Bernoulli Trials and Related Probability Distributions

BERNOULLI TRIALS and RELATED PROBABILITY DISTRIBUTIONS

A. BERNOULLI TRIALS
Consider tossing a coin several times. It is generally agreed that the following apply here.

1). Each time the coin is tossed there are two possible outcomes that can be observed: Heads and Tails.

2). At each toss of the coin, the probability of observing Heads is the same, namely one-half. That is, there is a constant probability of observing Heads.

3). What happens (i.e. Heads or Tails) on any one toss of the coin does not affect what happens on another toss. That is, outcomes at successive tosses are independent of each other.

Repeated trials of a simple experiment having these three properties are called Bernoulli Trials, and an experiment of this type is called a Bernoulli Experiment.

Definition: Bernoulli Trials are repeated trials of a simple experiment having the following three properties.

(1). At each trial there are two possible outcomes, commonly referred to as Success (the event of interest or being studied occurs) and Failure (the event does not occur).

(2). The probability of a Success is constant at each trial - and hence so is the probability of Failure. The probability of Success is usually denoted \( p \) and that of Failure \( q = 1 - p \).

(3). Outcomes at successive trials are independent of one another.

Example 1: Suppose a coin is tossed several times. What is the probability that Heads is observed for the first time on the fourth toss of the coin?

Let H denote Heads (Success) and T denote Tails (Failure). Then \( p = 1/2 \) and \( q = 1/2 \). In order for Heads (Success) to occur for the first time on the fourth toss, the sequence of outcomes must have the form TTTH or FFFS. [Note: We don't need to worry about outcomes at the fifth and later tosses of the coin since they do not impact on the stated event.] The probability of the sequence TTTH is determined as follows.

\[
P[TTTH] = P[T_1 \cap T_2 \cap T_3 \cap H_4]
\]

in which the subscript refers to the trial number. Because of the independence assumption (property (3)) which tells us that, if \( A \) and \( B \) are independent events then \( P[A \cap B] = P[A]P[B] \) (with this multiplication property being extendable to more than two events), this becomes

\[
\]

Property (2) of constant probabilities then allows us to obtain

\[
= P[T_1]P[T_2]P[T_3]P[H_4]
= qqqp = \frac{1 \cdot 1 \cdot 1}{2 \cdot 2 \cdot 2} = \left( \frac{1}{2} \right)^4 = \frac{1}{16}.
\]
Example 2: Suppose a die is rolled several times. What is the probability that a "multiple of 3" is observed for the first time on the fourth roll? This is similar to Example 1.

Let S denote the event "a multiple of 3 is observed" so that F is the event "a multiple of 3 is NOT observed". The multiples of 3 that are possible when rolling a die are "3" and "6". Hence \[ p = P[S] = \frac{2}{6} = \frac{1}{3} \] and \[ q = P[F] = 1 - p = 2/3. \] The required probability is then determined to be

\[
P[FFFS] = P[F_1 \cap F_2 \cap F_3 \cap S_4] = P[F_1]P[F_2]P[F_3]P[S_4] = q^3p = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{8}{81}.
\]

B. Situations resulting in Bernoulli trials.

Bernoulli trials are considered to exist in the following situations.

a). In situations like tossing a coin or rolling a die in which the number of possible outcomes is obviously fixed from trial to trial (e.g. the numbers on a die do not disappear once seen). One might construct dice that have different numbers of sides. For example, a pyramid built with equilateral triangles is like a 4-sided die.

b). If items are drawn one-after-another with replacement from a group of \( N \) items, and if \( M \) of the items have a certain property, then Success can be defined as obtaining an item with the certain property and its probability is then \( p = M/N \).

c). In the previous case, if items are selected one-after-another without replacement but relatively few are drawn in comparison to the sizes of \( N, M \) and \( N-M \), then it is usually assumed that the Bernoulli trials situation is a very good approximate model. [See the section on the Hypergeometric Distribution for the case in which the number drawn is not considered to be small relative to these values.] A useable guideline is to require that the number drawn \( n \) be not more than 5% of either \( M \) or \( N-M \).

d). If information is provided about the proportion or percentage (i.e. \( p \) or \( 100p\% \)) of the elements of a population of infinite size that have a certain property, then sampling from such a population is effectively the same as sampling with replacement.

e). If information is provided about the proportion or percentage (i.e. \( p \) or \( 100p\% \)) of the elements of a very large population that have a certain property but no mention is made about the actual size of the population, assume that the population is infinite as in the previous case.

