Random Variables

P. Sam Johnson

Random Variables

Frequently, when an experiment is performed, we are interested mainly in some function of the outcome as opposed to the actual outcome itself. For instance, in tossing dice, we are often interested in the sum of the two dice and are not really concerned about the separate values of each die. That is, we may be interested in knowing that the sum is 7 and may not be concerned over whether the actual outcome was $(1, 6)$, $(2, 5)$, $(3, 4)$, $(4, 3)$, $(5, 2)$, or $(6, 1)$. Also, in flipping a coin, we may be interested in the total number of heads that occur and not care at all about the actual head-tail sequence that results. These quantities of interest, or, more formally, these **real-valued functions defined on the sample space**, are known as random variables.

Because the value of a random variable is determined by the outcome of the experiment, we may assign probabilities to the possible values of the random variable.

Example

Example 1.

Suppose that our experiment consists of tossing 3 *fair coins. If we let* Y *denote the number of heads that appear, then* Y *is a random variable taking on one of the values* $0, 1, 2,$ *and* 3 *with respective probabilities*

$$P\{Y = 0\} = P\{(T, T, T)\} = \frac{1}{8}$$
$$P\{Y = 1\} = P\{(T, T, H), (T, H, T), (H, T, T)\} = \frac{3}{8}$$
$$P\{Y = 2\} = P\{(T, H, H), (H, T, H), (H, H, T)\} = \frac{3}{8}$$
$$P\{Y = 3\} = P\{(H, H, H)\} = \frac{1}{8}.$$

Example

Example 2.

Three balls are to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20. If we bet that at least one of the balls that are drawn has a number as large as or larger than 17, what is the probability that we win the bet?

Solution. Let X denote the largest number selected. Then X is a random variable taking on one of the values $3, 4, \ldots, 20$. Furthermore, if we suppose that each of the $\binom{20}{3}$ possible selections are equally likely to occur, then

$$P\{X = i\} = \frac{\binom{i-1}{2}}{\binom{20}{3}} \quad i = 3, \ldots, 20. \tag{1}$$

Example (contd...)

Equation 7 follows because the number of selections that result in the event $\{X = i\}$ is just the number of selections that result in the ball numbered i and two of the balls numbered 1 through $i - 1$ being chosen. Because there are clearly $\binom{1}{1}\binom{i-1}{2}$ such selections, we obtain the probabilities expressed in Equation (7), from which we see that

$$P\{X = 20\} = \frac{\binom{19}{2}}{\binom{20}{3}} = \frac{3}{20} = .150$$

$$P\{X = 19\} = \frac{\binom{18}{2}}{\binom{20}{3}} = \frac{51}{380} \approx .134$$

Example (contd...)

$$P\{X = 18\} = \frac{\binom{17}{2}}{\binom{20}{3}} = \frac{34}{285} \approx .119$$

$$P\{X = 17\} = \frac{\binom{16}{2}}{\binom{20}{3}} = \frac{2}{19} \approx .105$$

Hence, since the event $\{X \geq 17\}$ is the union of the disjoint events $\{X = i\}, i = 17, 18, 19, 20$, it follows that the probability of our winning the bet is given by

$$P\{X \geq 17\} \approx .105 + .119 + .134 + .150 = .508.$$

Example

Example 3.

Independent trials consisting of the flipping of a coin having probability p of coming up heads are continually performed until either a head occurs or a total of n flips is made. If we let X denote the number of times the coin is flipped, then X is a random variable taking on one of the values $1, 2, 3, \ldots, n$ *with respective probabilities*

$$
\begin{aligned}
P\{X = 1\} &= P\{H\} = p \\
P\{X = 2\} &= P\{(T, H)\} = (1 - p)p \\
P\{X = 3\} &= P\{(T, T, H)\} = (1 - p)^2 p \\
&\vdots \\
P\{X = n - 1\} &= P\{(\underbrace{T, T, \ldots, T,}_{n-2} H)\} = (1 - p)^{n-2} p
\end{aligned}
$$

Example (contd...)

$$P\{X = n\} = P\{(\underbrace{T, T, \ldots, T,}_{n-1} T)(\underbrace{T, T, \ldots, T,}_{n-1} H)\} = (1-p)^{n-1}$$

As a check, note that

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n}\{X = i\}\right) &= \sum_{i=1}^{n} P\{X = i\} \\
&= \sum_{i=1}^{n-1} p(1-p)^{i-1} + (1-p)^{n-1} \\
&= p\left(\frac{1-(1-p)^{n-1}}{1-(1-p)}\right) + (1-p)^{n-1} \\
&= 1-(1-p)^{n-1} + (1-p)^{n-1} \\
&= 1.
\end{aligned}$$

Example 4.

Three balls are randomly chosen from an urn containing 3 *white,* 3 *red, and* 5 *black balls. Suppose that we win* \$1 *for each white ball selected and lose* \$1 *for each red ball selected. If we let* X denote our total winnings from the experiment, *then* X *is a random variable taking on the possible values* $0, \pm, 1, \pm 2, \pm 3$ *with respective probabilities*

$$P\{X = 0\} = \frac{\binom{5}{3} + \binom{3}{1}\binom{3}{1}\binom{5}{1}}{\binom{11}{3}} = \frac{55}{165}$$

$$P\{X = 1\} = P\{X = -1\} = \frac{\binom{3}{1}\binom{5}{2} + \binom{3}{2}\binom{3}{1}}{\binom{11}{3}} = \frac{39}{165}$$

Example (contd...)

$$P\{X = 2\} = P\{X = -2\} = \frac{\binom{3}{2}\binom{5}{1}}{\binom{11}{3}} = \frac{15}{165}$$

$$P\{X = 3\} = P\{X = -3\} = \frac{\binom{3}{3}}{\binom{11}{3}} = \frac{1}{165}.$$

These probabilities are obtained, for instance, by noting that in order for X to equal 0, either all 3 balls selected must be black or 1 ball of each color must be selected. Similarly, the event $\{X = 1\}$ occurs either if 1 white and 2 black balls are selected or if 2 white and 1 red is selected.

Example (contd...)

As a check, we note that

$$\sum_{i=0}^{3} P\{X = i\} + \sum_{i=1}^{3} P\{X = -i\} = \frac{55 + 39 + 15 + 1 + 39 + 15 + 1}{165} = 1.$$

The probability that we win money is given by

$$\sum_{i=1}^{3} P\{X = i\} = \frac{55}{165} = \frac{1}{3}.$$

Example 5.

Suppose that there are N distinct types of coupons and that each time one obtains a coupon, it is, independently of previous selections, equally likely to be any one of the N types. One random variable of interest is T, the number of coupons that needs to be collected until one obtains a complete set of at least one of each type. Rather than derive $P\{T = n\}$ directly, let us start by considering the probability that T is greater than n. To do so, fix n and define the events $A_1, A_2, \ldots, A_N$ as follows: A_j is the event that no type j coupon is contained among the first n coupons collected, $j = 1, \ldots, N$. Hence,

$$
\begin{aligned}
P\{T > n\} &= P\left(\bigcup_{j=1}^{N} A_j\right) \\
&= \sum_j P(A_j) - \sum_{j_1<j_2}\sum P(A_{j_1}A_{j_2}) + \cdots \\
&+ (-1)^{k+1} \sum_{j_1<j_2<\cdots<j_k}\sum\sum P(A_{j_1}A_{j_2}\ldots A_{j_k}) \ldots \\
&+ (-1)^{N+1} P(A_1A_2\cdots A_N).
\end{aligned}
$$

Example (contd...)

Now, A_j will occur if each of the n coupons collected is not of type j. Since each of the coupons will not be of type j with probability $(N-1)/N$, we have, by the assumed independence of the types of successive coupons,

$$P(A_j) = \left(\frac{N-1}{N}\right)^n.$$

Also, the event $A_{j_1}A_{j_2}$ will occur if none of the first n coupons collected is of either type j_1 or type j_2. Thus, again using independence, we see that

$$P(A_{j_1}A_{j_2}) = \left(\frac{N-2}{N}\right)^n.$$

Example (contd...)

The same reasoning gives

$$P(A_{j_1} A_{j_2} \cdots A_{_k}) = \left(\frac{N-k}{N} \right)^n$$

and we see that, for $n > 0$,

$$\begin{aligned} P\{T > n\} &= N \left(\frac{N-1}{N} \right)^n - \binom{N}{2} \left(\frac{N-2}{N} \right)^n + \binom{N}{3} \left(\frac{N-3}{N} \right)^n - \cdots \\ &\quad + (-1)^N \binom{N}{N-1} \left(\frac{1}{N} \right)^n \\ &= \sum_{i=1}^{N-1} \binom{N}{i} \left(\frac{N-i}{N} \right)^n (-1)^{i+1}. \end{aligned} \tag{2}$$

Example (contd...)

