 1 to sets A in S.
- Thus for every $A \subset S$, the model gives us a probability $P[A]$, where $0 \leq P[A] \leq 1$.

Terminology

- We will examine probability models that assign numbers to the outcomes in the sample space.
- **Random variable**, X
  - That is, the observation
- **Range**, S_X
  - The set of possible values of X is the range of X
  - Analog to S, the set of all possible outcomes of an experiment.

Probability Model and Experiment

- A probability model always begins with an experiment. Each random variable is related directly to this experiment. There are three types of relationships.

1. The random variable is the observation.
   - Example 2.1
2. The random variable is a function of the observation.
   - Example 2.2
3. The random variable is a function of another random variable.
   - Example 2.3

Example 2.1 (The random variable is the observation)

Example 2.1

The experiment is to attach a photo detector to an optical fiber and count the number of photons arriving in a one microsecond time interval. Each observation is a random variable X. The range of X is $S_X = \{0, 1, 2, \ldots\}$. In this case, S_X, the range of X, and the sample space S are identical.

Example 2.2 (The random variable is a function of the observation)

Example 2.2

The experiment is to test six integrated circuits and after each test observe whether the circuit is accepted (a) or rejected (r). Each observation is a sequence of six letters where each letter is either a or r. For example, $s_8 =$ aaraaa. The sample space S consists of the 64 possible sequences. A random variable related to this experiment is N, the number of accepted circuits. For outcome s_8, $N = 5$ circuits are accepted. The range of N is $S_N = \{0, 1, \ldots, 6\}$.

Example 2.3 (The random variable is a function of another random variable)

Example 2.3

In Example 2.2, the net revenue R obtained for a batch of six integrated circuits is \$5 for each circuit accepted minus \$7 for each circuit rejected. (This is because for each bad circuit that goes out of the factory, it will cost the company \$7 to deal with the customer's complaint and supply a good replacement circuit.) When N circuits are accepted, $6-N$ circuits are rejected so that the net revenue R is related to N by the function

$$R = g(N) = 5N - 7(6-N) = 12N - 42 \text{ dollars.}$$

Since $S_N = \{0, \ldots, 6\}$, the range of R is

$$S_R = \{-42, -30, -18, -6, 6, 18, 30\}.$$

Definition 2.1 Random Variable

A random variable *consists of an experiment with a probability measure* $P[\cdot]$ *defined on a sample space* S *and a function that assigns a real number to each outcome in the sample space of the experiment.*

- A random variable is the result of an underlying experiment, but it also permit us to separate the experiment, in particular, the observations, from the process of assigning numbers to outcomes.
- In some definitions of experiments, the procedures contain variable parameters. In these experiments, there can be values of the parameters for which it is impossible to perform the observations specified in the experiments. In these cases, the experiments do not produce random variables. We refer to experiments with parameter settings that do not produce random variables as **improper experiments**.

Example 2.4 Problem

The procedure of an experiment is to fire a rocket in a vertical direction from the Earth's surface with initial velocity V km/h. The observation is T seconds, the time elapsed until the rocket returns to Earth. Under what conditions is the experiment improper?

Example 2.4 Solution

At low velocities, V, the rocket will return to Earth at a random time T seconds that depends on atmospheric conditions and small details of the rocket's shape and weight. However, when $V > v^* \approx 40{,}000$ km/hr, the rocket will not return to Earth. Thus, the experiment is improper when $V > v^*$ because it is impossible to perform the specified observation.

Notation

- On occasion, it is important to identify the random variable X by the function $X(s)$ that maps the sample outcomes s to the corresponding value of the random variable X.
- As needed we will write $\{X = x\}$ to emphasize that there is a set of sample points $s \in S$ for whith $X(s) = x$, that is, we have adopted the shorthand notation

$$\{X = x\} = \{s \in S \mid X(s) = x\}$$

- More random variables:
  - A, the number of students asleep in the next probability lecture;
  - C, the number of phone calls you answer in the next hour;
  - M, the number of minutes you wait until you next answer the phone.
  - Random variables A and C are discrete, M is continuous. (Chapter 3)

Definition 2.2 Discrete Random Variable

X *is a* discrete *random variable if the range of* X *is a countable set*

$$S_X = \{x_1, x_2, \ldots\}\,.$$

Definition 2.3 Finite Random Variable

X *is a* finite random variable *if the range is a finite set*

$$S_X = \{x_1, x_2, \cdots, x_n\}\,.$$

– By contrast a random variable Y that can take on any real number y in an interval $a \le y \le b$ is a continuous random variable.

- Often, but not always, a discrete random variable takes on integer values. An exception is the random variables related to your probability grade.
  - The experiment is to take this course and observe your grade.
  - At Rutgers, the sample space is

$$S = \{F, D, C, C^+, B, B^+, A\}.$$

  - The function $G(\cdot)$ that transform this sample space into a random variable, G, is
    $G(F) = 0, G(D) = 1, G(C) = 2, G(C^+) = 2.5, G(B) = 3, G(B^+) = 3.5, G(A) = 4$
  - G is a finite random variable. Its values are in the set $S_G = \{0, 1, 2, 2.5, 3, 3.5, 4\}$
  - Random variables allow us to computer averages. In the mathematics of probability, average are called expectations or expected values of random variables.

Example 2.5

Suppose we observe three calls at a telephone switch where voice calls (v) and data calls (d) are equally likely. Let X denote the number of voice calls, Y the number of data calls, and let $R = XY$. The sample space of the experiment and the corresponding values of the random variables X, Y, and R are

	Outcomes	ddd	ddv	dvd	dvv	vdd	vdv	vvd	vvv
	$P[\cdot]$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$
Random	X	0	1	1	2	1	2	2	3
Variables	Y	3	2	2	1	2	1	1	0
	R	0	2	2	2	2	2	2	0

Quiz 2.1

A student takes two courses. In each course, the student will earn a B with probability 0.6 or a C with probability 0.4, independent of the other course. To calculate a grade point average (GPA), a B is worth 3 points and a C is worth 2 points. The student's GPA is the sum of the GPA for each course divided by 2. Make a table of the sample space of the experiment and the corresponding values of the student's GPA, G.

Quiz 2.1 Solution

The sample space, probabilities and corresponding grades for the experiment are

Outcome	$P[\cdot]$	G
BB	0.36	3.0
BC	0.24	2.5
CB	0.24	2.5
CC	0.16	2

2.2 Probability Mass Function

Definition 2.4 Probability Mass Function (PMF)

The probability mass function *(PMF) of the discrete random variable* X *is*

$$P_X(x) = P[X = x]$$

- Recall that a discrete probability model assigns a number between 0 and 1 to each outcome in a sample space.
- When we have a discrete random variable X, we express the probability model as a probability mass function (PMF).
- The argument of PMF ranges over all the real numbers.

Note and Notation

- Note that $X = x$ is an event consisting of all outcomes s of the underlying experiment for which $X(s) = x$. On the other hand, $P_X(x)$ is a function ranging over all real numbers x. For any value of x, the function $P_X(x)$ is the probability of the event $X = x$.
- Notation

Notation	Meaning
X	The name of a random variable
x	A possible value of the random variable
$P_X(\cdot)$	The PMF of random variable X

Example 2.6 Problem

From Example 2.5, what is the PMF of R?

Example 2.5

Suppose we observe three calls at a telephone switch where voice calls (v) and data calls (d) are equally likely. Let X denote the number of voice calls, Y the number of data calls, and let $R = XY$. The sample space of the experiment and the corresponding values of the random variables X, Y, and R are

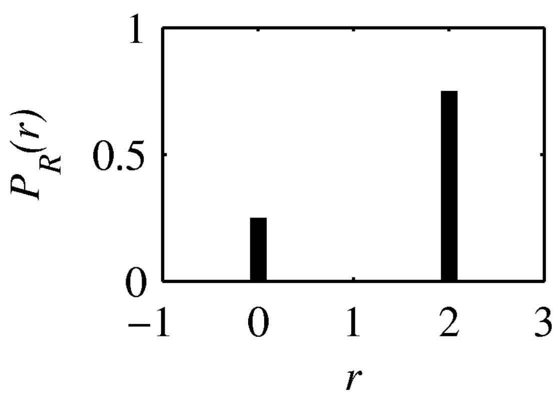
	Outcomes	ddd	ddv	dvd	dvv	vdd	vdv	vvd	vvv
	$P[\cdot]$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$
Random	X	0	1	1	2	1	2	2	3
Variables	Y	3	2	2	1	2	1	1	0
	R	0	2	2	2	2	2	2	0

Example 2.6 Solution

From Example 2.5, we see that $R = 0$ if either outcome, DDD or VVV, occurs so that

$$P[R=0] = P[DDD] + P[VVV] = 1/4.$$

For the other six outcomes of the experiment, $R = 2$ so that $P[R = 2] = 6/8$. The PMF of R is



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

The final line is necessary to specify the function at all other numbers.
It is helpful to keep this part of the definition in mind when working with the PMF.

Example 2.7 Problem

When the basketball player Wilt Chamberlain shot two free throws, each shot was equally likely either to be good (g) or bad (b). Each shot that was good was worth 1 point. What is the PMF of X, the number of points that he scored?

Example 2.7 Solution

There are four outcomes of this experiment: gg, gb, bg, and bb. A simple tree diagram indicates that each outcome has probability $1/4$. The random variable X has three possible values corresponding to three events:

$$\{X = 0\} = \{bb\}, \qquad \{X = 1\} = \{gb, bg\}, \qquad \{X = 2\} = \{gg\}.$$

Since each outcome has probability $1/4$, these three events have probabilities

$$P[X = 0] = 1/4, \qquad P[X = 1] = 1/2, \qquad P[X = 2] = 1/4.$$

We can express the probabilities of these events as the probability mass function

Theorem 2.1

For a discrete random variable X with PMF $P_X(x)$ and range S_X:

(a) For any x, $P_X(x) \geq 0$.