Example 3: Suppose that, within the adult population of Canada (say of size \( N = 20,000,000 \)), 70\% of the adult citizens believe that governments are inefficient in their handling of taxpayers' money. Thus \( n = 14,000,000 \) have the property of having this opinion. If three people are selected at random and without replacement one-after-another from this population, what is the probability that only the second selected person believes this inefficiency idea?

Let S denote the event that a selected person believes this inefficiency idea. The sequence of selections must then result in FSF and the probability required here is \( P[FSF] \). If it is assumed that the Bernoulli Trials model provides a good approximation, then \( p = 0.7 \) and \( q = 0.3 \) and
How close to the correct answer is the approximation obtained assuming a Bernoulli Trials model? Noting that sampling is actually done without replacement, the effective population size changes after each selection and the number of elements having the certain property may also change after a selection. Hence conditional probabilities are required. The solution is then obtained as

\[
P[FSF] = P[F_1 \cap S_2 \cap F_3]
= P[F_1]P[S_2 | F_1]P[F_3 | (F_1 \cap S_2)]
= \frac{6,000,000}{20,000,000} \times \frac{14,000,000}{19,999,999} \times \frac{5,999,999}{19,999,998}
= 0.062999999
\]

so that the Bernoulli Trials solution gives a value very close to this correct value suggesting that it is a useful approximate method.

When can one assume that the Binomial Trials solution (which effectively assumes sampling is done with replacement) provides an adequate approximation to the correct value based on sampling without replacement? A reasonable guideline to follow (or "rule-of-thumb") is that the approximation is adequate if both the population size \(N\), the number having the property \(M\) and the number not having the property \(N - M\) are all "large" with the number of trials or selections \(n\) being not more than 5% of \(M\) or \(N - M\). In Example 3 above, \(N = 20,000,000\), \(M = 14,000,000\), \(N - M = 6,000,000\) and the number \(n = 3\) selected is a very tiny percentage of each of these values.

**Exercises:**

1. A die is rolled repeatedly until the number "5" is first seen. What is the probability that exactly 8 rolls are required? [Answer: 0.046513608]

2. A die is rolled repeatedly until a "multiple of 3" is seen exactly twice. What is the probability that exactly four rolls are required? [Answer: 0.148148148]

3. Fifty-five percent of the students registered at a very large Canadian university are female. If four students are selected one-after-another at random and without replacement from the student body of this university, what is the probability that the sexes of the selected students will alternate between male and female (or female and male)? [Answer: 0.1225125]

**C. BINOMIAL EXPERIMENTS and the BINOMIAL DISTRIBUTION**

A Binomial experiment is a Bernoulli Trials experiment in which the number of trials is fixed in advance. For example, a coin is tossed 5 times, or a four dice are rolled simultaneously, or a
sample of 1,000 people is taken from the adult population of Canada. The number of trials (here 5, 4 and 1000 respectively) are known/fixed in advance.

**A Binomial Experiment** has the following four properties.

1. There are a fixed number \( n \) of repeated trials of a simple experiment.
2. At each trial there are two possible outcomes, commonly referred to as *Success* (the event of interest or being studied occurs) and *Failure* (the event does not occur).
3. The probability of a Success is constant at each trial - and hence so is the probability of Failure. The probability of Success is usually denoted \( p \) and that of Failure \( q = 1 - p \).
4. Outcomes at successive trials are independent of one another.

When a Binomial experiment is performed, the usual random variable of interest is "the number of Successes observed in the \( n \)-trial Binomial experiment". If \( X \) denotes such a random variable, how are probabilities of events such as \( X = 3 \) found in such a case? What is the probability distribution of the random variable \( X \) here? The following two examples will help in understanding the result.

**Example 4:** A coin is tossed 5 times. What is the probability that exactly 3 Heads are observed in the 5 tosses? Let a random variable be defined as \( X = \) "the number of Heads observed in 5 tosses of a coin". Then what is \( P[X = 3] \)?