The probability that T equals n can now be obtained from the preceding formula by the use of

$$P\{T > n-1\} = P\{T = n\} + P\{T > n\}$$

or, equivalently,

$$P\{T = n\} = P\{T > n-1\} - P\{T > n\}.$$

Remark

Since one must collect at least N coupons to obtain a compete set, it follows that $P\{T > n\} = 1$ if $n < N$. Therefore, from Equation (2), we obtain the interesting combinatorial identity that, for integers $1 \leq n < N$,

$$\sum_{i=1}^{N-1} \binom{N}{i} \left(\frac{N-i}{N}\right)^n (-1)^{i+1} = 1$$

which can be written as

$$\sum_{i=1}^{N-1} \binom{N}{i} \left(\frac{N-i}{N}\right)^n (-1)^{i+1} = 0$$

or, upon multiplying by $(-1)^N N^n$ and letting $j = N - i$,

$$\sum_{j=1}^{N} \binom{N}{j} j^n (-1)^{j-1} = 0 \qquad 1 \leq n < N.$$

Cumulative Distribution Function

For a random variable X, the function F defined by

$$F(x) = P\{X \leq x\} \quad -\infty < x < \infty$$

is called the cumulative distribution function, or, more simply, the distribution function, of X. Thus, the distribution function specifies, for all real values x, the probability that the random variable is less than or equal to x.

$F(x)$ is a nondecreasing function of x.

Discrete Random Variables

A random variable that can take on at most a countable number of possible values is said to be discrete. For a discrete random variable X, we define the **probability mass function** p(a) of X by

$$p(a) = P\{X = a\}.$$

The probability mass function $p(a)$ is positive for at most a countable number of values of a. That is, if X must assume one of the values $x_1, x_2, \dots,$ then

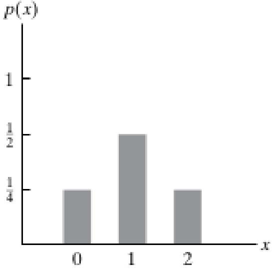
$$p(x_i) \geq 0 \quad for\, i = 1, 2, \dots$$
$$p(x) = 0 \quad \text{for all other values of } x.$$

Since X must take on one of the values x_i, we have

$$\sum_{i=1}^{\infty} p(x_i) = 1.$$

It is often instructive to present the probability mass function in a graphical format by plotting $p(x_i)$ on the y-axis against x_i on the x-axis. For instance, if the probability mass function of X is

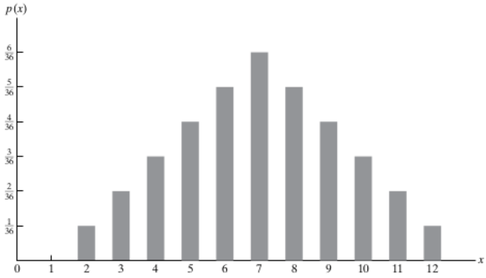
$$p(0) = \frac{1}{4}, \quad p(1) = \frac{1}{2}, \quad p(2) = \frac{1}{4}.$$



Exercise 6.

How to check whether the given function $p(x_i)$, $i = 1, 2, \ldots$ is a probability mass function?

Answer : We need to verifty $\sum_{i=1}^{\infty} p(x_i) = 1$.



We can represent this function graphically as shown in the above figure.

Example 7.

The probability mass function of a random variable X is given by $p(i) = c\lambda^i/i!, i = 0, 1, 2, \ldots,$ where λ is some positive value. Find (a) $P\{X = 0\}$ and (b) $P\{X > 2\}$.

Solution. Since $\sum_{i=0}^{\infty} p(i) = 1,$, we have $c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$ which, because $e^x = \sum_{i=0}^{\infty} x^i/i!$, implies that

$$ce^{\lambda} = 1 \quad \text{or} \quad c = e^{-\lambda}.$$

Hence,

(a) $P\{X = 0\} = e^{-\lambda}\lambda^0/0! = e^{-\lambda}$.

(b)

$$\begin{aligned} P\{X > 2\} = 1 - P\{X \leq 2\} &= 1 - P\{X = 0\} - P\{X = 1\} - P\{X = 2\} \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}. \end{aligned}$$

Cumulative Distribution Function

The cumulative distribution function F can be expressed in terms of $p(a)$ by

$$F(a) = \sum_{\text{all } x \leq a} p(x).$$

If X is a discrete random variable whose possible values are $x_1, x_2, x_3, \ldots,$ where $x_1 < x_2 < x_3 < \cdots$, then the distribution function F of X is a step function. That is, the value of F is constant in the intervals $[x_{i-1}, x_i)$ and then takes a step (or jump) of size $p(x_i)$ at x_i. For instance, let X be a probability mass function given by

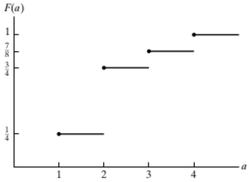
$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8}.$$

Cumulative Distribution Function

Then its cumulative distribution function is

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \geq a. \end{cases}$$

This function is depicted graphically in the following figure.



Note that the size of the step at any of the values $1, 2, 3$, and 4 is equal to the probability that X assumes that particular value.

Expected Value

If X is a discrete random variable having a probability mass function $p(x)$, then the expectation, or the expected value, of X, denoted by $E[X]$, is defined by

$$E[X] = \sum_{x:p(x)>0} xp(x).$$

In words, the expected value of X is a weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it. For instance, on the one hand, if the probability mass function of X is given by $p(0) = \frac{1}{2} = p(1)$, then

$$E[X] = 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{1}{2}$$

is just the ordinary average of the two possible values, 0 and 1, that X can assume.

Expected Value

On the other hand, if

$$p(0) = \frac{1}{3} \quad p(1) = \frac{2}{3}$$

then

$$E[x] = 0\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right) = \frac{2}{3}$$

is a weighted average of the two possible values 0 and 1, where the value 1 is given twice as much weight as the value 0, since $p(1) = 2p(0)$.

Motivation

Another motivation of the definition of expectation is provided by the frequency interpretation of probabilities. This interpretation assumes that if an infinite sequence of independent replications of an experiment is performed, then, for any event E, the proportion of time that E occurs will be $P(E)$. Now, consider a random variable X that must take on one of the values $x_1, x_2, \ldots, x_n$ with respective probabilities $p(x_1), p(x_2), \ldots, p(x_n)$, and think of X as representing our winnings in a single game of chance. That is, with probability $p(x_i)$ **we shall win** $\mathrm{x_i}$ **units** $i = 1, 2, \ldots, n$. By the frequency interpretation, if we play this game continually, then the proportion of time that we win x_i will be $p(x_i)$. Since this is true for all $i, i = 1, 2, \ldots, n$, it follows that **our average winnings per game** will be

$$\sum_{i=1}^{n} x_i p(x_i) = E[x].$$

Example 8.

Find $E[X]$, where X is the outcome when we roll a fair die.

Solution. Since $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$, we obtain

$$E[X] = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + +4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + +6\left(\frac{1}{6}\right) = \frac{7}{2}.$$

Example

Example 9.

We say that I *is an indicator variable for the event* A *if*

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs.} \end{cases}$$

Find $E[I]$.

Solution. Since $p(1) = P(A), p(0) = 1 - P(A)$, we have

$$E[I] = P(A).$$

That is, the expected value of the indicator variable for the event A is equal to the probability that A occurs.

Example 10.

A school class of 120 *students is driven in* 3 *buses to a symphonic performance. There are* 36 *students in one of the buses,* 40 *in another, and* 44 *in the third bus. When the buses arrive, one of the* 120 *students is randomly chosen. Let* X *denote the number of students on the bus of that randomly chosen student, and find* $E[X]$.

Solution. Since the randomly chosen student is equally likely to be any of the 120 students, it follows that

$$P\{X = 36\} = \frac{36}{120} \quad P\{X = 40\} = \frac{40}{120} \quad P\{X = 44\} = \frac{44}{120}.$$

Hence,

$$E[X] = 36\left(\frac{3}{10}\right) + 40\left(\frac{1}{3}\right) + 44\left(\frac{11}{30}\right) = \frac{1208}{30} = 40.2667.$$

Example (contd...)

However, the average number of students on a bus is $120/3 = 40$, showing that the expected number of students on the bus of a randomly chosen student is larger than the average number of students on a bus.

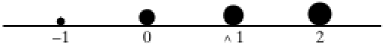
This is a general phenomenon, and it occurs because the more students there are on a bus, the more likely it is that a randomly chosen student would have been on that bus.

As a result, buses with many students are given more weight than those with fewer students.

The probability concept of expectation is analogous to the physical concept of the **center of gravity of a distribution of mass**. Consider a discrete random variable X having probability mass function $p(x_i), i \geq 1$.

Example (contd...)

If we now imagine a weightless rod in which weights with mass $p(x_i), i \geq 1$, are located at the points $x_i, i \geq 1$, then the point at which the rod would be in balance is known as the center of gravity. For those readers acquainted with elementary statics, it is now a simple matter to show that this point is at $E[X]$[1].



$$p(-1) = .10, p(0) = .25, p(1) = .30, p(2) = .35$$

$$\wedge = \text{center of gravity} = .9$$

[1] To prove this, we must show that the sum of the torques tending to turn the point around $E[X]$ is equal to 0. That is, we must show that $0 = \sum_i (x_i - E[X])p(x_i)$, which is immediate.