(b) $\sum_{x \in S_X} P_X(x) = 1$.

(c) For any event $B \subset S_X$, the probability that X is in the set B is

$$P[B] = \sum_{x \in B} P_X(x).$$

Proof: Theorem 2.1

All three properties are consequences of the axioms of probability (Section 1.3). First, $P_X(x) \geq 0$ since $P_X(x) = P[X = x]$. Next, we observe that every outcome $s \in S$ is associated with a number $x \in S_X$. Therefore, $P[x \in S_X] = \sum_{x \in S_X} P_X(x) = P[s \in S] = P[S] = 1$. Since the events $\{X = x\}$ and $\{X = y\}$ are disjoint when $x \neq y$, B can be written as the union of disjoint events $B = \bigcup_{x \in B} \{X = x\}$. Thus we can use Axiom 3 (if B is countably infinite) or Theorem 1.4 (if B is finite) to write

$$P[B] = \sum_{x \in B} P[X = x] = \sum_{x \in B} P_X(x).$$

Quiz 2.2

The random variable N has PMF

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

Find

(1) The value of the constant c

(2) $P[N = 1]$

(3) $P[N \geq 2]$

(4) $P[N > 3]$

Quiz 2.2 Solution

(1) To find c, we recall that the PMF must sum to 1. That is,

$$\sum_{n=1}^{3} P_N(n) = c\left(1 + \frac{1}{2} + \frac{1}{3}\right) = 1$$

This implies $c = 6/11$. Now that we have found c, the remaining parts are straightforward.

(2) $P[N = 1] = P_N(1) = c = 6/11$

(3) $P[N \geq 2] = P_N(2) + P_N(3) = c/2 + c/3 = 5/11$

(4) $P[N > 3] = \sum_{n=4}^{\infty} P_N(n) = 0$

2.3 Families of Discrete Random Variables

6 families of discrete random variables

Bernoulli, Geometric, Binomial, Pascal,
Discrete Uniform, Poisson

- In this section, we define 6 families of discrete random variables.
- There is one formula for the PMF of all the random variables in a family. Depending on the family, the PMF formula contains one or two parameters. By assigning numerical values to the parameters, we obtain a specific random variable.
  - *Binomial*(n, p) → *Binomial* $(7, 0.1)$, given $n = 7$, $p = 0.1$.
- Appendix A summarizes important properties of 17 families of random variables.

Example 2.8

- Consider the following experiments:
  1. Flip a coin and let it land on a table. Observe whether the side facing up is heads or tails. Let X be the number of heads observed.
  2. Select a student at random and find out her telephone number. Let X = 0 if the last digit is even. Otherwise, let X = 1.
  3. Observe one bit transmitted by a modem that is downloading a file from the Internet. Let X be the value of the bit (0 or 1).
- All three experiments lead to the probability mass function

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

Definition 2.5 Bernoulli (p) Random Variable

X *is a* Bernoulli (p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \textit{otherwise} \end{cases}$$

where the parameter p *is in the range* $0 < p < 1$.

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

Example 2.9 Problem

Suppose you test one circuit. With probability p, the circuit is rejected. Let X be the number of rejected circuits in one test. What is $P_X(x)$?

Example 2.9 Solution

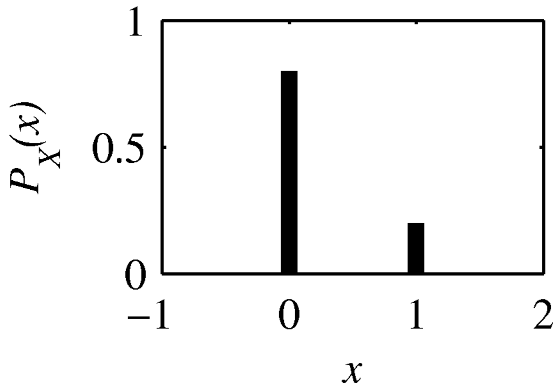
Because there are only two outcomes in the sample space, $X = 1$ with probability p and $X = 0$ with probability $1 - p$.

$$P_X(x) = \begin{cases} 1 - p & x = 0 \\ p & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, the number of circuits rejected in one test is a Bernoulli (p) random variable.

Example 2.10

If there is a 0.2 probability of a reject,



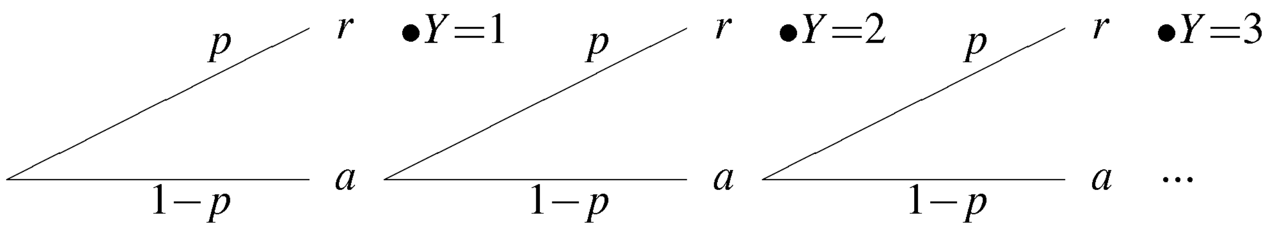
$$P_X(x) = \begin{cases} 0.8 & x = 0 \\ 0.2 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Example 2.11 Problem

In a test of integrated circuits there is a probability p that each circuit is rejected. Let Y equal the number of tests up to and including the first test that discovers a reject. What is the PMF of Y?

Example 2.11 Solution

The procedure is to keep testing circuits until a reject appears. Using a to denote an accepted circuit and r to denote a reject, the tree is



From the tree, we see that $P[Y = 1] = p$, $P[Y = 2] = p(1 - p)$, $P[Y = 3] = p(1 - p)^2$, and, in general, $P[Y = y] = p(1 - p)^{y-1}$. Therefore,

$$P_Y(y) = \begin{cases} p(1-p)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Y is referred to as a *geometric random variable* because the probabilities in the PMF constitute a geometric series.

Definition 2.6 Geometric (p) Random Variable

X *is a* geometric (p) *random variable if the PMF of* X *has the form*

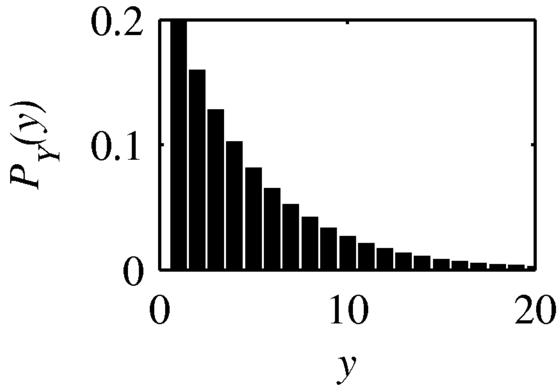
$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots \\ 0 & \textit{otherwise.} \end{cases}$$

where the parameter p *is in the range* $0 < p < 1$*.*

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

Example 2.12

If there is a 0.2 probability of a reject,



$$P_Y(y) = \begin{cases} (0.2)(0.8)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Example 2.13 Problem

Suppose we test n circuits and each circuit is rejected with probability p independent of the results of other tests. Let K equal the number of rejects in the n tests. Find the PMF $P_K(k)$.

Example 2.13 Solution

Adopting the vocabulary of Section 1.9, we call each discovery of a defective circuit a *success*, and each test is an independent trial with success probability p. The event $K = k$ corresponds to k successes in n trials, which we have already found, in Equation (1.18), to be the binomial probability

$$P_K(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

K is an example of a *binomial random variable*.

Definition 2.7 Binomial (n, p) Random Variable

X *is a* binomial (n, p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

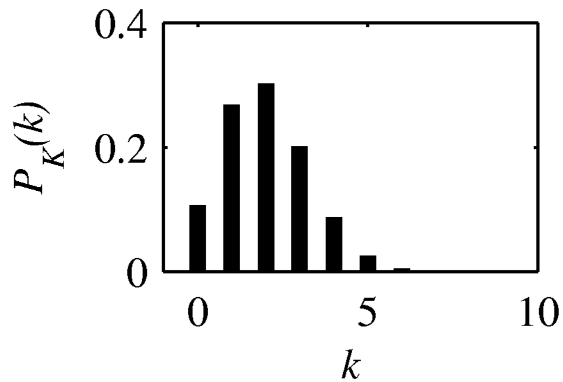
where $0 < p < 1$ *and* n *is an integer such that* $n \geq 1$.

- When ever we have a sequence of n independent trials each with success probability p, the number of successes is a binomial random variable.
- *Bernoulli*(p) random variable is a *binomial*$(n = 1, p)$ random variable.

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \binom{n}{x} = 0, \forall x \notin \{0, 1, \cdots, n\}$$

Example 2.14

If there is a 0.2 probability of a reject and we perform 10 tests,



$$P_K(k) = \binom{10}{k}(0.2)^k(0.8)^{10-k}.$$

Example 2.15 Problem

Suppose you test circuits until you find k rejects. Let L equal the number of tests. What is the PMF of L?

Example 2.15 Solution

For large values of k, the tree becomes difficult to draw. Once again, we view the tests as a sequence of independent trials where finding a reject is a success. In this case, $L = l$ if and only if there are $k-1$ successes in the first $l-1$ trials, *and* there is a success on trial l so that

$$P[L = l] = P\Bigg[\underbrace{k-1 \text{ rejects in } l-1 \text{ attempts}}_{A}, \underbrace{\text{success on attempt } l}_{B}\Bigg]$$

The events A and B are independent since the outcome of attempt l is not affected by the previous $l-1$ attempts. Note that $P[A]$ is the binomial probability of $k-1$ successes in $l-1$ trials so that

$$P[A] = \binom{l-1}{k-1} p^{k-1}(1-p)^{l-1-(k-1)}$$

Finally, since $P[B] = p$,

$$P_L(l) = P[A]\,P[B] = \binom{l-1}{k-1} p^{k}(1-p)^{l-k}$$

L is an example of a *Pascal* random variable.