One way in which 5 tosses of a coin can result in exactly 3 Heads (i.e. Successes) is if the sequence HHTTH occurs (note that this is just one possible sequence). What is the probability of this particular sequence? The sequence can be considered as the outcome of five Bernoulli trials in which \( p = q = 1/2 \) and its probability found as in examples 1, 2 and 3. Hence

\[
P[HHTTH] = P[H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5] = P[H_1]P[H_2]P[H_3]P[H_4]P[H_5] = ppqqp = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{32}.
\]

But there are other sequences that also lead to exactly 3 Heads. For example, TTHHH and HTHTH are two others. Note that each of these sequences each involves three H's and two T's and has exactly the same probability as our first sequence, namely \( 1/32 \). The required probability \( P[3 \text{ Heads in } 5 \text{ tosses}] = P[X = 3] \) is thus the sum of the probability \( 1/32 \) as many times as there are different sequences involving 3 Heads and 2 Tails. A complete listing of the 10 possible sequences is as follows.

HHHTT, HHTHT, HHTTH, HTHTH, HTHHT, HTHTH, HTTHH, THHHT, THHTH, TTHHH

Each of these 10 sequences has the probability \( 1/32 \) so \( P[X = 3] = 10/32 \), the sum of \( 1/32 \) ten times.

Is there a systematic way of counting the number of sequences each of which has the same probability? The answer is yes and the number is determined as follows. The sequence of five Bernoulli trials is to involve exactly 3 Successes. Each way that 3 of the five trials can be selected so as to be the trials at which Success occurs yields a different sequence. The number of
ways that 3 trials (or 3 objects) can be selected from 5 trials (or 5 objects) is $\binom{5}{3} = 10$.

Thus the answer to this problem could be found as

$$P[3 \text{ } H's \text{ in } 5 \text{ trials}] = \binom{5}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32}.$$ 

One simplicity in this example is that $p = q = 1/2$. The next example is more typical in that the values of $p$ and $q$ are not the same.

**Example 5:** Four dice are rolled simultaneously. What is the probability that exactly two "multiples of 3" are observed on the four dice? If a random variable is defined as $Y = "the number of "multiples of 3" observed on four dice"$, then what is $P[Y = 2]$?

One sequence that gives exactly two "multiples of 3" is SFSF, where Success denotes obtaining a "multiple of 3". The probability of this sequence is

$$P[SFSF] = P[S_1 \cap F_2 \cap S_3 \cap F_4] = P[S_1]P[F_2]P[S_3]P[F_4] = pqpq = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{4}{81}.$$ 

Adopting the counting approach in example 4, the number of different sequences each having two Successes equals the number of ways that 2 trials can be selected from the 4 Bernoulli trials; that is $\binom{4}{2} = 6$. Hence $P[Y = 2] = 6 \times \frac{4}{81} = \frac{24}{81} = \frac{8}{27}$. Alternatively,

$$P[Y = 2] = \binom{4}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 = \frac{24}{81}.$$ 

**The BINOMIAL DISTRIBUTION**

Let $X$ be a random variable counting the number of Successes observed in a Binomial experiment having $n$ trials with a constant probability of Success $p$. The possibilities for the number of successes observed are $0, 1, 2, ..., n$ so that the value space for this random variable is $V_X = \{0,1,2,\cdots,n\}$ and the probability mass function for $X$ is

$$p_X(x) = P[X = x] = \binom{n}{x} p^x q^{n-x} \quad (1.1)$$

for each $x \in V_X = \{0,1,2,\cdots,n\}$. A short-hand indication of this probability distribution is $X \sim B(n,p)$. 

Page 5 of 15
Comments:

(1). The parameters of this distribution are \( n \) and \( p \). Knowing them and that \( X \sim B(n, p) \) is enough to be able to proceed with probability calculations.