Expectation of a Function of a Random Variable

Suppose that we are given a discrete random variable along with its probability mass function and that we want to compute the expected value of some function of X, say, $g(X)$.

How can we accomplish this? One way is as follows: Since $g(X)$ is itself a discrete random variable, it has a probability mass function, which can be determined from the probability mass function of X.

Once we have determined the probability mass function of $g(X)$, we can compute $E[g(X)]$ by using the definition of expected value.

Example 11.

Let X denote a random variable that takes on any of the values $-1, 0,$ and 1 with respective probabilities

$$P\{X = -1\} = .2, \quad P\{X = 0\} = .5, \quad P\{X = 1\} = .3.$$

Compute $E[X^2]$.

Solution. Let $Y = X^2$. Then the probability mass function of Y is given by

$$\begin{aligned} P\{Y = 1\} &= P\{X = -1\} + P\{X = 1\} = .5 \\ P\{Y = 0\} &= P\{X = 0\} = .5. \end{aligned}$$

Hence,

$$E[X^2] = E[Y] = 1(.5) + 0(.5) = .5.$$

Example (contd...)

Note that

$$.5 = E[X^2] \neq (E[X])^2 = .01.$$

Although the preceding procedure will always enable us to compute the expected value of any function of X from a knowledge of the probability mass function of X, there is another way of thinking about $E[g(X)]$: Since $g(X)$ will equal $g(x)$ whenever X is equal to x, it seems reasonable that $E[g(X)]$ should just be a weighted average of the values $g(x)$, with $g(x)$ being weighted by the probability that X is equal to x.

Proposition

That is, the following result is quite intuitive:

Proposition 12.

If X is a discrete random variable that takes on one of the values $x_i, i \geq 1$, with respective probabilities $p(x_i)$, then, for any real-valued function g,

$$E[g(X)] = \sum_i g(x_i)p(x_i).$$

Applying it to that example yields

$$\begin{aligned} E\{X^2\} &= (-1)^2(.2) + 0^2(.5) + 1^2(.3) \\ &= 1(.2 + .3) + 0(.5). \\ &= .5. \end{aligned}$$

Proof of Proposition

The proof of Proposition (12) proceeds, as in the preceding verification, by grouping together all the terms in $\sum_i g(x_i)p(x_i)$ having the same value of $g(x_i)$. Specifically, suppose that $y_j, j \geq 1$, represent the different values of $g(x_i), i \geq 1$. Then, grouping all the $g(x_i)$ having the same value gives

$$\begin{aligned}
\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\
&= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \\
&= \sum_j y_j P\{g(X) = y_j\} \\
&= E[g(x)].
\end{aligned}$$

Corollary 13.

If a and b are constants, then

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

Proof.

$$\begin{aligned} E[aX + b] &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= a \sum_{x:p(x)>0} xp(x) + b \sum_{x:p(x)>0} p(x) \\ &= aE[X] + b. \end{aligned}$$

Mean or the first moment of X

The expected value of a random variable $X, E[X]$, is also referred to as the mean or the first moment of X.

The quantity $E[X^n], n \geq 1$, is called the nth moment of X. By Proposition (12), we note that

$$E[X^n] = \sum_{x:p(x)>0} x^n p(x).$$

Definition

Definition 14.

If X is a random variable with mean μ, then the variance of X, denoted by $Var(X)$, is defined by $Var\,(X) = E[(X-\mu)^2]$.

An alternative formula for $Var(X)$ is derived as follows:

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] = \sum_x (x-\mu)^2 p(x) = \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\ &= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\ &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu. \end{aligned}$$

That is, $Var(X) = E[X^2] - (E[X])^2$. In words, the variance of X is equal to the expected value of X^2 minus the square of its expected value. In practice, this formula frequently offers the easiest way to compute $Var(X)$.

Example

Example 15.

Calculate $Var(X)$ if X represents the outcome when a fair die is rolled.

Solution. We showed that $E[X] = \frac{7}{2}$. Also,

$$\begin{aligned} E[X^2] &= 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right) \\ &= \left(\frac{1}{6}\right)(91). \end{aligned}$$

Hence,

$$Var(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Useful Identity

A useful identity is that, for any constants a and b,

$$Var(aX + b) = a^2 Var(X).$$

Let $\mu = E[X]$. Then $E[aX + b] = a\mu + b$. Therefore,

$$\begin{aligned} Var(aX + b) &= E[(aX + b - a\mu - b)^2] \\ &= E[a^2(X - \mu)^2] = a^2 E[(X - \mu)^2] = a^2 Var(X). \end{aligned}$$

Remarks :

(a) Analogous to the means being the center of gravity of a distribution of mass, the variance represents, in the terminology of mechanics, the moment of inertia.

(b) The square root of the $Var(X)$ is called the standard deviation of X, and we denote it by $SD(X)$. That is, $SD(X) = \sqrt{Var(X)}$.

Bernoulli and Binomial Random Variables

Discrete random variables are often classified according to their probability mass functions.

Suppose that a trial, or an experiment, whose outcome can be classified as either a success or a failure is performed. If we let $X = 1$ when the outcome is a success and $X = 0$ when it is a failure, then the probability mass function of X is given by

$$\begin{aligned} p(0) &= P\{X = 0\} = 1 - p \\ p(1) &= P\{X = 1\} = p \end{aligned} \tag{3}$$

where $p, 0 \leq p \leq 1$, is the probability that the trial is a success.

A random variable X is said to be a Bernoulli random variable (after the Swiss mathematician James Bernoulli) if its probability mass function is given by Equations (3) for some $p \in (0, 1)$.

Bernoulli and Binomial Random Variables

Suppose now that n independent trials, each of which results in a success with probability p and in a failure with probability $1 - p$, are to be performed. If X represents the number of successes that occur in the n trials, then X is said to be a binomial random variable with parameters (n, p). Thus, a Bernoulli random variable is just a binomial random variable with parameters $(1, p)$.

The probability mass function of a binomial random variable having parameters (n, p) is given by

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \ldots, n. \tag{4}$$

The validity of Equation (4) may be verified by first noting that the probability of any particular sequence of n outcomes containing i successes and $n-i$ failures is, by the assumed independence of trials, $p^i(1-p)^{n-i}$. Equation (4) then follows, since there are $\binom{n}{i}$ different sequences of the n outcomes leading to i successes and $n-i$ failures. This perhaps can most easily be seen by noting that there are $\binom{n}{i}$ different choices of the i trials that result in successes.

Note that, by the binomial theorem, the probabilities sum to 1; that is,

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{n} \binom{n}{i} p^i(1-p)^{n-i} = [p+(1-p)]^n = 1.$$

Example 16.

Five fair coins are flipped. If the outcomes are assumed independent, find the probability mass function of the number of heads obtained.

Solution. If we let X equal the number of heads (successes) that appear, then X is a binomial random variable with parameters $(n = 5, p = \frac{1}{2})$. Hence, by Equation (4), we have

$$P\{X = 0\} = \binom{5}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

$$P\{X = 1\} = \binom{5}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 = \frac{5}{32}$$

$$P\{X = 2\} = \binom{5}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 = \frac{10}{32}$$

$$P\{X = 3\} = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{10}{32}$$

$$P\{X = 4\} = \binom{5}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1 = \frac{5}{32}$$

$$P\{X = 5\} = \binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 = \frac{1}{32}.$$

Few Questions

Consider the same problem with $p = \frac{1}{2}$.

1. We found that $P\{X = k\}$ reached its largest value when $k = 2$ and 3. If $p = \frac{1}{3}$, where does its largest value attain?
2. If $P\{X = k\}$ is known, can be find $P(X = k + 1)$?

Example 17.

It is known that screws produced by a certain company will be defective with probability .01, independently of each other. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

Solution. If X is the number of defective screws in a package, then X is a binomial random variable with parameters (10, .01). Hence, the probability that a package will have to be replaced is

$$1 - P\{X = 0\} - P\{X = 1\} - \binom{10}{0}(.01)^0(.99)^{10} - \binom{10}{1}(.01)^1(.99)^9.$$

Thus, only .4 percent of the packages will have to be replaced.

Example

Example 18.

The following gambling game, known as the wheel of fortune (or huck-a-luck), is quite popular at many carnivals and gambling casinos: A player bets on one of the numbers 1 through 6*. Three dice are then rolled, and if the number bet by the player appears* i *times,* $i = 1, 2, 3$*, then the player wins* i *units; if the number bet by the player does not appear on any of the dice, then the player loses* 1 *unit. Is this game fair to the player? (Actually, the game is played by spinning a wheel that comes to rest on a slot labeled by three of the numbers* 1 *through* 6*, but this variant is mathematically equivalent to the dice version.)*

Solution

If we assume that the dice are fair and act independently of each other, then the number of times that the number bet appears is a binomial random variable with parameters $\left(3, \frac{1}{6}\right)$. Hence, letting X denote the player's winnings in the game, we have

$$
\begin{aligned}
P\{X=-1\} &= \binom{3}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3 = \frac{125}{216}\\
P\{X=1\} &= \binom{3}{1}\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^1 = \frac{75}{216}\\
P\{X=2\} &= \binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1 = \frac{15}{216}\\
P\{X=3\} &= \binom{3}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0 = \frac{1}{216}.
\end{aligned}
$$

In order to determine whether or not this is a fair game for the player, let us calculate $E[X]$. From the preceding probabilities, we obtain

$$E[X] = \frac{-125+75+30+3}{216} = \frac{-17}{216}.$$

Hence, in the long run, the player will lose 17 units per every 216 games he plays.