Definition 2.8 Pascal (k, p) Random Variable

Negative Binomial (k, p) Random Variable

X is a Pascal (k, p) *random variable if the PMF of X has the form*

$$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

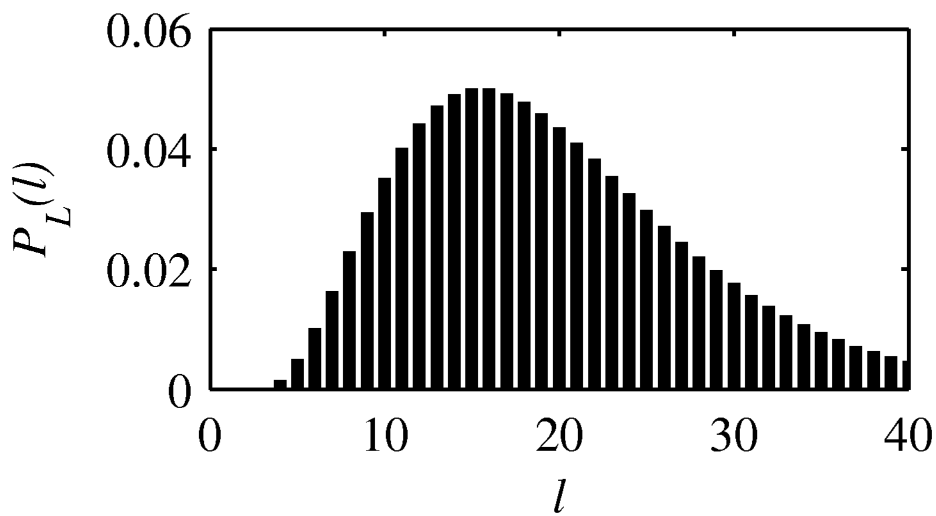
where $0 < p < 1$ and k is an integer such that $k \geq 1$.

- For a sequence of n independent trials each with success probability p, a Pascal random variable is the number of trials up to and including the k-th success.
- *geometric* (p) random variable is a *Pascal* $(k = 1, p)$ random variable.

$$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

Example 2.16

If there is a 0.2 probability of a reject and we seek four defective circuits, the random variable L is the number of tests necessary to find the four circuits. The PMF is



$$P_L(l) = \binom{l-1}{3}(0.2)^4(0.8)^{l-4}.$$

Example 2.17

In an experiment with equiprobable outcomes, the random variable N has the range $S_N = \{k, k+1, k+2, \cdots, l\}$, where k and l are integers with $k < l$. The range contains $l-k+1$ numbers, each with probability $1/(l-k+1)$. Therefore, the PMF of N is

$$P_N(n) = \begin{cases} 1/(l-k+1) & n = k, k+1, k+2, \ldots, l \\ 0 & \text{otherwise} \end{cases}$$

N is an example of a *discrete uniform* random variable.

Discrete Uniform (k, l) Random

Definition 2.9 Variable

X *is a* discrete uniform (k, l) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, k+2, \ldots, l \\ 0 & \textit{otherwise} \end{cases}$$

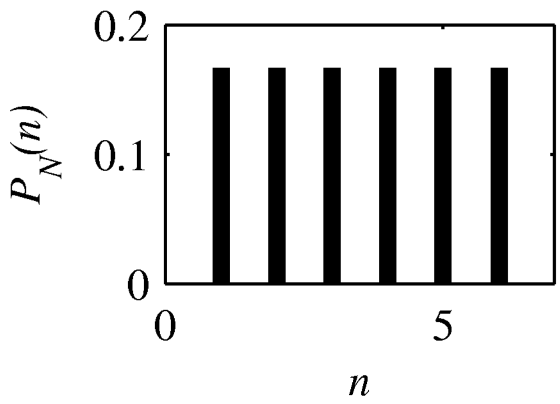
where the parameters k *and* l *are integers such that* $k < l$.

- X is uniformly distributed between k and l.

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, k+2, \cdots, l \\ 0 & \text{otherwise} \end{cases}$$

Example 2.18

Roll a fair die. The random variable N is the number of spots that appears on the side facing up. Therefore, N is a discrete uniform $(1, 6)$ random variable and



$$P_N(n) = \begin{cases} 1/6 & n = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise.} \end{cases}$$

Poisson Random Variable

- The probability model of a Poisson random variable describes phenomena that occur randomly in time.
- While the time of each occurrence is complete random, there is a known average number of occurrences per unit time.
- The Poisson model is used widely in many fields
  - Arrival of information requests at the WWW server
  - Initiation of telephone calls
  - Emission of particle from a radio active source

Definition 2.10 Poisson (α) Random Variable

X *is a* Poisson (α) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & \textit{otherwise}, \end{cases}$$

where the parameter α *is in the range* $\alpha > 0$.

More

- To describe a Poisson random variable, we will call the occurrence of the phenomenon of interest an **arrival**.
- λ: average rate (arrivals per second), T: (seconds)
  $\alpha = \lambda T$

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0,1,2,\cdots \\ 0 & \text{otherwise.} \end{cases}$$

Poisson ⇔ Binomial

Example 2.19 Problem

The number of hits at a Web site in any time interval is a Poisson random variable. A particular site has on average $\lambda = 2$ hits per second. What is the probability that there are no hits in an interval of 0.25 seconds? What is the probability that there are no more than two hits in an interval of one second?

Example 2.19 Solution

In an interval of 0.25 seconds, the number of hits H is a Poisson random variable with $\alpha = \lambda T = $ (2 hits/s) $\times$ (0.25 s) $=$ 0.5 hits. The PMF of H is

$$P_H(h) = \begin{cases} 0.5^h e^{-0.5}/h! & h = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

The probability of no hits is

$$P[H = 0] = P_H(0) = (0.5)^0 e^{-0.5}/0! = 0.607.$$

In an interval of 1 second, $\alpha = \lambda T = $ (2 hits/s) $\times$ (1s) $=$ 2 hits. Letting J denote the number of hits in one second, the PMF of J is

$$P_J(j) = \begin{cases} 2^j e^{-2}/j! & j = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

To find the probability of no more than two hits, we note that $\{J \le 2\} = \{J = 0\} \cup \{J = 1\} \cup \{J = 2\}$ is the union of three mutually exclusive events. Therefore,

$$\begin{aligned} P[J \le 2] &= P[J = 0] + P[J = 1] + P[J = 2] \\ &= P_J(0) + P_J(1) + P_J(2) \\ &= e^{-2} + 2^1 e^{-2}/1! + 2^2 e^{-2}/2! = 0.677. \end{aligned}$$

Example 2.20 Problem

The number of database queries processed by a computer in any 10-second interval is a Poisson random variable, K, with $\alpha = 5$ queries. What is the probability that there will be no queries processed in a 10-second interval? What is the probability that at least two queries will be processed in a 2-second interval?

Example 2.20 Solution

The PMF of K is

$$P_K(k) = \begin{cases} 5^k e^{-5}/k! & k = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Therefore $P[K = 0] = P_K(0) = e^{-5} = 0.0067$. To answer the question about the 2-second interval, we note in the problem definition that $\alpha = 5$ queries $= \lambda T$ with $T = 10$ seconds. Therefore, $\lambda = 0.5$ queries per second. If N is the number of queries processed in a 2-second interval, $\alpha = 2\lambda = 1$ and N is the Poisson (1) random variable with PMF

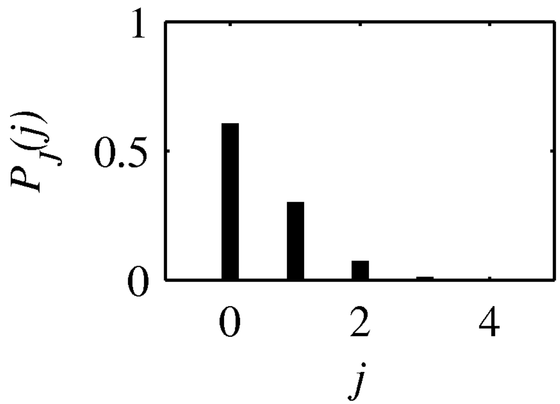
$$P_N(n) = \begin{cases} e^{-1}/n! & n = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$P[N \geq 2] = 1 - P_N(0) - P_N(1) = 1 - e^{-1} - e^{-1} = 0.264.$$

Example 2.21

Calls arrive at random times at a telephone switching office with an average of $\lambda = 0.25$ calls/second. The PMF of the number of calls that arrive in a $T = 2$-second interval is the Poisson (0.5) random variable with PMF

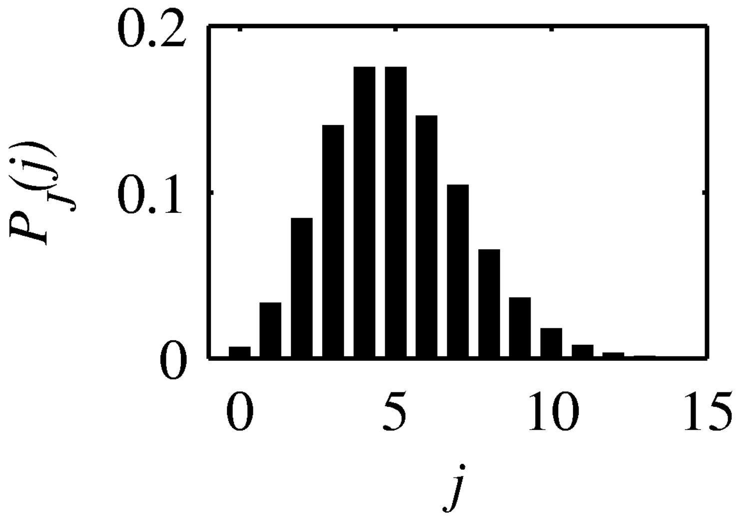


$$P_J(j) = \begin{cases} (0.5)^j e^{-0.5}/j! & j = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Note that we obtain the same PMF if we define the arrival rate as $\lambda = 60 \cdot 0.25 = 15$ calls per minute and derive the PMF of the number of calls that arrive in $2/60 = 1/30$ minutes.