(2). The expression (1.1) is obtainable as follows. In order for the random variable \( X \) to have the particular value \( x \), the \( n \) Bernoulli trials must result in exactly \( x \) Successes and hence exactly \( n - x \) Failures. The probability of any one such sequence is then \( p^x q^{n-x} \). The number of different sequences of \( n \) trials that have \( x \) Successes and \( n - x \) Failures equals the number of ways of selecting \( x \) of the \( n \) trials to be "Success" trials with the remaining \( n - x \) being "Failure" trials, that is \( \binom{n}{x} \). Each of these has the probability \( p^x q^{n-x} \) so that \( p_X(x) = \binom{n}{x} p^x q^{n-x} \).

**Example 6:** Suppose a die is rolled 6 times and Success is defined as obtaining a "multiple of 3". If \( X \) counts the number of successes in the 6 rolls, then \( X \sim B(6, 1/3) \). The probability distribution of this random variable can be presented as in the following table.

<table>
<thead>
<tr>
<th>Possible value: ( x )</th>
<th>Probability Formula: ( p_X(x) = P[X = x] = \left(\binom{6}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{6-x}\right) )</th>
<th>Calculated Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( \left(\binom{6}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^6\right) = \frac{64}{729} )</td>
<td>0.087791</td>
</tr>
<tr>
<td>1</td>
<td>( \left(\binom{6}{1} \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^5\right) = 6 \times \frac{32}{729} = \frac{192}{729} )</td>
<td>0.263374</td>
</tr>
<tr>
<td>2</td>
<td>( \left(\binom{6}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^4\right) = 15 \times \frac{16}{729} = \frac{240}{729} )</td>
<td>0.329218</td>
</tr>
<tr>
<td>3</td>
<td>( \left(\binom{6}{3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^3\right) = 20 \times \frac{8}{729} = \frac{160}{729} )</td>
<td>0.219479</td>
</tr>
<tr>
<td>4</td>
<td>( \left(\binom{6}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^2\right) = 15 \times \frac{4}{729} = \frac{60}{729} )</td>
<td>0.082305</td>
</tr>
<tr>
<td>5</td>
<td>( \left(\binom{6}{5} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^1\right) = 6 \times \frac{2}{729} = \frac{12}{729} )</td>
<td>0.016461</td>
</tr>
<tr>
<td>6</td>
<td>( \left(\binom{6}{6} \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^0\right) = \frac{1}{729} )</td>
<td>0.001372</td>
</tr>
<tr>
<td>Total</td>
<td>( \frac{729}{729} = 1 )</td>
<td>1.000000</td>
</tr>
</tbody>
</table>
Example 7: A professional "darts" player successfully hits the region of the dart board at which he aims with constant probability of 0.90. Because of his experience in playing the game, it is reasonable to assume that his results on successive throws (i.e. hitting the region at which he aims, or not) are independent events. What is the probability that in his next nine throws he hits the region at which he is aiming exactly six times? What is the probability that in these nine throws he hits his targeted region at least eight times?

Let $U$ be a random variable counting the number of times the dart player hits the region at which he is aiming in nine throws of the darts. Then $U \sim B(9, 0.9)$.

(i). What is the probability that in his next eight throws he hits the region at which he is aiming exactly six times?

$$P[U = 6] = \binom{9}{6} (0.9)^6 (0.1)^3 = 0.04464$$

(ii). What is the probability that in these nine throws he hits his targeted region at least eight times?

$$P[U \geq 8] = \sum_{u=8}^{9} \binom{9}{u} (0.9)^u (0.1)^{9-u}$$

$$= \binom{9}{8} (0.9)^8 (0.1)^1 + \binom{9}{9} (0.9)^9 (0.1)^0$$

$$= 0.38742 + 0.038742 = 0.426164$$

Example 8: Recent reports suggest that only 20% of Canadians are for changes to the medicare system that would allow privately run hospitals. A random sample of 12 Canadians is obtained. Find the probabilities that, for this random sample, (i) exactly 2 are for such changes; (ii) at most 2 are for such changes; (iii) more than two are for such changes.

Let $V$ be a random variable that counts the number in the sample of size 12 who are for such changes. Then $V \sim B(12, 0.2)$.