Example

In the next example, we consider the simplest form of the theory of inheritance as developed by Gregor Mendel (1822-1884).

Example 19.

Suppose that a particular trait (such as eye color or left-handedness) of a person is classified on the basis of one pair of genes, and suppose also that d represents a dominant gene and r a recessive gene. Thus, a person with dd genes is purely dominant, one with rr is purely recessive, and one with rd is hybrid. The purely dominant and the hybrid individuals are alike in appearance. Children receive 1 *gene from each parent. If, with respect to a particular trait,* 2 *hybrid parents have a total of* 4 *children, what is the probability that* 3 *of the* 4 *children have the outward appearance of the dominant gene?*

Solution

Solution. If we assume that each child is equally likely to inherit either of 2 genes from each parent, the probabilities that the child of 2 hybrid parents will have *dd*, *rr*, and *rd* pairs of genes are, respectively, $\frac{1}{4}\frac{1}{4}$, and $\frac{1}{2}$.

Hence, since an offspring will have the outward appearance of the dominant gene if its gene pair is either *dd* or *rd*, it follows that the number of such children is binomially distributed with parameters $(4, \frac{3}{4})$.

Thus, the desired probability is

$$\binom{4}{3}\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^1 = \frac{27}{64}.$$

Properties of Binomial Random Variables

We will now examine the properties of a binomial random variable with parameters n and p. To begin, let us compute its expected value and variance. Now,

$$E[X^k] = \sum_{i=0}^{n} i^k \binom{n}{i} p^i (1-p)^{n-1} = \sum_{i=1}^{n} i^k \binom{n}{i} p^i (1-p)^{n-1}.$$

Using the identity $i\binom{n}{i} = n\binom{n-1}{i-1}$ gives

$$\begin{aligned} E[X^k] &= np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} \\ &= np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \text{ by letting } j = i-1 \\ &= npE[(Y+1)^{k-1}] \end{aligned}$$

where Y is a binomial random variable with parameters $n-1, p$. Setting $k = 1$ in the preceding equation yields

$$E[X] = np.$$

That is, the expected number of successes that occur in n independent trials when each is a success with probability p is equal to np.

Properties of Binomial Random Variables

Setting $k = 2$ in the preceding equation, and using the preceding formula for the expected value of a binomial random variable yields

$$\begin{aligned} E[X^2] &= npE[Y + 1] \\ &= np[(n-1)p + 1]. \end{aligned}$$

Since $E[X] = np$, we obtain

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \\ &= np[(n-1)p + 1] - (np)^2 \\ &= np(1-p). \end{aligned}$$

Summing up, we have shown the following:

If X is a binomial random variable with parameters n and p, then

$$E[X] = np$$
$$Var(X) = np(1-p).$$

Proposition

The following proposition details how the binomial probability mass function first increases and then decreases.

Proposition 20.

If X is a binomial random variable with parameters (n, p), where $0 < p < 1$, then as k goes from 0 to $n, P\{X = k\}$ first increases monotonically and then decreases monotonically, reaching its largest value when k is the largest integer less than or equal to $(n+1)p$.

Proof. *We prove the proposition by considering $P\{X = k\}/P\{X = k-1\}$ and determining for what values of k it is greater or less than 1. Now,*

$$\frac{P\{X=K\}}{P\{X=k-1\}} = \frac{\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}}{\frac{n!}{(n-k+1)!(k-1)!}p^{k-1}(1-p)^{n-k+1}} = \frac{(n-k+1)p}{k(1-p)}.$$

Proposition (contd...)

Hence, $P\{X = k\} \geq P\{X = k - 1\}$ if and only if

$$(n - k + 1)p \geq k(1 - p)$$

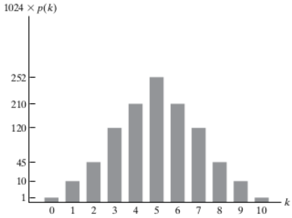
or, equivalently, if and only if

$$k \leq (n + 1)p$$

and the proposition is proved.

Graph of the probability mass function

The graph of the probability mass function of a binomial random variable with parameters $(10, \frac{1}{2})$ is shown in the following figure.



Graph of $p(k) = \binom{10}{k} \left(\frac{1}{2}\right)^{10}$

Exercise 21.

Draw the graph of the probability mass function of a binomial random variable with parameters $(10, \frac{1}{3})$.

Computing the Binomial Distibution Function

Suppose that X is binomial with parameters (n, p). The key to computing its distribution function

$$P\{X \geq i\} = \sum_{k=0}^{i} \binom{n}{k} p^k (1-p)^{n-k} \quad i = 0, 1, \ldots, n$$

is to utilize the following relationship between $P\{X = k+1\}$ and $P\{X = k\}$, which was established in the proof of Proposition (20):

$$P\{X = k+1\} = \frac{p}{1-p} \frac{n-k}{k+1} P\{X = k\}. \tag{5}$$

A computer program that utilizes the recursion (5) to compute the binomial distribution function is easily written. To compute $P\{X \leq i\}$, the program should first compute $P\{X = i\}$ and then use the recursion to successively compute $P\{X = i-1\}, P\{X = i-2\}$, and so on.

Example 22.

Let X be a binomial random variable with parameters $n = 6, p = .4$. Then, starting with $P\{X = 0\} = (.6)6$ and recursively employing Equation 5, we obtain

$$\begin{aligned}
P\{X = 0\} &= (.6)^6 \approx .0467 \\
P\{X = 1\} &= \frac{4}{6}\frac{6}{1} P\{X = 0\} \approx .1866 \\
P\{X = 2\} &= \frac{4}{6}\frac{5}{2} P\{X = 1\} \approx .3110 \\
P\{X = 3\} &= \frac{4}{6}\frac{4}{3} P\{X = 2\} \approx .2765 \\
P\{X = 4\} &= \frac{4}{6}\frac{3}{4} P\{X = 3\} \approx .1382 \\
P\{X = 5\} &= \frac{4}{6}\frac{2}{5} P\{X = 4\} \approx .0369 \\
P\{X = 6\} &= \frac{4}{6}\frac{1}{6} P\{X = 5\} \approx .0041.
\end{aligned}$$

Historical Note

Independent trials having a common probability of success p were first studied by the Swiss mathematician Jacques Bernoulli (1654-1705). In his book Ars Conjectandi (The Art of Conjecturing), published by his nephew Nicholas eight years after his death in 1713, Bernoulli showed that if the number of such trials were large, then the proportion of them that were successes would be close to p with a probability near 1.

Jacques Bernoulli was from the first generation of the most famous mathematical family of all time. Altogether, there were between 8 and 12 Bernoullis, spread over three generations, who made fundamental contributions to probability, statistics, and mathematics. One difficulty in knowing their exact number is the fact that several had the same name.

Historical Note

For example, two of the sons of Jacques's brother Jean were named Jacques and Jean. Another difficulty is that several of the Bernoullis were known by different names in different places. Our Jacques (sometimes written Jaques) was, for instance, also known as Jakob (sometimes written Jacob) and as James Bernoulli.

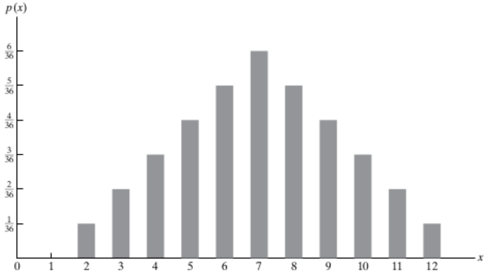
But whatever their number, their influence and output were prodigious. Like the Bachs of music, the Bernoullis of mathematics were a family for the ages!

Example 23.

If X is a binomial random variable with parameters $n = 100$ and $p = .75$, find $P\{X = 70\}$ and $P\{X \leq 70\}$.

Solution.

The answer is shown here in the following figure.



The Poisson Random Variable

A random variable X that takes on one of the values $0, 1, 2, \ldots$ is said to be a Poisson random variable with parameter λ if, for some $\lambda > 0$,

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!} \quad i = 0, 1, 2, \ldots. \tag{6}$$

Equation 6 defines a probability mass function, since

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1.$$

The Poisson random variable has a tremendous range of applications in diverse areas because it may be used as **an approximation for a binomial random variable with parameters** (n, p) **when** n **is large and** p **is small enough so that** np **is of moderate size.**

The Poisson Random Variable

To see this, suppose that X is a binomial random variable with parameters (n, p), and let $\lambda = np$. Then

$$
\begin{aligned}
P\{X = i\} &= \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \\
&= \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\
&= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{1-\lambda/n)^n}{(1-\lambda/n)^i}.
\end{aligned}
$$

Now, for n large and λ moderate,

$$
\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}; \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1; \left(1 - \frac{\lambda}{n}\right)^i \approx 1.
$$

The Poisson Random Variable

Hence, for n large and λ moderate,

$$P\{X = i\} \approx e^{-\lambda}\frac{\lambda^i}{i!}.$$

In other words, if n independent trials, each of which results in a success with probability p, are performed, then, when n is large and p is small enough to make np moderate, the **number of successes occurring is approximately a Poisson random variable with parameter** $\lambda = \text{np}$. This value λ (which will later be shown to equal the expected number of successes) will usually be determined empirically ()by means of observation or experience rather than theory of pure logic).