Example 2.22

Calls arrive at random times at a telephone switching office with an average of $\lambda = 0.25$ calls per second. The PMF of the number of calls that arrive in any $T = 20$-second interval is the Poisson (5) random variable with PMF



$$P_J(j) = \begin{cases} 5^j e^{-5}/j! & j = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 2.3

Each time a modem transmits one bit, the receiving modem analyzes the signal that arrives and decides whether the transmitted bit is 0 or 1. It makes an error with probability p, independent of whether any other bit is received correctly.

(1) If the transmission continues until the receiving modem makes its first error, what is the PMF of X, the number of bits transmitted?

(2) If $p = 0.1$, what is the probability that $X = 10$? What is the probability that $X \geq 10$?

(3) If the modem transmits 100 bits, what is the PMF of Y, the number of errors?

(4) If $p = 0.01$ and the modem transmits 100 bits, what is the probability of $Y = 2$ errors at the receiver? What is the probability that $Y \leq 2$?

(5) If the transmission continues until the receiving modem makes three errors, what is the PMF of Z, the number of bits transmitted?

(6) If $p = 0.25$, what is the probability of $Z = 12$ bits transmitted?

Quiz 2.3 Solution

Decoding each transmitted bit is an independent trial where we call a bit error a "success." Each bit is in error, that is, the trial is a success, with probability p. Now we can interpret each experiment in the generic context of independent trials.

(1) The random variable X is the number of trials up to and including the first success. Similar to Example 2.11, X has the geometric PMF

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

(2) If $p = 0.1$, then the probability exactly 10 bits are sent is

$$P[X = 10] = P_X(10) = (0.1)(0.9)^9 = 0.0387$$

The probability that at least 10 bits are sent is $P[X \geq 10] = \sum_{x=10}^{\infty} P_X(x)$. This sum is not too hard to calculate. However, its even easier to observe that $X \geq 10$ if the first 10 bits are transmitted correctly. That is,

$$P[X \geq 10] = P\left[\text{first 10 bits are correct}\right] = (1-p)^{10}$$

For $p = 0.1$, $P[X \geq 10] = 0.9^{10} = 0.3487$.

(3) The random variable Y is the number of successes in 100 independent trials. Just as in Example 2.13, Y has the binomial PMF

$$P_Y(y) = \binom{100}{y} p^y (1-p)^{100-y}$$

[Continued]

Quiz 2.3 Solution (continued)

If $p = 0.01$, the probability of exactly 2 errors is

$$P[Y = 2] = P_Y(2) = \binom{100}{2}(0.01)^2(0.99)^{98} = 0.1849$$

(4) The probability of no more than 2 errors is

$$\begin{aligned}P[Y \le 2] &= P_Y(0) + P_Y(1) + P_Y(2)\\ &= (0.99)^{100} + 100(0.01)(0.99)^{99} + \binom{100}{2}(0.01)^2(0.99)^{98}\\ &= 0.9207\end{aligned}$$

(5) Random variable Z is the number of trials up to and including the third success. Thus Z has the Pascal PMF (see Example 2.15)

$$P_Z(z) = \binom{z-1}{2} p^3 (1-p)^{z-3}$$

Note that $P_Z(z) > 0$ for $z = 3, 4, 5, \ldots$.

(6) If $p = 0.25$, the probability that the third error occurs on bit 12 is

$$P_Z(12) = \binom{11}{2}(0.25)^3(0.75)^9 = 0.0645$$

2.4 Cumulative Distribution Function (CDF)

Definition 2.11 Cumulative Distribution Function (CDF)

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

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

- For any real number x, the CDF is the probability that the random variable X is no larger than x.
- All random variables have cumulative distribution functions but only discrete random variables have probability mass function.

Theorem 2.2

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

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

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

(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 $x_i \leq x < x_{i+1}$.

Equivalent Statements for Theorem 2.2

(a) Going from left to right on the x-axis, $F_X(x)$ starts at zero and ends at one.

(b) The CDF never decreases as it goes from left to right.

(c) For a discrete random variable X, there is a jump (discontinuity) at each value of $x_i \in S_X$. The height of the jump at x_i is $P_X(x_i)$.

(d) Between jumps, the graph of the CDF of the discrete random variable X is a horizontal line.

Theorem 2.3

For all $b \geq a$,

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

Proof: Theorem 2.3

To prove this theorem, express the event $E_{ab} = \{a < X \leq b\}$ as a part of a union of disjoint events. Start with the event $E_b = \{X \leq b\}$. Note that E_b can be written as the union

$$E_b = \{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\} = E_a \cup E_{ab}$$

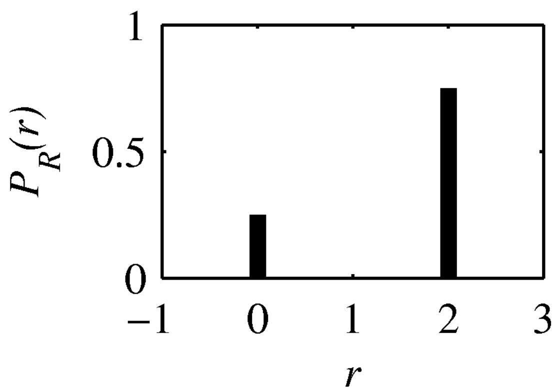
Note also that E_a and E_{ab} are disjoint so that $P[E_b] = P[E_a] + P[E_{ab}]$. Since $P[E_b] = F_X(b)$ and $P[E_a] = F_X(a)$, we can write $F_X(b) = F_X(a) + P[a < X \leq b]$. Therefore $P[a < X \leq b] = F_X(b) - F_X(a)$.

Note

- It is necessary to pay careful attention to the nature of inequalities, strict ($<$) or loose ($\leq$).
- The definition of the CDF contains a loose inequality, which means that function is continuous from the right.

Example 2.23 Problem

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

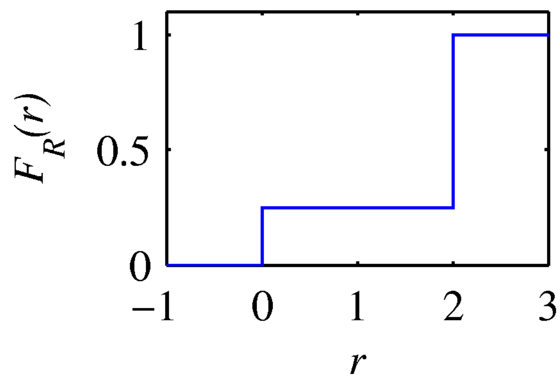


$$P_R(r) = \begin{cases} 1/4 & r = 0, \\ 3/4 & r = 2, \\ 0 & \text{otherwise.} \end{cases}$$

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 $F_R(r)$.

Note

- Consider any finite random variable X with possible values (nonzero probability) between $x_{\min}$ and $x_{\max}$.
- For this random variable, the numerical specification of the CDF begins with
$$F_X(x) = 0 \quad x < x_{\min}$$
and ends with
$$F_X(x) = 1 \quad x \geq x_{\max}$$

Example 2.24 Problem

In Example 2.11, let the probability that a circuit is rejected equal $p = 1/4$. The PMF of Y, the number of tests up to and including the first reject, is the geometric $(1/4)$ random variable with PMF

$$P_Y(y) = \begin{cases} (1/4)(3/4)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

What is the CDF of Y?

Example 2.24 Solution

Y is an infinite random variable, with nonzero probabilities for all positive integers. For any integer $n \geq 1$, the CDF is

$$F_Y(n) = \sum_{j=1}^{n} P_Y(j) = \sum_{j=1}^{n} \frac{1}{4}\left(\frac{3}{4}\right)^{j-1}.$$

Equation (2.41) is a geometric series. Familiarity with the geometric series is essential for calculating probabilities involving geometric random variables. Appendix B summarizes the most important facts. In particular, Math Fact B.4 implies $(1-x)\sum_{j=1}^{n} x^{j-1} = 1 - x^n$. Substituting $x = 3/4$, we obtain

$$F_Y(n) = 1 - \left(\frac{3}{4}\right)^{n}.$$

The complete expression for the CDF of Y must show $F_Y(y)$ for all integer *and noninteger* values of y.