(i). exactly 2 are for such changes

$$P[V = 2] = \binom{12}{2} (0.2)^2 (0.8)^{10} = 0.2835$$

(ii). at most 2 are for such changes

$$P[V \leq 2] = \sum_{v=0}^{2} \binom{12}{v} (0.2)^v (0.8)^{12-v}$$

$$= \binom{12}{0} (0.2)^0 (0.8)^{12} + \binom{12}{1} (0.2)^1 (0.8)^{11} + \binom{12}{2} (0.2)^2 (0.8)^{10}$$

$$= 0.06872 + 0.20616 + 0.28347 = 0.5583$$

(iii). more than two are for such changes

$$P[V > 2] = 1 - P[V \leq 2] = 1 - 0.5583 = 0.4417$$
Mean and Standard Deviation of a Binomial Distribution

If $X$ is a Binomially distributed random variable with parameters $n$ and $p$, that is $X \sim B(n, p)$, then its mean (expected value), variance and standard deviations are defined respectively as

$$
\mu = E[X] = \sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x},
$$

$$
\sigma^2 = Var[X] = E[(X - \mu)^2] = \sum_{x=0}^{n} (x - \mu)^2 p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \text{ and }
$$

$$
\sigma = SD[X] = \sqrt{\sigma^2}.
$$

These are rather complicated looking expressions but can be evaluated in given situations. The following table outlines how these calculations can be done in table form for the random variable given in Example 6. [Comment: The calculations in the column headed "$(x - \mu)^2 p(x) = "$ are easier in this example than they usually are since the mean $\mu$ has an integer value.]

<table>
<thead>
<tr>
<th>Value: $x =$</th>
<th>Probability: $p(x) =$</th>
<th>$x \cdot p(x) =$</th>
<th>$(x - \mu)^2 \cdot p(x) =$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\frac{64}{729}$</td>
<td>$0 \times \frac{64}{729} = \frac{0}{729}$</td>
<td>$(0 - \mu)^2 \times \frac{64}{729} = \frac{256}{729}$</td>
</tr>
<tr>
<td>1</td>
<td>$\frac{192}{729}$</td>
<td>$1 \times \frac{192}{729} = \frac{192}{729}$</td>
<td>$(1 - \mu)^2 \times \frac{192}{729} = \frac{192}{729}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{240}{729}$</td>
<td>$2 \times \frac{240}{729} = \frac{480}{729}$</td>
<td>$(2 - \mu)^2 \times \frac{240}{729} = \frac{0}{729}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{160}{729}$</td>
<td>$3 \times \frac{160}{729} = \frac{480}{729}$</td>
<td>$(3 - \mu)^2 \times \frac{160}{729} = \frac{160}{729}$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{60}{729}$</td>
<td>$4 \times \frac{60}{729} = \frac{240}{729}$</td>
<td>$(4 - \mu)^2 \times \frac{60}{729} = \frac{240}{729}$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{12}{729}$</td>
<td>$5 \times \frac{12}{729} = \frac{60}{729}$</td>
<td>$(5 - \mu)^2 \times \frac{12}{729} = \frac{108}{729}$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{1}{729}$</td>
<td>$6 \times \frac{1}{729} = \frac{6}{729}$</td>
<td>$(6 - \mu)^2 \times \frac{1}{729} = \frac{16}{729}$</td>
</tr>
<tr>
<td>Total</td>
<td>$\frac{729}{729} = 1$</td>
<td>$\mu = \frac{1458}{729} = 2.00$</td>
<td>$\sigma^2 = \frac{972}{729} = \frac{4}{3}$</td>
</tr>
</tbody>
</table>

Fortunately it is not necessary to go through this amount of work. The following theorem indicates how the mean, variance and standard deviation can always be obtained when dealing with a Binomial random variable $X \sim B(n, p)$. 
Theorem 1:
Suppose $X$ is a Binomial random variable with parameters $n$ and $p$; that is, $X \sim B(n, p)$.
Then the mean or expected value of $X$ is $\mu = np$,
the variance of $X$ is $\sigma^2 = npq$, and
the standard deviation of $X$ is $\sigma = \sqrt{npq}$.