The Poisson Random Variable

Some examples of random variables that generally obey the Poisson probability law [that is, they obey Equation (6) are as follows:

1. The number of misprints on a page (or a group of pages) of a book .
2. The number of people in a community who survive to age 100.
3. The number of wrong telephone numbers that are dialed in a day.
4. The number of packages of dog biscuits sold in a particular store each day.
5. The number of customers entering a post office on a given day.
6. The number of vacancies occurring during a year in the federal judicial system.
7. The number of a-particles discharged in a fixed period of time from some radioactive material.

The Posson Random Variable

Each of the preceding, and numerous other random variables, are **approximately Poisson for the same reason** - namely, because of the Poisson approximation to the binomial.

For instance, we can suppose that **there is a small probability** p **that each letter typed on a page will be misprinted**. Hence, the number of misprints on a page will be approximately Poisson with $\lambda = np$, where n is the number of letters on a page.

Similarly, we can suppose that **each person in a community has some small probability of reaching age** 100.

Also, each person entering a store may be thought of as **having some small probability of buying a package of dog biscuits, and so on**.

Example

Example 24.

Suppose that the number of typographical errors on a single page of this book has a Poisson distribution with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on this page.

Solution.

Letting X denote the number of errors on this page, we have

$$P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - e^{-1/2} \approx .393.$$

Example 25.

Suppose that the probability that an item produced by a certain machine will be defective is .1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Solution. The desired probability is

$$\binom{10}{0}(.1)^0(.9)^{10} + \binom{10}{1}(.1)^1(.9)^9 = .7361,$$

whereas the Poisson approximation yields the value

$$P\{X=0\} + P\{X=1\} = e^{-\lambda}\frac{\lambda^0}{0!} + e^{-\lambda}\frac{\lambda^1}{1!} = e^{-1} + e^{-1} \approx .7358.$$

Expected value and variance of the Poisson random variable

Before computing the expected value and variance of the Poisson random variable with parameter λ, recall that this random variable approximates a binomial random variable with parameters n and p when n is large, p is small, and $\lambda = np$. Since such a binomial random variable has expected value $np = \lambda$ and variance $np(1-p) = \lambda(1-p) \approx \lambda$ (since p is small), it would seem that both the expected value and the variance of a Poisson random variable would equal its parameter λ. We now verify this result:

$$
\begin{aligned}
E[X] &= \sum_{i=0}^{\infty} \frac{ie^{-\lambda}\lambda^i}{i!} = \lambda \sum_{i=0}^{\infty} \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} \\
&= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \text{ by letting } j = i-1 \\
&= \lambda, \quad \text{since } \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^{\lambda}.
\end{aligned}
$$

Example

Thus, the expected value of a Poisson random variable X is indeed equal to its parameter λ. To determine its variance, we first compute $E[X^2]$:

$$E[X^2] = \sum_{i=0}^{\infty} \frac{i^2 e^{-\lambda} \lambda^i}{i!} = \lambda \sum_{i=1}^{\infty} \frac{i e^{-\lambda} \lambda^{i-1}}{(i-1)!}$$

$$= \lambda \sum_{j=0}^{\infty} \frac{(j+1) e^{-\lambda} \lambda^j}{j!} \text{ by letting } j = i - 1$$

$$= \lambda \left[\sum_{j=0}^{\infty} \frac{j e^{-\lambda} \lambda^j}{j!} + \sum_{j=0}^{\infty} \frac{e^{-\lambda} \lambda^j}{j!} \right] = \lambda(\lambda + 1)$$

where the final equality follows because the first sum is the expected value of a Poisson random variable with parameter λ and the second is the sum of the probabilities of this random variable.

Example

Therefore, since we have shown that $E[X] = \lambda$, we obtain

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

Hence, the expected value and variance of a Poisson random variable are both equal to its parameter λ.

Example

We have shown that the Poisson distribution with parameter np is a very good approximation to the distribution of the number of successes in n independent trials when each trial has probability p of being a success, provided that n is large and p small. In fact, it remains a good approximation even when the trials are not independent, provided that their dependence is weak.

For instance, recall the matching problem in which n men randomly select hats from a set consisting of one hat from each person. From the point of view of the number of men who select their own hat, we may regard the random selection as the result of n trials where we say that trial i is a success if person i selects his own hat, $i = 1, \ldots, n$.

Example

Defining the events $E_i, i = 1, \dots, n$, by

$$E_i = \{\text{trial i is a success}\}$$

it is easy to see that

$$P\{E_i\} = \frac{1}{n} \quad \text{and} \quad P\{E_i|E_j\} = \frac{1}{n-1}, \quad j \neq i.$$

Thus, we see that, although the events $E_i, i = 1, \dots, n$ are not independent, their dependence, for large n, appears to be weak. Because of this it seems reasonable to expect that the number of successes will approximately have a Poisson distribution with parameter $n \times 1/n = 1$.

Example

For a second illustration of the strength of the Poisson approximation when the trials are weakly dependent, let us consider again the birthday problem.

In this example, we suppose that each of n people is equally likely to have any of the 365 days of the year as his or her birthday, and the problem is to determine the probability that a set of n independent people all have different birthdays.

A combinatorial argument was used to determine this probability, which was shown to be less than $\frac{1}{2}$ when $n = 23$.

Poisson Random Variable

We can approximate the preceding probability by using the Poisson approximation as follows: Imagine that we have a trial for each of the $\binom{n}{2}$ pairs of individuals i and j, $i \neq j$, and say that trial i, j is a success if persons i and j have the same birthday.

If we let E_{ij} denote the event that trial i, j is a success, then, whereas the $\binom{n}{2}$ events E_{ij}, $1 \leq i < j \leq n$, are not independent, their dependence appears to be rather weak. Indeed, these events are even *pairwise independent*, in that any 2 of the events E_{ij} and E_{kl} are independent-again. Since $P(E_{ij}) = 1/365$, it is reasonable to suppose that the number of successes should approximately have a Poisson distribution with mean $\binom{n}{2}/365 = n(n-1)/730$.

Poisson Random Variable

Therefore,

$$P\{\text{no 2 people have the same birthday}\} = P\{0 \text{ successes}\} \approx \exp\left\{\frac{-n(n-1)}{730}\right\}.$$

To determine the smallest integer n for which this probability is less than $\frac{1}{2}$, note that

$$\exp\left\{\frac{-n(n-1)}{730}\right\} \leq \frac{1}{2}$$

is equivalent to

$$\exp\left\{\frac{n(n-1)}{730}\right\} \geq 2.$$

Taking logarithms of both sides, we obtain

$$n(n-1) \geq 730 \log 2 \approx 505.997$$

which yields the solution $n = 23$.

The Poisson Random Variable

Suppose now that we wanted the probability that, among the n people, no 3 of them have their birthday on the same day. **Whereas this now becomes a difficult combinatorial problem, it is a simple matter to obtain a good approximation.**

To begin, imagine that we have a trial for each of the $\binom{n}{3}$ triplets i, j, k, where $1 \leq i < j < k \leq n$, and call the i, j, k trial a success if persons i, j, and k all have their birthday on the same day. As before, we can then conclude that the number of successes is approximately a Poisson random variable with parameter

$$\begin{aligned}\binom{n}{3} P\{i, j, k \text{ have the same birthday}\} &= \binom{n}{3}\left(\frac{1}{365}\right)^2 \\ &= \frac{n(n-1)(n-2)}{6 \times (365)^2}.\end{aligned}$$

The Poisson Random Variable

Hence,

$$P\{no \text{ 3 have the same birthday}\} \approx \exp\left\{\frac{-n(n-1)(n-2)}{799350}\right\}.$$

This probability will be less than $\frac{1}{2}$ when n is such that

$$n(n-1)(n-2) \geq 799350 \log 2 \approx 554067.1$$

which is equivalent to $n \geq 84$. Thus, the approximate probability that at least 3 people in a group of size 84 or larger will have the same birthday exceeds $\frac{1}{2}$.

Poisson Paradigm

For the number of events to occur to approximately have a Poisson distribution, it is not essential that all the events have the same probability of occurrence, but only that all of these probabilities be small. The following is referred to as the *Poisson paradigm*.

Consider n events, with p_i equal to the probability that event i occurs, $i = 1, \ldots, n$. If all the p_i are "small" and the trials are either independent or at most "weakly dependent," then the number of these events that occur approximately has a Poisson distribution with mean $\sum_{i=1}^{n} p_i$.