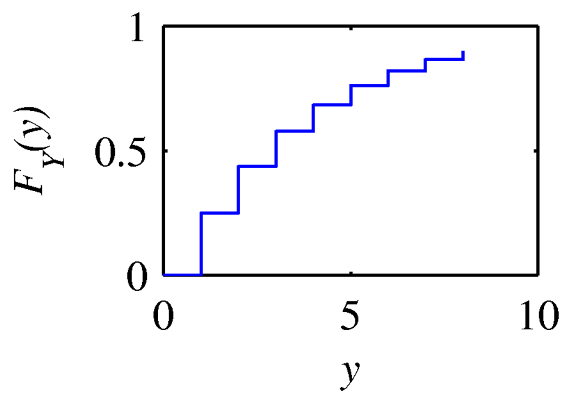
[Continued]

Example 2.24 Solution (continued)

For an integer-valued random variable Y, we can do this in a simple way using the *floor function* $\lfloor y \rfloor$, which is the largest integer less than or equal to y. In particular, if $n \leq y < n - 1$ for some integer n, then $n = \lfloor y \rfloor$ and

$$F_Y(y) = P[Y \leq y] = P[Y \leq n] = F_Y(n) = F_Y(\lfloor y \rfloor).$$

In terms of the floor function, we can express the CDF of Y as



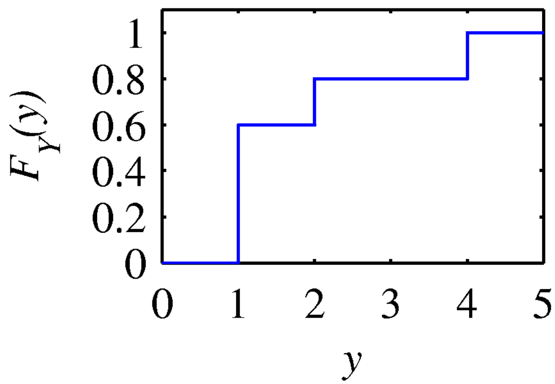
$$F_Y(y) = \begin{cases} 0 & y < 1, \\ 1 - (3/4)^{\lfloor y \rfloor} & y \geq 1. \end{cases}$$

To find the probability that Y takes a value in the set $\{4, 5, 6, 7, 8\}$, we refer to Theorem 2.3 and compute

$$P[3 < Y \leq 8] = F_Y(8) - F_Y(3) = (3/4)^3 - (3/4)^8 = 0.322.$$

Quiz 2.4

Use the CDF $F_Y(y)$ to find the following probabilities:



(1) $P[Y < 1]$

(2) $P[Y \leq 1]$

(3) $P[Y > 2]$

(4) $P[Y \geq 2]$

(5) $P[Y = 1]$

(6) $P[Y = 3]$

Quiz 2.4 Solution

Each of these probabilities can be read from the graph of the CDF $F_Y(y)$. However, we must keep in mind that when $F_Y(y)$ has a discontinuity at y_0, $F_Y(y)$ takes the upper value $F_Y(y_0^+)$.

(1) $P[Y < 1] = F_Y(1^-) = 0$

(2) $P[Y \le 1] = F_Y(1) = 0.6$

(3) $P[Y > 2] = 1 - P[Y \le 2] = 1 - F_Y(2) = 1 - 0.8 = 0.2$

(4) $P[Y \ge 2] = 1 - P[Y < 2] = 1 - F_Y(2^-) = 1 - 0.6 = 0.4$

(5) $P[Y = 1] = P[Y \le 1] - P[Y < 1] = F_Y(1^+) - F_Y(1^-) = 0.6$

(6) $P[Y = 3] = P[Y \le 3] - P[Y < 3] = F_Y(3^+) - F_Y(3^-) = 0.8 - 0.8 = 0$

2.5 Averages

Averages

- The average value of a collection of numerical observations is a statistic of the collection, a single number (parameter) that describes the entire collection.
- **Mean**: adding up all the numbers in the collection and dividing by the number of terms in the sum.
- **Median**: the median is a number in the middle of the set of numbers, in the sense that an equal number of members of the set are below the median and above the median.
- **Mode**: the mode is the most common number in the collection of observations. If there are two or more numbers with this property, the collection of observations is called multimodal.

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.

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

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*

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

Comments on Mode and Median

- A random variable can have several modes?
- A random variable can have several medians?
- A random variable can have several modes or medians.

Definition 2.14 Expected Value

The expected value *of* X *is*

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

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

Comments on Expectation

- expected value ≡ expectation ≡ mean value
- μ_X looks like the center of mass. That is why $P_X(x)$ is called PMF.
- X: A random variable,
  n: number of independent trials,
  $x(1), \ldots, x(n)$: sample values, $x \in S_X$ occurs N_x times
  m_n: sample average

$$m_n = \frac{1}{n}\sum_{i=1}^{n} x(i) = \frac{1}{n}\sum_{x \in S_X} N_x x = \sum_{x \in S_X} \frac{N_x}{n} x$$

- Recall section 1.3 relative frequency

$$P[A] = \lim_{n\to\infty} \frac{N_A}{n} \longrightarrow P_X(x) = \lim_{n\to\infty} \frac{N_x}{n}$$

$$\lim_{n\to\infty} m_n = \sum_{x \in S_X} x P_x(x) = E[X]$$

Theorem 2.4, ~ 2.7

Random Variable	**Expected Value, $E[X]$**
Bernoulli (p)	p
Geometric (p)	$1 / p$
Poisson (α)	α
Binomial (n, p)	np
Pascal (k, p)	k / p
Discrete uniform (k, l)	$(k + l) / 2$

Theorem 2.4

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

Proof: Theorem 2.4

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

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

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

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

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

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.

Theorem 2.7

(a) For the binomial (n, p) random variable X of Definition 2.7,

$$E[X] = np.$$

(b) For the Pascal (k, p) random variable X of Definition 2.8,

$$E[X] = k/p.$$

(c) For the discrete uniform (k, l) random variable X of Definition 2.9,

$$E[X] = (k + l)/2.$$

Theorem 2.8

Perform n Bernoulli trials. In each trial, let the probability of success be α/n, where $\alpha > 0$ is a constant and $n > \alpha$. Let the random variable K_n be the number of successes in the n trials. As $n \to \infty$, $P_{K_n}(k)$ converges to the PMF of a Poisson (α) random variable.

$$\lim_{n\to\infty} Binomial\left(n, p = \frac{\alpha}{n}\right) = Poisson(\alpha)$$

$$\Downarrow$$

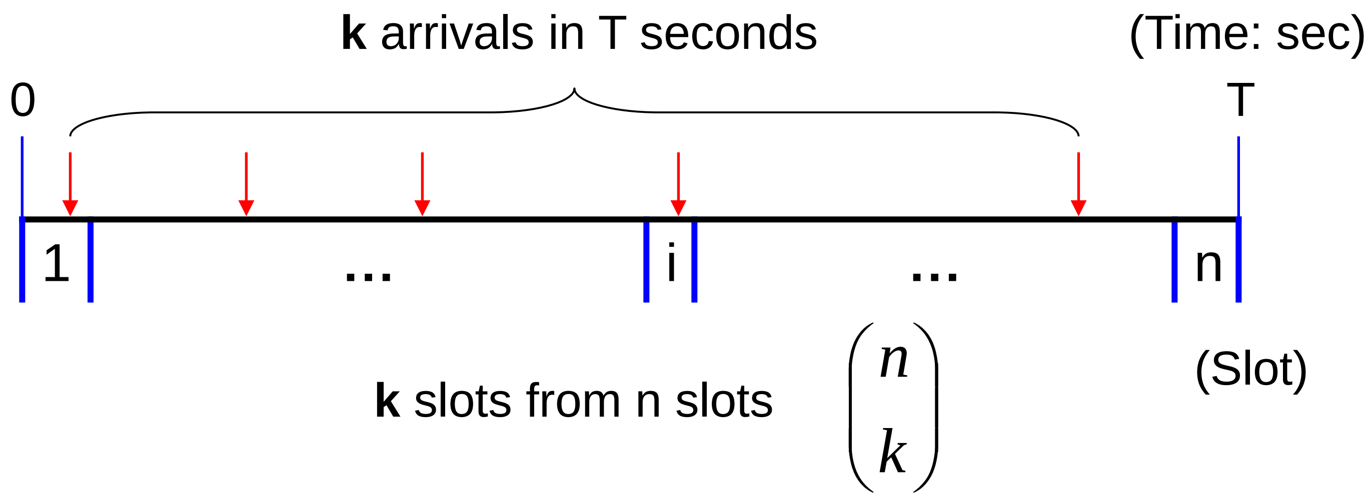
$$\lim_{n\to\infty}\left[\binom{n}{k}\left(\frac{\alpha}{n}\right)^k \left(1-\frac{\alpha}{n}\right)^{n-k}\right] = \frac{\alpha^k}{k!}e^{-\alpha}$$

Hints for Theorem 2.8

Poisson(α): Probability of **k** arrivals in **T** seconds

Given average arrival rate: α/T

$\alpha = \lambda T$



$$n \to \infty$$
$$\Downarrow$$
$$\frac{T}{n} \to 0$$

Binomial(n, p): Probability of select **k** slots from **n** slots

Given average probability of one arrival in one slot: **p**

p $= \lambda(T/n) = \alpha/n$

Proof of Theorem 2.8

- K_n is the binomial (n, α/n) random variable with PMF. For $k = 0, \ldots, n$

$$P_{K_n}(n) = \binom{n}{k}\left(\frac{\alpha}{n}\right)^k \left(1-\frac{\alpha}{n}\right)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{\alpha^k}{k!}\left(1-\frac{\alpha}{n}\right)^{n-k}$$

- As $n \to \infty$,

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = 1$$

$$\lim_{n\to\infty} \left(1-\frac{\alpha}{n}\right)^n = e^{-\alpha}$$

$$\lim_{n\to\infty} \left(1-\frac{\alpha}{n}\right)^k = 1$$

$$\lim_{n\to\infty} P_{K_n}(n) = \begin{cases} \dfrac{\alpha^k}{k!}e^{-\alpha} & k = 0,1,\cdots \\ 0 & \text{otherwise} \end{cases}$$

Supplements

$$\left(1+\frac{x}{n}\right)^n \to e^x \text{ as } n \to \infty$$

- Proof

$$\ln\left(1+\frac{x}{n}\right)^n = n\ln\left(1+\frac{x}{n}\right) = x\left[\frac{\ln\left(1+\frac{x}{n}\right)-\ln 1}{\frac{x}{n}}\right]$$

$$\lim_{n\to\infty} x\left[\frac{\ln\left(1+\frac{x}{n}\right)-\ln 1}{\frac{x}{n}}\right] = \lim_{n\to\infty} x\left[\frac{\ln(1+h)-\ln 1}{h}\right] = 1$$

$$\lim_{n\to\infty} x\left[\frac{\ln(t+h)-\ln t}{h}\right] = \frac{1}{t}$$

Quiz 2.5

The probability that a call is a voice call is $P[V] = 0.7$. The probability of a data call is $P[D] = 0.3$. Voice calls cost 25 cents each and data calls cost 40 cents each. Let C equal the cost (in cents) of one telephone call and find

(1) The PMF $P_C(c)$

(2) The expected value $E[C]$

Quiz 2.5 Solution

(1) With probability 0.7, a call is a voice call and $C = 25$. Otherwise, with probability 0.3, we have a data call and $C = 40$. This corresponds to the PMF

$$P_C(c) = \begin{cases} 0.7 & c = 25 \\ 0.3 & c = 40 \\ 0 & \text{otherwise} \end{cases}$$

(2) The expected value of C is

$$E[C] = 25(0.7) + 40(0.3) = 29.5 \text{ cents}$$

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

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

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

Example 2.28 Problem

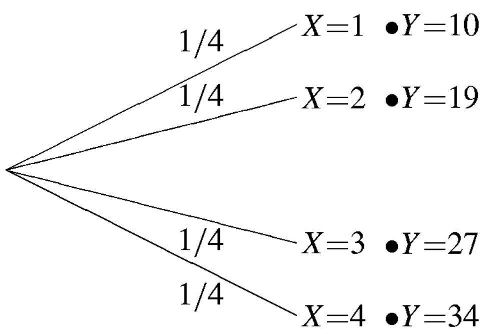
In Example 2.27, suppose all your faxes contain 1, 2, 3, or 4 pages with equal probability. Find the PMF and expected value of Y, the charge for a fax.