Applying these results in the previous example gives
$$
\mu = np = 6 \times \frac{1}{3} = 2 \quad \sigma^2 = npq = 6 \times \frac{1}{3} \times \frac{2}{3} = \frac{4}{3} \quad \sigma = \sqrt{npq} = \sqrt{\frac{4}{3}} \approx 1.1547.
$$

Comment: A Binomial random variable $X$ counts the number of successes in $n$ trials. Hence the mean or expected value may be referred to as the expected number of successes. For example, what is the expected number of Heads if a coin is tossed 9 times? The number of Heads in 9 tosses of a coin is a $B(9, 0.5)$ random variable, so the expected number of Heads is $\mu = E[X] = np = 9 \times 0.5 = 4.5$. The expected number of Heads is the mean number of Heads to be expected when a fair coin is tossed 9 times. Remember that the mean or expected number need not be a possible number (as in this case in which the expected number 4.5 is not a number of Heads that would ever be achieved in 9 tosses of a coin).

Exercises 2:

(1). In attempting free throws during basketball games, a certain player has a constant probability of 0.58 of making the shot. Furthermore his successes (or lack of same) from shot to shot are independent events. What is the probability that he makes exactly three-quarters of the 12 free throws he attempts one game? What is the expected number of free throws that he would make in these 12 attempts? What is the standard deviation of the number of successful free throws in 12 attempts?

(2). An urn contains 5 red balls and 4 blue balls. Balls are drawn one after another with replacement from this urn. Suppose a total of 7 draws are made. What is the probability that red balls are drawn exactly 4 times? What is the probability that red balls are drawn at most once? What is the probability that red balls are drawn at least twice? What is the expected number of red balls in the 7 draws?

B. The GEOMETRIC DISTRIBUTION

A Binomial experiment involves a fixed number $n$ of Bernoulli trials. Another interesting experiment is to perform Bernoulli trials repeatedly until a "Success" is observed for the first time. The question of interest is: how many rolls are required for this to happen?

Example 9: A die is rolled repeatedly until the first time that a "multiple of 3" (a Success) is observed. How many rolls are required? What is the probability that exactly five rolls are
required? In these Bernoulli trials, \( p = \frac{1}{3} \) is the probability of observing a "multiple of 3" at any particular roll. If the first Success occurs on the fifth roll, then each of the first four rolls must have resulted in Failure. Thus the sequence of outcomes would be \( FFFFFS \). The probability of this sequence is just

\[
P[FFFFS] = P[F_1 \cap F_2 \cap F_3 \cap F_4 \cap S_5] = P[F_1]P[F_2]P[F_3]P[F_4]P[S_5] = \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right) = \frac{16}{243}.
\]

See examples 1, 2 and 3 for similar Bernoulli trials solutions.

In the above example, a random variable \( Y \) could be defined by "\( Y = \) the roll number on which a multiple of 3 is first obtained". The probability calculated here is then

\[
p_Y(5) = P[Y = 5] = q^4 p = \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right) = \frac{16}{243}.
\]

This random variable \( Y \) is said to have a Geometric distribution with single parameter \( p = \frac{1}{3} \). Its value space is the set of all possible values for \( Y \), namely all possible numbers of trials that could be required in order to obtain a Success for the first time. Thus

\[
V_Y = \{1,2,3,\ldots\},
\]

the set of all positive integers. The number of trials required could be any positive integer without upper bound. It is conceivable (but unlikely) to continue to roll the die repeatedly without ever observing a Success, and hence the value space has no upper limit.

If \( y \in V_Y = \{1,2,3,\ldots\} \), what is \( p_Y(y) \equiv P[Y = y] \)? In order for the \( y \)th Bernoulli trial to be the trial at which the first Success occurs, all \( y - 1 \) previous trials must result in Failure. Thus the sequence of outcomes yielding this is \( FFFFFS \) in which there are \( y - 1 \) F’s followed by one S. Thus

\[
p_Y(y) \equiv P[Y = y] = q^{y-1} p
\]

for \( y \in V_Y = \{1,2,3,\ldots\} \).

Definition: A random variable \( Y \) is said to have a Geometric distribution with parameter \( p \) - that is, \( Y \sim Ge(p) \) - if its value space is \( V_Y = \{1,2,3,\ldots\} \) and its probability function is given by

\[
p_Y(y) \equiv P[Y = y] = q^{y-1} p
\]

for \( y \in V_Y = \{1,2,3,\ldots\} \).