Our next example not only makes use of the Poisson paradigm, but also illustrates a variety of the techniques we have studied so far.

Other Discrete Probability Distributions : Geometric Random Variable

Suppose that independent trials, each having a probability p, $0 < p < 1$, of being a success, are performed until a success occurs. If we let X equal the number of trials required, then

$$P\{X = n\} = (1-p)^{n-1}p \quad n = 1, 2, \ldots. \tag{7}$$

Equation (7) follows because, in order for X to equal n, it is necessary and sufficient that the first $n-1$ trials are failures and the nth trial is a success. Equation (7) then follows, since the outcomes of the successive trials are assumed to be independent. Since

$$\sum_{n=1}^{\infty} P\{X = n\} = p\sum_{n=1}^{\infty}(1-p)^{n-1} = \frac{p}{1-(1-p)} = 1$$

it follows that, with probability 1, a success will eventually occur. Any random variable X whose probability mass function is given by Equation (7) is said to be a *geometric* random variable with parameter p.

Example 26.

An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each ball selected is replaced before the next one is drawn, what is the probability that

(a) *exactly n draws are needed?*

(b) *at least k draws are needed?*

Solution: If we let X denote the number of draws needed to select a black ball, then X satisfies Equation (7) with $p = M/(M+N)$. Hence,

(a)

$$P\{X = n\} = \left(\frac{N}{M+N}\right)^{n-1} \frac{M}{M+N} = \frac{MN^{n-1}}{(M+N)^n}.$$

(b)

$$P\{X \geq k\} = \frac{M}{M+N} \sum_{n=k}^{\infty} \left(\frac{N}{M+N} \right)^{n-1}$$
$$= \left(\frac{M}{M+N} \right) \left(\frac{N}{M+N} \right)^{k-1} \Big/ \left[1 - \frac{N}{M+N} \right]$$
$$= \left(\frac{N}{M+N} \right)^{k-1}.$$

Of course, part (b) could have been obtained directly, since the probability that at least k trials are necessary to obtain a success is equal to the probability that the first $k-1$ trials are all failures. That is, for a geometric random variable,

$$P\{X \geq k\} = (1-p)^{k-1}.$$

Example 27.

Find the expected value of a geometric random variable.

Solution: With $q = 1 - p$, we have

$$E[X] = \sum_{i=1}^{\infty} iq^{i-1}p = \sum_{i=1}^{\infty}(i-1+1)q^{i-1}p$$
$$= \sum_{i=1}^{\infty}(i-1)q^{i-1}p + \sum_{i=1}^{\infty} q^{i-1}p$$
$$= \sum_{j=0}^{\infty} jq^j p + 1 = q\sum_{j=1}^{\infty} jq^{j-1}p + 1 = qE[X] + 1.$$

Solution (Contd...)

Hence,

$$pE[X] = 1$$

yielding the result

$$E[X] = \frac{1}{p}$$

In other words, if independent trials having a common probability p of being successful are performed until the first success occurs, then the expected number of required trials equals $1/p$. For instance, the expected number of rolls of a fair die that it takes to obtain the value 1 is 6.

Example 28.

Find the variance of a geometric random variable.

Solution: To determine $Var(X)$, let us first compute $E[X^2]$. With $q = 1 - p$, we have

$$
\begin{aligned}
E[X^2] &= \sum_{i=1}^{\infty} i^2 q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i - 1 + 1)^2 q^{i-1} p \\
&= \sum_{i=1}^{\infty} (i-1)^2 q^{i-1} p + \sum_{i=1}^{\infty} 2(i-1) q^{i-1} p + \sum_{i=1}^{\infty} q^{i-1} p \\
&= \sum_{j=0}^{\infty} j^2 q^j p + 2 \sum_{i=1}^{\infty} j q^j p + 1
\end{aligned}
$$

Solution (Contd...)

Using $E[X] = 1/p$, the equation for $E[X^2]$ yields

$$pE[X^2] = \frac{2q}{p} + 1.$$

Hence,

$$E[X^2] = \frac{2q+p}{p^2} = \frac{q+1}{p^2}$$

giving the result

$$Var(X) = \frac{q+1}{p^2} - \frac{1}{p^2} = \frac{q}{p^2} = \frac{1-p}{p^2}.$$

Exercise 29.

Two balls are chosen randomly from an urn containing 8 white, 4 black, and 2 orange balls. Suppose that we win \$2 for each black ball selected and we lose \$1 for each white ball selected. Let X denote our winnings. What are the possible values of X, and what are the probabilities associated with each value?

Solution:

$$P\{X=4\} = \frac{\binom{4}{2}}{\binom{14}{2}} = \frac{6}{91} \qquad P\{X=0\} = \frac{\binom{2}{2}}{\binom{14}{2}} = \frac{1}{91}$$

$$P\{X=2\} = \frac{\binom{4}{2}\binom{2}{1}}{\binom{14}{2}} = \frac{8}{91} \qquad P\{X=-1\} = \frac{\binom{8}{1}\binom{2}{1}}{\binom{14}{2}} = \frac{16}{91}$$

$$P\{X=1\} = \frac{\binom{4}{1}\binom{8}{1}}{\binom{14}{2}} = \frac{32}{91} \qquad P\{X=-2\} = \frac{\binom{8}{2}}{\binom{14}{2}} = \frac{28}{91}.$$

Exercise 30.

Five men and 5 women are ranked according to their scores on an examination. Assume that no two scores are alike and all 10! possible rankings are equally likely. Let X denote the highest ranking achieved by a woman. (For instance, $X = 1$ if the top-ranked person is female.) Find $P\{X = i\}$, $i \in range(X)$.

Solution:

$P\{X = 1\} = 1/2, \quad P\{X = 2\} = \frac{5}{10}\frac{5}{9} = \frac{5}{18},$

$P\{X = 3\} = \frac{5}{10}\frac{4}{9}\frac{5}{8} = \frac{5}{36}, \quad P\{X = 4\} = \frac{5}{10}\frac{4}{9}\frac{3}{8}\frac{5}{7} = \frac{10}{168},$

$P\{X = 5\} = \frac{5}{10}\frac{4}{9}\frac{3}{8}\frac{2}{7}\frac{5}{6} = \frac{5}{252}, \quad P\{X = 6\} = \frac{5}{10}\frac{4}{9}\frac{3}{8}\frac{2}{7}\frac{1}{6} = \frac{1}{252}.$

Exercise 31.

Let X represent the difference between the number of heads and the number of tails obtained when a coin is tossed n times. What are the possible values of X?

Solution:

$n - 2i,\ i = 0, 1, \ldots, n.$

Exercise 32.

Suppose that a die is rolled twice. What are the possible values that the following random variables can take on:

(a) *the maximum value to appear in the two rolls;*

(b) *the minimum value to appear in the two rolls;*

(c) *the sum of the two rolls;*

(d) *the value of the first roll minus the value of the second roll?*

Exercise 33.

In the game of Two-Finger Morra, 2 players show 1 or 2 fingers and simultaneously guess the number of fingers their opponent will show. If only one of the players guesses correctly, he wins an amount (in dollars) equal to the sum of the fingers shown by him and his opponent. If both players guess correctly or if neither guesses correctly, then no money is exchanged. Consider a specified player, and denote by X the amount of money he wins in a single game of Two-Finger Morra.

(a) *If each player acts independently of the other, and if each player makes his choice of the number of fingers he will hold up and the number he will guess that his opponent will hold up in such a way that each of the 4 possibilities is equally likely, what are the possible values of X and what are their associated probabilities?*

(b) *Suppose that each player acts independently of the other. If each player decides to hold up the same number of fingers that he guesses his opponent will hold up, and if each player is equally likely to hold up 1 or 2 fingers, what are the possible values of X and their associated probabilities?*

Exercise 34.

Five distinct numbers are randomly distributed to players numbered 1 through 5. Whenever two players compare their numbers, the one with the higher one is declared the winner. Initially, players 1 and 2 compare their numbers; the winner then compares her number with that of player 3, and so on. Let X denote the number of times player 1 is a winner. Find $P\{X = i\}$, $i = 0, 1, 2, 3, 4$.

Solution:

$$
\begin{aligned}
P\{X = 0\} &= P\{1 \text{ loses to } 2\} = 1/2 \\
P\{X = 1\} &= P\{\text{of } 1, 2, 3 : 3 \text{ has largest, then } 1, \text{ then } 2\} \\
&= (1/3)(1/2) = 1/6 \\
P\{X = 2\} &= P\{\text{of } 1, 2, 3, 4 : 4 \text{ has largest and } 1 \text{ has next largest}\} \\
&= (1/4)(1/3) = 1/12 \\
P\{X = 3\} &= P\{\text{of } 1, 2, 3, 4, 5 : 5 \text{ has largest then } 1\} \\
&= (1/5)(1/4) = 1/20 \\
P\{X = 4\} &= P\{1 \text{ has largest}\} = 1/5.
\end{aligned}
$$

Exercise 35.

A gambling book recommends the following "winning strategy" for the game of roulette: Bet \$1 on red. If red appears (which has probability $\frac{18}{38}$), then take the \$1 profit and quit. If red does not appear and you lose this bet (which has probability $\frac{20}{38}$ of occurring), make additional \$1 bets on red on each of the next two spins of the roulette wheel and then quit. Let X denote your winnings when you quit.