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 $P_Y(y)$.

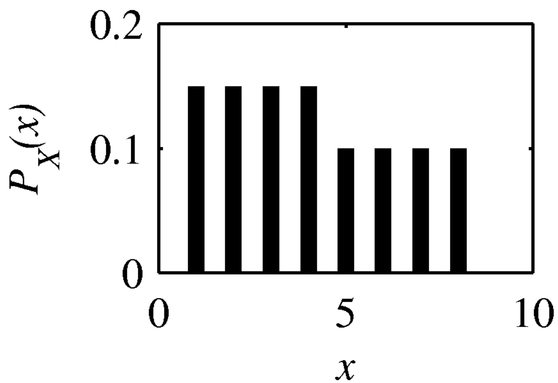


$$P_Y(y) = \begin{cases} 1/4 & y = 10, 19, 27, 34, \\ 0 & \text{otherwise.} \end{cases}$$

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

Example 2.29 Problem

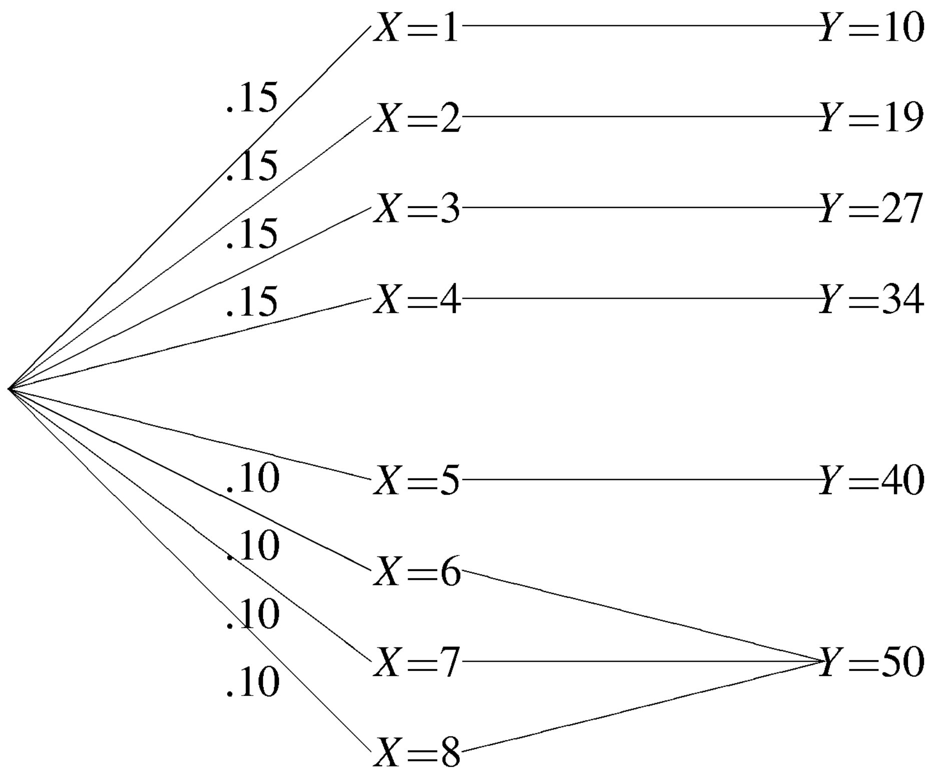
Suppose the probability model for the number of pages X of a fax in Example 2.28 is



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

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

Figure 2.1



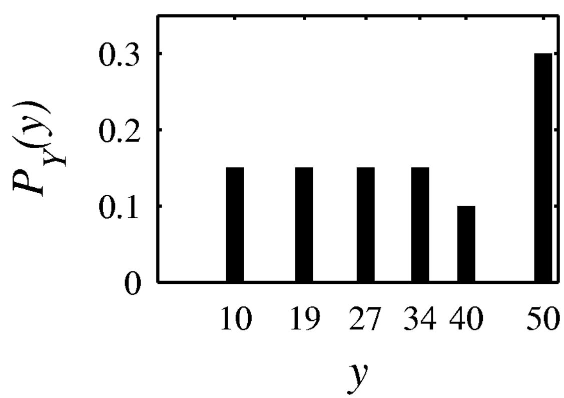
The derived random variable $Y = g(X)$ for Example 2.29.

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

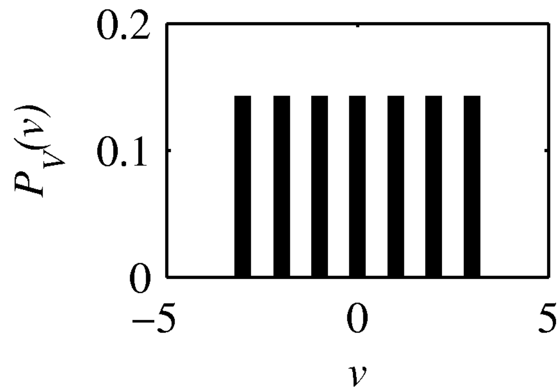


$$P_Y(y) = \begin{cases} 0.15 & y = 10, 19, 27, 34, \\ 0.10 & y = 40, \\ 0.30 & y = 50, \\ 0 & \text{otherwise.} \end{cases}$$

$E[Y] = 0.15\,(10 + 19 + 27 + 34) + 0.10\,(40) + 0.30\,(50) = 32.5$

Example 2.30 Problem

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

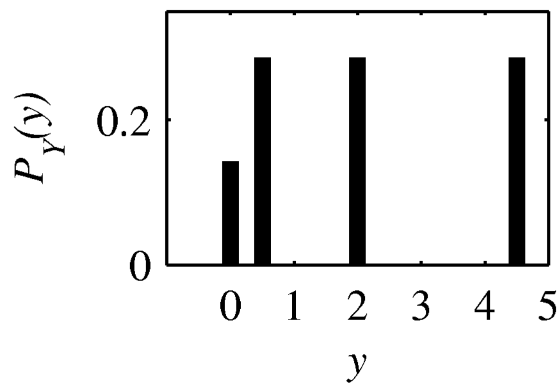


$$P_V(v) = \begin{cases} 1/7 & v = -3, -2, \ldots, 3 \\ 0 & \text{otherwise} \end{cases}$$

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.

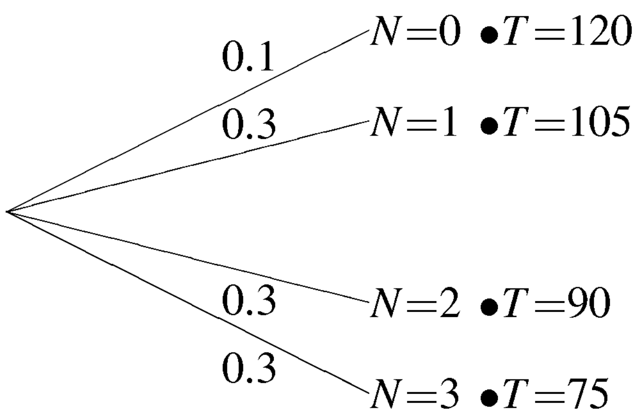
(2) Find $P_T(t)$ and $E[T]$.

Quiz 2.6 Solution

(1) As a function of N, the cost T is

$$T = 25N + 40(3 - N) = 120 - 15N$$

(2) To find the PMF of T, we can draw the following tree and write down the corresponding PMF:



$$P_T(t) = \begin{cases} 0.3 & t = 75, 90, 105 \\ 0.1 & t = 120 \\ 0 & \text{otherwise} \end{cases}$$

From the PMF $P_T(t)$, the expected value of T is

$$\begin{aligned} E[T] &= 75P_T(75) + 90P_T(90) + 105P_T(105) + 120P_T(120) \\ &= (75 + 90 + 105)(0.3) + 120(0.1) = 62 \end{aligned}$$

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,

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

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 $h(E[W]) = (3/2)^2$.

Quiz 2.7

The number of memory chips M needed in a personal computer depends on how many application programs, A, the owner wants to run simultaneously. The number of chips M and the number of application programs A are described by

$$M = \begin{cases} 4 & \text{chips for 1 program,} \\ 4 & \text{chips for 2 programs,} \\ 6 & \text{chips for 3 programs,} \\ 8 & \text{chips for 4 programs,} \end{cases} \qquad P_A(a) = \begin{cases} 0.1(5-a) & a = 1, 2, 3, 4, \\ 0 & \text{otherwise.} \end{cases}$$

(1) What is the expected number of programs $\mu_A = E[A]$?

(2) Express M, the number of memory chips, as a function $M = g(A)$ of the number of application programs A.

(3) Find $E[M] = E[g(A)]$. Does $E[M] = g(E[A])$?

Quiz 2.7 Solution

(1) Using Definition 2.14, the expected number of applications is

$$E[A] = \sum_{a=1}^{4} a P_A(a) = 1(0.4) + 2(0.3) + 3(0.2) + 4(0.1) = 2$$

(2) The number of memory chips is $M = g(A)$ where

$$g(A) = \begin{cases} 4 & A = 1, 2 \\ 6 & A = 3 \\ 8 & A = 4 \end{cases}$$

(3) By Theorem 2.10, the expected number of memory chips is

$$E[M] = \sum_{a=1}^{4} g(A) P_A(a) = 4(0.4) + 4(0.3) + 6(0.2) + 8(0.1) = 4.8$$

Since $E[A] = 2$, $g(E[A]) = g(2) = 4$. However, $E[M] = 4.8 \neq g(E[A])$. The two quantities are different because $g(A)$ is not of the form $\alpha A + \beta$.

2.8 Variance and Standard Deviation

Dispersion

- Average is one number that summarizes an entire probability model.
- Question
  - How typical is the average?
  - What are the chances of observing an event far from the average?
- A measure of dispersion is an answer to these questions wrapped up in a single number.
  - Small measure → observations are likely to be near the average.
- The most important measures of dispersion are the **standard deviation** and its close relative, the **variance**.

$$E[X-\mu_X]=0 \longrightarrow E\left[(X-\mu_X)^2\right]$$

Definition 2.16 Variance

The variance *of random variable* X *is*

$$\mathrm{Var}[X] = E\left[(X - \mu_X)^2\right].$$

Definition 2.17 Standard Deviation

The standard deviation *of random variable* X *is*

$$\sigma_X = \sqrt{\mathrm{Var}[X]}.$$

Comments

- σ_X has the same units as X.
- σ_X can be compared directly with the expected value.
- Informally, we think of outcomes within $\pm\sigma_X$ of μ_X as being in the center of the distribution.
- Informally, we think of sample values within σ_X of the expected value, $x \in [\mu_X - \sigma_X, \mu_X + \sigma_X]$ as "typical" values of X and other values as "unusual".

$$\mathrm{Var}[X] = \sigma_X^2 = \sum_{x \in S_X} (x - \mu_X)^2 P_X(x)$$

Theorem 2.13

$$\text{Var}[X] = E\left[X^2\right] - \mu_X^2 = E\left[X^2\right] - (E[X])^2$$

Proof: Theorem 2.13

Expanding the square in (2.91), we have

$$\begin{aligned}\text{Var}[X] &= \sum_{x \in S_X} x^2 P_X(x) - \sum_{x \in S_X} 2\mu_X x P_X(x) + \sum_{x \in S_X} \mu_X^2 P_X(x) \\ &= E[X^2] - 2\mu_X \sum_{x \in S_X} x P_X(x) + \mu_X^2 \sum_{x \in S_X} P_X(x) \\ &= E[X^2] - 2\mu_X^2 + \mu_X^2\end{aligned}$$

Definition 2.18 Moments

For random variable X:

(a) The nth moment *is $E[X^n]$.*

(b) The nth central moment *is $E[(X - \mu_X)^n]$.*

Comments

- $E[X]$: first moment
- $E[X^2]$: second moment
- The set of moments of X is a complete probability model. (Section 6.3, moment generating function)

Example 2.34 Problem

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

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

In Example 2.26, we calculated $E[R] = \mu_R = 3/2$. What is the variance of R?

Example 2.34 Solution

In order of increasing simplicity, we present three ways to compute $\mathrm{Var}[R]$.

- From Definition 2.16, define

$$W = (R - \mu_R)^2 = (R - 3/2)^2$$

The PMF of W is

$$P_W(w) = \begin{cases} 1/4 & w = (0 - 3/2)^2 = 9/4, \\ 3/4 & w = (2 - 3/2)^2 = 1/4, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\mathrm{Var}[R] = E[W] = (1/4)(9/4) + (3/4)(1/4) = 3/4.$$

Example 2.34

- Recall that in Examples 2.6, we found that R has PMF

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

In Example 2.26, we calculated $E[R] = \mu_R = 3/2$. What is the variance of R?

Solution:

- To apply Theorem 2.13, we find that
- $E[R^2] = 0^2\, P_R(0) + 3^2\, P_R(3) = 3$
- $\text{Var}[R] = E[R^2] - \mu_R^2 = 3 - (3/2)^2 = 3/4$

Theorem 2.14

$$\text{Var}\,[aX + b] = a^2\,\text{Var}\,[X]$$

Proof: Theorem 2.14

We let $Y = aX + b$ and apply Theorem 2.13. We first expand the second moment to obtain

$$E\left[Y^2\right] = E\left[a^2X^2 + 2abX + b^2\right] = a^2E\left[X^2\right] + 2ab\mu_X + b^2.$$

Expanding the right side of Theorem 2.12 yields

$$\mu_Y^2 = a^2\mu_X^2 + 2ab\mu_x + b^2.$$

Because $\text{Var}[Y] = E[Y^2] - \mu_Y^2$, Equations (2.101) and (2.102) imply that

$$\text{Var}\,[Y] = a^2E\left[X^2\right] - a^2\mu_X^2 = a^2(E\left[X^2\right] - \mu_X^2) = a^2\,\text{Var}\,[X]\,.$$

Example 2.35 Problem

A new fax machine automatically transmits an initial cover page that precedes the regular fax transmission of X information pages. Using this new machine, the number of pages in a fax is $Y = X + 1$. What are the expected value and variance of Y?

Example 2.35 Solution

The expected number of transmitted pages is $E[Y] = E[X] + 1$. The variance of the number of pages sent is $\mathrm{Var}[Y] = \mathrm{Var}[X]$.

Example 2.36 Problem

In Example 2.30, the amplitude V in volts has PMF

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

A new voltmeter records the amplitude U in millivolts. What is the variance of U?

Example 2.36 Solution

Note that $U = 1000V$. To use Theorem 2.14, we first find the variance of V. The expected value of the amplitude is

$$\mu_V = 1/7[-3 + (-2) + (-1) + 0 + 1 + 2 + 3] = 0 \text{ volts.}$$

The second moment is

$$E\left[V^2\right] = 1/7[(-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2] = 4 \text{ volts}^2$$

Therefore the variance is $\text{Var}[V] = E[V^2] - \mu_V^2 = 4$ volts^2. By Theorem 2.14,

$$\text{Var}[U] = 1000^2 \, \text{Var}[V] = 4{,}000{,}000 \text{ millivolts}^2.$$

Theorem 2.15

(a) If X is Bernoulli (p), then $\text{Var}[X] = p(1-p)$.

(b) If X is geometric (p), then $\text{Var}[X] = (1-p)/p^2$.

(c) If X is binomial (n, p), then $\text{Var}[X] = np(1-p)$.

(d) If X is Pascal (k, p), then $\text{Var}[X] = k(1-p)/p^2$.

(e) If X is Poisson (α), then $\text{Var}[X] = \alpha$.

(f) If X is discrete uniform (k, l), then $\text{Var}[X] = (l-k)(l-k+2)/12$.

Theorem 2.15

Random Variable	Variance, $\mathrm{Var}[X]$
Bernoulli (p)	$p(1-p)$
Geometric (p)	$(1-p) / p^2$
Poisson (α)	α
Binomial (n, p)	$np(1-p)$
Pascal (k, p)	$k(1-p) / p^2$
Discrete uniform (k, l)	$(l-k)(l-k+2) / 12$

Quiz 2.8

In an experiment to monitor two calls, the PMF of N the number of voice calls, is

$$P_N(n) = \begin{cases} 0.1 & n = 0, \\ 0.4 & n = 1, \\ 0.5 & n = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Find

(1) The expected value $E[N]$

(2) The second moment $E[N^2]$

(3) The variance $\mathrm{Var}[N]$

(4) The standard deviation σ_N

Quiz 2.8 Solution

The PMF $P_N(n)$ allows us to calculate each of the desired quantities.

(1) The expected value of N is

$$E[N] = \sum_{n=0}^{2} n P_N(n) = 0(0.1) + 1(0.4) + 2(0.5) = 1.4$$

(2) The second moment of N is

$$E\left[N^2\right] = \sum_{n=0}^{2} n^2 P_N(n) = 0^2(0.1) + 1^2(0.4) + 2^2(0.5) = 2.4$$

(3) The variance of N is

$$\mathrm{Var}[N] = E\left[N^2\right] - (E[N])^2 = 2.4 - (1.4)^2 = 0.44$$

(4) The standard deviation is $\sigma_N = \sqrt{\mathrm{Var}[N]} = \sqrt{0.44} = 0.663$.

2.9 Conditional Probability Mass Function

Example 2.37

Let N equal the number of bytes in a fax. A conditioning event might be the event I that the fax contains an image. A second kind of conditioning would be the event $\{N > 10{,}000\}$ which tells us that the fax required more than $10{,}000$ bytes. Both events I and $\{N > 10{,}000\}$ give us information that the fax is likely to have many bytes.

Definition 2.19 Conditional PMF

Given the event B, *with* $P[B] > 0$, *the* conditional probability mass function *of* X *is*

$$P_{X|B}(x) = P[X = x|B].$$

Theorem 2.16

A random variable X resulting from an experiment with event space $B_1, \ldots, B_m$ has PMF

$$P_X(x) = \sum_{i=1}^{m} P_{X|B_i}(x) \, P[B_i].$$

Proof: Theorem 2.16

The theorem follows directly from Theorem 1.10 with A denoting the event $\{X = x\}$.