(a) *Find $P\{X > 0\}$.*

(b) *Are you convinced that the strategy is indeed a "winning" strategy? Explain your answer!*

(c) *Find $E[X]$.*

Solution

(a)

$$P\{x > 0\} = P\{\text{win first bet}\} + P\{\text{lose,win,win}\}$$
$$= 18/38 + (20/38)(18/38)^2 \approx .5918$$

(b) No, because if the gambler wins then he or she wins \$1. However, a loss would either be \$1 or \$3.

(c) $E[X] = 1[18/38 + (20/38)(18/38)^2] - [(20/38)2(20/38)(18/38)] - 3(20/38)^3 \approx -.108.$

Exercise 36.

Four buses carrying 148 students from the same school arrive at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying the randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on her bus.

(a) *Which of $E[X]$ or $E[Y]$ do you think is larger? Why?*

(b) *Compute $E[X]$ and $E[Y]$.*

Solution:

(a) $E[X]$ since whereas the bus driver selected is equally likely to be from any of the 4 buses, the student selected is more likely to have come from a bus carrying a large number of students.

(b) $P\{X = i\} = i/148, \ i = 40, 33, 25, 50$

$$E[X] = [(40)^2 + (33)^2 + (25)^2 + (50)^2]/148 \approx 39.28$$
$$E[Y] = (40 + 33 + 25 + 50)/4 = 37.$$

Exercise 37.

Two coins are to be flipped. The first coin will land on heads with probability .6, the second with probability .7. Assume that the results of the flips are independent, and let X equal the total number of heads that result.

(a) *Find $P\{X = 1\}$.*

(b) *Determine $E[X]$.*

Solution:

(a) $\frac{1}{10}(1 + 2 + \ldots + 10) = \frac{11}{2}$

(b) after 2 questions, there are 3 remaining possibilities with probability 3/5 and 2 with probability 2/5. Hence.

$$E[\text{Number}] = \frac{2}{5}(3) + \frac{3}{5}\left[2 + \frac{1}{3} + 2\frac{2}{3}\right] = \frac{17}{5}.$$

The above assumes that when 3 remain, you choose 1 of the 3 and ask if that is the one.

Exercise 38.

One of the numbers 1 *through* 10 *is randomly chosen. You are to try to guess the number chosen by asking questions with "yes-no" answers. Compute the expected number of questions you will need to ask in each of the following two cases:*

(a) *Your* i*th question is to be "Is it* i*?"* $i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$.

(b) *With each question you try to eliminate one-half of the remaining numbers, as nearly as possible.*

Exercise 39.

An insurance company writes a policy to the effect that an amount of money A must be paid if some event E occurs within a year. If the company estimates that E will occur within a year with probability p, what should it charge the customer in order that its expected profit will be 10 percent of A?

Solution:

$$C - Ap = \frac{A}{10} \implies C = A\left(p + \frac{1}{10}\right).$$

Exercise 40.

A person tosses a fair coin until a tail appears for the first time. If the tail appears on the nth flip, the person wins 2^n dollars. Let X denote the player's winnings. Show that $E[X] = +\infty$. This problem is known as the St. Petersburg paradox.

(a) *Would you be willing to pay \$1 million to play this game once?*

(b) *Would you be willing to pay \$1 million for each game if you could play for as long as you liked and only had to settle up when you stopped playing?*

Solution:

$$E[X] = \sum_{n=1}^{\infty} 2^n (1/2)^n = \infty$$

(a) probably not

(b) yes, if you could play an arbitrarily large number of games.

Exercise 41.

Each night different meteorologists give us the probability that it will rain the next day. To judge how well these people predict, we will score each of them as follows: If a meteorologist says that it will rain with probability p, then he or she will receive a score of

$$1-(1-p)^2 \qquad \textit{if it does rain}$$

$$1-p^2 \qquad \textit{if it does not rain.}$$

We will then keep track of scores over a certain time span and conclude that the meteorologist with the highest average score is the best predictor of weather. Suppose now that a given meteorologist is aware of our scoring mechanism and wants to maximize his or her expected score. If this person truly believes that it will rain tomorrow with probability p^*, what value of p should he or she assert so as to maximize the expected score?

Solution: $E[score] = p^*[1-(1-P)^2 + (1-p^*)(1-p^2)]$

$$\frac{d}{dp} = 2(1-p)p^* - 2p(1-p^*) = 0$$

$$\implies p = p^*.$$

Exercise

Exercise 42.

A newsboy purchases papers at 10 cents and sells them at 15 cents. However, he is not allowed to return unsold papers. If his daily demand is a binomial random variable with $n = 10$, $p = \frac{1}{3}$, approximately how many papers should he purchase so as to maximize his expected profit?

Exercise 43.

If $E[X] = 1$ *and* $Var(X) = 5$, *find*

(a) $E[(2 + X)^2]$;

(b) $Var(4 + 3X)$.

Solution:

(a) $E[(2 + X)^2] = Var(2 + X) + (E[2 + X])^2 = Var(X) + 9 = 14$

(b) $Var(4 + 3X) = 9Var(X) = 45$.

Exercise 44.

A communications channel transmits the digits 0 *and* 1. *However, due to static, the digit transmitted is incorrectly received with probability .2. Suppose that we want to transmit an important message consisting of one binary digit. To reduce the chance of error, we transmit* 00000 *instead of* 0 *and* 11111 *instead of* 1. *If the receiver of the message uses "majority" decoding, what is the probability that the message will be wrong when decoded? What independence assumptions are you making?*

Solution:

$$\binom{5}{3}(.2)^3(.8)^2 + \binom{5}{4}(.2)^4(.8) + (.2)^5.$$

Exercise 45.

A satellite system consists of n components and functions on any given day if at least k of the n components function on that day. On a rainy day each of the components independently functions with probability p_1, whereas on a dry day they each independently function with probability p_2. If the probability of rain tomorrow is α, what is the probability that the satellite system will function?

Solution:

$$\alpha \sum_{i=k}^{n} \binom{n}{i} p_1^i (1-p_1)^{n-i} + (1-\alpha) \sum_{i=k}^{n} \binom{n}{i} p_2^i (1-p_2)^{n-i}.$$

Exercise 46.

Suppose that a biased coin that lands on heads with probability p is flipped 10 times. Given that a total of 6 heads results, find the conditional probability that the first 3 outcomes are

(a) *h, t, t (meaning that the first flip results in heads, the second in tails, and the third in tails);*

(b) *t, h, t.*

Solution:

(a)

$$\begin{aligned} P\{H, T, T|6 \text{ heads}\} &= P\{H, T, T \text{ and } 6 \text{ heads}\}/P\{6 \text{ heads}\} \\ &= P\{H, T, T\}P\{6 \text{ heads}|H, T, T\}/P\{6 \text{ heads}\} \\ &= pq^2 \binom{7}{5} p^5 q^2 \Big/ \binom{10}{6} p^6 q^4 = 1/10. \end{aligned}$$

(b)

$$\begin{aligned} P\{T, H, T|6 \text{ heads}\} &= P\{T, H, T \text{ and } 6 \text{ heads}\}/P\{6 \text{ heads}\} \\ &= P\{T, H, T\}P\{6 \text{ heads}|T, H, T\}/P\{6 \text{ heads}\} \\ &= q^2 p \binom{7}{5} p^5 q^2 \Big/ \binom{10}{6} p^6 q^4 = 1/10. \end{aligned}$$

Exercise 47.

The monthly worldwide average number of airplane crashes of commercial airlines is 3.5. What is the probability that there will be

(a) *at least 2 such accidents in the next month;*

(b) *at most 1 accident in the next month?*

Explain your reasoning!

Solution:

(a) $1 - e^{-3.5} - 3.5e^{-3.5} = 1 - 4.5e^{-3.5}$

(b) $4.5e^{-3.5}$.

Exercise 48.

Consider n independent trials, each of which results in one of the outcomes $1, \ldots, k$ with respective probabilities $p_1, \ldots, p_k$, $\sum_{i=1}^{k} p_i = 1$. Show that if all the p_i are small, then the probability that no trial outcome occurs more than once is approximately equal to $\exp(-n(n-1)\sum_i p_i^2/2)$.

Solution: If A_i is the event that couple number i are seated next to each other, then these events are, when n is large, roughly independent. As $P(A_i) = 2/(2n-1)$ it follows that, for n large, the number of wives that sit next to their husbands is approximately Poisson with mean $2n/(2n-1) \approx 1$. Hence, the desired probability is $e^{-1} = .368$ which is not particularly close to the exact solution of .2656, thus indicating that $n = 10$ is not large enough for the approximation to be a good one.

Exercise 49.

People enter a gambling casino at a rate of 1 every 2 minutes.

(a) *What is the probability that no one enters between* 12 : 00 *and* 12 : 05*?*

(b) *What is the probability that at least* 4 *people enter the casino during that time?*

Solution:

(a) $e^{-2.5}$

(b) $1 - e^{-2.5} - 2.5e^{-2.5} - \frac{(2.5)^2}{2}e^{-2.5} - \frac{(2.5)^3}{3!}e^{-2.5}$.