Example 2.38 Problem

Let X denote the number of additional years that a randomly chosen 70 year old person will live. If the person has high blood pressure, denoted as event H, then X is a geometric $(p = 0.1)$ random variable. Otherwise, if the person's blood pressure is regular, event R, then X has a geometric $(p = 0.05)$ PMF with parameter. Find the conditional PMFs $P_{X|H}(x)$ and $P_{X|R}(x)$. If 40 percent of all seventy year olds have high blood pressure, what is the PMF of X?

Example 2.38 Solution

The problem statement specifies the conditional PMFs in words. Mathematically, the two conditional PMFs are

$$P_{X|H}(x) = \begin{cases} 0.1(0.9)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

$$P_{X|R}(x) = \begin{cases} 0.05(0.95)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Since H, R is an event space, we can use Theorem 2.16 to write

$$\begin{aligned} P_X(x) &= P_{X|H}(x)\, P[H] + P_{X|R}(x)\, P[R] \\ &= \begin{cases} (0.4)(0.1)(0.9)^{x-1} + (0.6)(0.05)(0.95)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Conditional PMF

- When a conditioning event $B \subset S_X$, the PMF $P_X(x)$ determines both the probability of B as well as the conditional PMF:

$$P_{X|B}(x) = \frac{P[X = x, B]}{P[B]}$$

- Now either the event $X = x$ is contained in the event B or it is not.
- If $x \in B$, then $\{X = x\} \cap B = \{X = x\}$ and $P[X = x, B] = \mathrm{P}_X(x)$.
- If $x \notin B$, then $\{X = x\} \cap B) = \phi$, and $P[X = x, B] = 0$.

Theorem 2.17

$$P_{X|B}(x) = \begin{cases} \dfrac{P_X(x)}{P[B]} & x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

- The theorem states that when we learn that an outcome $x \in B$, the probabilities of all $x \notin B$ are zero in our conditional model and the probabilities of all $x \in B$ are proportionally higher than there were before we learned $x \in B$.

Example 2.39 Problem

In the probability model of Example 2.29, the length of a fax X has PMF

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

Suppose the company has two fax machines, one for faxes shorter than five pages and the other for faxes that have five or more pages. What is the PMF of fax length in the second machine?

Example 2.39 Solution

Relative to $P_X(x)$, we seek a conditional PMF. The condition is $x \in L$ where $L = \{5, 6, 7, 8\}$. From Theorem 2.17,

$$P_{X|L}(x) = \begin{cases} \dfrac{P_X(x)}{P[L]} & x = 5, 6, 7, 8, \\ 0 & \text{otherwise.} \end{cases}$$

From the definition of L, we have

$$P[L] = \sum_{x=5}^{8} P_X(x) = 0.4.$$

With $P_X(x) = 0.1$ for $x \in L$,

$$P_{X|L}(x) = \begin{cases} 0.1/0.4 = 0.25 & x = 5, 6, 7, 8, \\ 0 & \text{otherwise.} \end{cases}$$

Thus the lengths of long faxes are equally likely. Among the long faxes, each length has probability 0.25.

Example 2.40 Problem

Suppose X, the time in integer minutes you must wait for a bus, has the uniform PMF

$$P_X(x) = \begin{cases} 1/20 & x = 1, 2, \ldots, 20, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose the bus has not arrived by the eighth minute, what is the conditional PMF of your waiting time X?

Example 2.40 Solution

Let A denote the event $X > 8$. Observing that $P[A] = 12/20$, we can write the conditional PMF of X as

$$P_{X|X>8}(x) = \begin{cases} \dfrac{1/20}{12/20} = \dfrac{1}{12} & x = 9, 10, \ldots, 20, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2.18

(a) For any $x \in B$, $P_{X|B}(x) \geq 0$.

(b) $\sum_{x \in B} P_{X|B}(x) = 1$.

(c) For any event $C \subset B$, $P[C|B]$, the conditional probability that X is in the set C, is

$$P\left[C|B\right] = \sum_{x \in C} P_{X|B}\left(x\right).$$

Definition 2.20 Conditional Expected Value

The conditional expected value *of random variable X given condition B is*

$$E[X|B] = \mu_{X|B} = \sum_{x \in B} x P_{X|B}(x).$$

Theorem 2.19

For a random variable X resulting from an experiment with event space $B_1, \ldots, B_m$,

$$E[X] = \sum_{i=1}^{m} E[X|B_i]\, P[B_i]\,.$$

Proof: Theorem 2.19

Since $E[X] = \sum_x x\, P_X(x)$, we can use Theorem 2.16 to write

$$P_X(x) = \sum_{i=1}^{m} P_{X|B_i}(x) P[B_i]$$

$$\begin{aligned} E[X] &= \sum_x x \sum_{i=1}^{m} P_{X|B_i}(x)\, P[B_i] \\ &= \sum_{i=1}^{m} P[B_i] \sum_x x\, P_{X|B_i}(x) = \sum_{i=1}^{m} P[B_i]\, E[X|B_i]\,. \end{aligned}$$

Theorem 2.20

The conditional expected value of $Y = g(X)$ given condition B is

$$E\left[Y|B\right] = E\left[g(X)|B\right] = \sum_{x \in B} g(x) P_{X|B}\left(x\right).$$

Example 2.41 Problem

Find the conditional expected value, the conditional variance, and the conditional standard deviation for the long faxes defined in Example 2.39.

Example 2.41 Solution

$$E\left[X|L\right] = \mu_{X|L} = \sum_{x=5}^{8} x\, P_{X|L}\left(x\right) = 0.25 \sum_{x=5}^{8} x = 6.5 \text{ pages}$$

$$E\left[X^2|L\right] = 0.25 \sum_{x=5}^{8} x^2 = 43.5 \text{ pages}^2$$

$$\mathrm{Var}\left[X|L\right] = E\left[X^2|L\right] - \mu_{X|L}^2 = 1.25 \text{ pages}^2$$

$$\sigma_{X|L} = \sqrt{\mathrm{Var}\left[X|L\right]} = 1.12 \text{ pages}$$

Quiz 2.9

On the Internet, data is transmitted in packets. In a simple model for World Wide Web traffic, the number of packets N needed to transmit a Web page depends on whether the page has graphic images. If the page has images (event I), then N is uniformly distributed between 1 and 50 packets. If the page is just text (event T), then N is uniform between 1 and 5 packets. Assuming a page has images with probability $1/4$, find the

(1) conditional PMF $P_{N|I}(n)$

(2) conditional PMF $P_{N|T}(n)$

(3) PMF $P_N(n)$

(4) conditional PMF $P_{N|N\leq 10}(n)$

(5) conditional expected value $E[N|N \leq 10]$

(6) conditional variance $\mathrm{Var}[N|N \leq 10]$

Quiz 2.9 Solution

(1) From the problem statement, we learn that the conditional PMF of N given the event I is

$$P_{N|I}(n) = \begin{cases} 0.02 & n = 1, 2, \ldots, 50 \\ 0 & \text{otherwise} \end{cases}$$

(2) Also from the problem statement, the conditional PMF of N given the event T is

$$P_{N|T}(n) = \begin{cases} 0.2 & n = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$$

(3) The problem statement tells us that $P[T] = 1 - P[I] = 3/4$. From Theorem 1.10 (the law of total probability), we find the PMF of N is

$$\begin{aligned} P_N(n) &= P_{N|T}(n)\, P[T] + P_{N|I}(n)\, P[I] \\ &= \begin{cases} 0.2(0.75) + 0.02(0.25) & n = 1, 2, 3, 4, 5 \\ 0(0.75) + 0.02(0.25) & n = 6, 7, \ldots, 50 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 0.155 & n = 1, 2, 3, 4, 5 \\ 0.005 & n = 6, 7, \ldots, 50 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

[Continued]

Quiz 2.9 Solution (continued)

(4) First we find

$$P[N \le 10] = \sum_{n=1}^{10} P_N(n) = (0.155)(5) + (0.005)(5) = 0.80$$

By Theorem 2.17, the conditional PMF of N given $N \le 10$ is

$$\begin{aligned} P_{N|N\le 10}(n) &= \begin{cases} \frac{P_N(n)}{P[N\le 10]} & n \le 10 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 0.155/0.8 = 0.19375 & n = 1,2,3,4,5 \\ 0.005/0.8 = 0.00625 & n = 6,7,8,9,10 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

(5) Once we have the conditional PMF, calculating conditional expectations is easy.

$$\begin{aligned} E[N|N \le 10] &= \sum_n n P_{N|N\le 10}(n) \\ &= \sum_{n=1}^{5} n(0.19375) + \sum_{n=6}^{10} n(0.00625) \\ &= 3.15625 \end{aligned}$$

[Continued]

Quiz 2.9 Solution (continued)

(6) To find the conditional variance, we first find the conditional second moment

$$
\begin{aligned}
E\left[N^2|N \le 10\right] &= \sum_n n^2 P_{N|N\le 10}(n) \\
&= \sum_{n=1}^{5} n^2(0.19375) + \sum_{n=6}^{10} n^2(0.00625) \\
&= 55(0.19375) + 330(0.00625) = 12.71875
\end{aligned}
$$

The conditional variance is

$$
\begin{aligned}
\mathrm{Var}[N|N \le 10] &= E\left[N^2|N \le 10\right] - (E\,[N|N \le 10])^2 \\
&= 12.71875 - (3.15625)^2 = 2.75684
\end{aligned}
$$

Summary of Discrete RV Families

- Theorem 2.4 ~ 2.7 and 2.15

Random Variable	Expected Value $E[X]$	Variance $\mathrm{Var}[X]$
Bernoulli (p)	p	$p(1-p)$
Geometric (p)	$1 / p$	$(1-p) / p^2$
Poisson (α)	α	α
Binomial (n, p)	np	$np(1-p)$
Pascal (k, p)	k / p	$k(1-p) / p^2$
Discrete uniform (k, l)	$(k+l) / 2$	$(l-k)(l-k+2) / 12$