Exercise 50.

A total of 2n people, consisting of n married couples, are randomly seated (all possible orderings being equally likely) at a round table. Let C_i denote the event that the members of couple i are seated next to each other, $i = 1, \ldots, n$.

1. *Find $P(C_i)$.*
2. *For $j \neq i$, find $P(C_j|C_i)$.*
3. *Approximate the probability, for n large, that there are no married couples who are seated next to each other.*

Solution: Assume $n > 1$.

1. $\frac{2}{2n-1}$
2. $\frac{2}{2n-2}$
3. $\exp\{-2n/(2n-1)\} \approx e^{-1}$.

Exercise 51.

In response to an attack of 10 missiles, 500 antiballistic missiles are launched. The missile targets of the antiballistic missiles are independent, and each antiballstic missile is equally likely to go towards any of the target missiles. If each antiballistic missile independently hits its target with probability .1, use the Poisson paradigm to approximate the probability that all missiles are hit.

Solution:

$\exp\{-10e^{-5}\}$.

Exercise 52.

Two athletic teams play a series of games; the first team to win 4 games is declared the overall winner. Suppose that one of the teams is stronger than the other and wins each game with probability .6, independently of the outcomes of the other games. Find the probability, for $i = 4, 5, 6, 7$, that the stronger team wins the series in exactly i games. Compare the probability that the stronger team wins with the probability that it would win a 2-out-of-3 series.

Solution:

$$P\{\text{wins in } i \text{ games}\} = \binom{i-1}{3}(.6)^4(.4)^{i-4}.$$

Exercise 53.

In the Banach matchbox problem, find the probability that, at the moment when the first box is emptied (as opposed to being found empty), the other box contains exactly k matches.

Solution:

$$2\binom{2N-k}{N}(1/2)^{2N-k}$$

$$2\binom{2N-k-1}{N-1}(1/2)^{2N-k-1}(1/2)$$

Exercise 54.

A purchaser of transistors buys them in lots of 20. It is his policy to randomly inspect 4 components from a lot and to accept the lot only if all 4 are nondefective. If each component in a lot is, independently, defective with probability .1, what proportion of lots is rejected?

Exercise 55.

There are three highways in the county. The number of daily accidents that occur on these highways are Poisson random variables with respective parameters .3, .5, and .7. Find the expected number of accidents that will happen on any of these highways today.

Exercise 56.

Suppose that 10 *balls are put into* 5 *boxes, with each ball independently being put in box* i *with probability* p_i, $\sum_{i=1}^{5} p_i = 1$.

(a) *Find the expected number of boxes that do not have any balls.*

(b) *Find the expected number of boxes that have exactly* 1 *ball.*

Exercise 57.

There are k types of coupons. Independently of the types of previously collected coupons, each new coupon collected is of type i with probability p_i, $\sum_{i=1}^{k} p_i = 1$. If n coupons are collected, find the expected number of distinct types that appear in this set. (That is, find the expected number of types of coupons that appear at least once in the set of n coupons.)

Exercise 58.

If X has distribution function F, what is the distribution function of e^X?

Solution:

Exercise 59.

If X has distribution function F, what is the distribution function of the random variable $\alpha X + \beta$, where α and β are constants, $\alpha \neq 0$?

Solution: $1 - \lim_{h \to 0} F(a - h)$

Exercise 60.

For a nonnegative integer-valued random variable N, show that

$$\sum_{i=0}^{\infty} iP\{N > i\} = \frac{1}{2}(E[N^2] - E[N]).$$

Solution:

$$\begin{aligned}\sum_{i=1}^{\infty} P\{N \geq i\} &= \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} P\{N = k\} \\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} P\{N = K\} \\ &= \sum_{k=1}^{\infty} kP\{N = k\} = E[N].\end{aligned}$$

Exercise 61.

Let X be such that

$$P\{X = 1\} = p = 1 - P\{X = -1\}$$

Find $c \neq 1$ such that $E[c^X] = 1$.

Solution:

$E[c^X] = cp + c^{-1}(1 - p)$.

Hence, $1 = E[c^X]$ if $cp + c^{-1}(1 - p) = 1$, or, equivalently $pc^2 - c + 1 - p = 0$ or $(pc - 1 + p)(c - 1) = 0$.

Thus, $c = (1 - p)/p$.

Exercise 62.

Find $Var(X)$ *if*

$$P(X = a) = p = 1 - P(X = b)$$

Exercise 63.

Let X be a binomial random variable with parameters (n, p). What value of p maximizes $P\{X = k\}$, $k = 0, 1, \ldots, n$? This is an example of a statistical method used to estimate p when a binomial (n, p) random variable is observed to equal k. If we assume that n is known, then we estimate p by choosing that value of p which maximizes $P\{X = k\}$. This is known as the method of maximum likelihood estimation.

Solution: Easiest to first take log and then determine the p that maximizes $\log P\{X = k\}$.

$$\log P\{X = k\} = \log \binom{n}{k} + k \log p + (n - k) \log(1 - p)$$

$$\frac{\partial}{\partial p} \log P\{x = k\} = \frac{k}{p} - \frac{n - k}{1 - p} = 0 \implies p = k/n \text{ maximizes.}$$

Exercise 64.

Let X be a Poisson random variable with parameter λ. Show that $P\{X = i\}$ increases monotonically and then decreases monotonically as i increases, reaching its maximum when i is the largest integer not exceeding λ.

Exercise 65.

Show that X is a Poisson random variable with parameter λ, then

$$E[X^n] = \lambda E[(X+1)^{n-1}]$$

Now use this result to compute $E[X^3]$.

Solution:

$$\begin{aligned} E[X^n] &= \sum_{i=0}^{\infty} i^n e^{-\lambda}\lambda^i/i! = \sum_{i=1}^{\infty} i^n e^{-\lambda}\lambda^i/i! \\ &= \sum_{i=1}^{\infty} i^{n-1} e^{-\lambda}\lambda^i/(i-1)! = \sum_{j=0}^{\infty}(j+1)^{n-1}e^{-\lambda}\lambda^{j+1}/j! \\ &= \lambda\sum_{j=0}^{\infty}(j+1)^{n-1}e^{-\lambda}\lambda^j/j! \\ &= \lambda E[(X+1)^{n-1}]. \end{aligned}$$

Solution (Contd...)

Hence

$$
\begin{aligned}
E[X^3] &= \lambda E[(X+1)^2] \\
&= \lambda \sum_{i=0}^{\infty} (i+1)^2 e^{-\lambda} \lambda^i / i! \\
&= \lambda \left[\sum_{i=0}^{\infty} i^2 e^{-\lambda} \lambda^i / i! + 2 \sum_{i=0}^{\infty} i e^{-\lambda} \lambda^i / i! + \sum_{i=0}^{\infty} e^{-\lambda} \lambda^i / i! \right] \\
&= \lambda [E[X^2] + 2E[X] + 1] \\
&= \lambda (Var(X) = E^2[X] + 2E[X] + 1) \\
&= \lambda (\lambda + \lambda^2 + 2\lambda + 1) = \lambda (\lambda^2 + 3\lambda + 1).
\end{aligned}
$$

Exercise 66.

Consider a random collection of n individuals. In approximating the probability that no 3 of these individuals share the same birthday, a better Poisson approximation than that obtained in the text (at least for values of n between 80 and 90) is obtained by letting E_i be the event that there are at least 3 birthdays on day i, $i = 1, \ldots, 365$.

(a) *Find $P(E_i)$.*

(b) *Give an approximation for the probability that no 3 individuals share the same birthday.*

(c) *Evaluate the preceding when $n = 88$ (which can be shown to be the smallest value of n for which the probability exceeds .5).*

Solution:

(a) $P(E_i) = 1 - \sum_{j=0}^{2} \binom{365}{j} (1/365)^j (364/365)^{365-j}$

(b) $\exp\{-365P(E_1)\}$.

Exercise 67.

Prove

$$\sum_{i=0}^{n} e^{-\lambda} \frac{\lambda^i}{i!} = \frac{1}{n!} \int_{\lambda}^{\infty} e^{-x} x^n dx.$$

Exercise 68.

For a hypergeometric random variable, determine

$$P\{X = k+1\}/P\{X = k\}.$$

Solution:

$$\frac{P\{X = k+1\}}{P\{X = k\}} = \frac{\binom{Np}{k+1}\binom{N-np}{n-k-1}}{\binom{Np}{k}\binom{N-Np}{n-k}} = \frac{(Np-k)(n-k)}{(k+1)(N-Np-n+k+1)}.$$

Exercise 69.

A jar contains n chips. Suppose that a boy successively draws a chip from the jar, each time replacing the one drawn before drawing another. The process continues until the boy draws a chip that he has previously drawn. Let X denote the number of draws, and compute its probability mass function.

Solution:

$$P\{X = k\} = \frac{k-1}{n} \prod_{i=0}^{k-2} \frac{n-i}{n}, k > 1.$$

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.