Comment: Since the experiment is to wait until the first Success occurs, this distribution might be referred to as a waiting time distribution.

**Example 10:** A student writing a multiple choice test is faced with 10 answer choices for each question. If he randomly selects an answer for each question (that is, he guesses), what is the probability that the first correct answer is obtained on the seventh question?
Let $Y$ be a random variable counting the number of guesses until he finally gets one right. Assuming his guesses are at random and done independently from question to question, it follows that $Y \sim Ge(1/10)$ and $p_Y(7) \equiv P[Y = 7] = \left( \frac{9}{10} \right)^{7-1} \left( \frac{1}{10} \right) = \frac{531,554}{10,000,000} = 0.05314$.

**Example 11:** Seventy percent of students hold jobs while attending university. Suppose students are selected one-after-another at random from a large class. What is the probability that the first student selected who does NOT hold a job this year is the fourth student selected?

The question asks about finding a student who does NOT hold a job. Thus Success is defined as finding such a student. Selecting students at random will be assumed to give us the "independence" of outcomes at various trials (selections), and the probability of a success at each trial is $0.30$ ($70\%$ hold jobs so $30\%$ do not). Letting $W \sim Ge(0.3)$, the required probability is $p_W(4) \equiv P[W = 4] = (0.7)^{4-1}(0.3) = 0.1029$.

How long should one expect to have to wait until the first success is observed? That is, what is the mean or expected value of a Geometric random variable?

**Theorem 2:**

Suppose $Y$ is a Geometric random variable with parameter $p$; that is, $Y \sim Ge(p)$. Then

- the mean or expected value of $Y$ is $\mu = \frac{1}{p}$,
- the variance of $Y$ is $\sigma^2 = \frac{q}{p^2}$, and
- the standard deviation of $Y$ is $\sigma = \frac{\sqrt{q}}{p}$.

**Example 11 (continued):**

How many students would one expect to have to select until one who does not work while attending university is found? If this procedure is repeated for different classes, what is the standard deviation of the numbers of students, from class to class, who would have to be selected in order to find one who does not work?

The mean or expected number is $\mu = \frac{1}{p} = \frac{1}{0.3} = \frac{10}{3} = 3.3333$ and the standard deviation is $\sigma = \frac{\sqrt{q}}{p} = \frac{\sqrt{0.7}}{(0.3)^2} = \frac{\sqrt{0.7}}{0.09} = \sqrt{7.7778} = 2.7889$. Thus, on average, 3.3333 students would have to be selected in order to find one who does not work, but there is a relatively large amount of variability from attempt to attempt as measured by the standard deviation of 2.7889 (which is 83.67\% of the mean).

Comments:
(1). The definition of a probability distribution requires that the probability function $p_Y(y)$ satisfy the two properties a) $p_Y(y) \geq 0$ for every $y \in V_Y$ and b) $\sum_{y \in V_Y} p_Y(y) = 1$. In the case of the Geometric distribution, the first of these is clearly satisfied since $p_Y(y) = q^{y-1}p$ is greater than zero since both of $p$ and $q$ are. Examining the second property produces the following summation.

$$\sum_{y \in V_Y} p_Y(y) = \sum_{y=1}^{\infty} q^{y-1}p = q^0p + q^1p + q^2p + q^3p + \cdots$$

$$= p(1 + q + q^2 + q^3 + \cdots)$$

In order for this sum to equal 1, the sum $1 + q + q^2 + q^3 + \cdots$ must equal $1/p$ for then the above would become $p(1 + q + q^2 + q^3 + \cdots) = \frac{1}{p} = 1$ as required. Since $p = 1 - q$, the ratio $1/p$ is equivalent to $\frac{1}{1-q}$ and thus $\frac{1}{1-q} = 1 + q + q^2 + q^3 + \cdots$. This latter expression is called a "series expansion" for $1/(1-q)$.

**Exercises 3:**

(1). Consider the "darts" player in Example 7. What is the expected number of throws until he first misses the region at which he aims? What is the probability that he first misses on his sixth throw? What is the probability that his first miss occurs on his third or fourth throw? What is the probability that his first miss comes later than his fifth throw?

(2). Consider the problem of sampling adult Canadians regarding their opinions on the issue of government inefficiency in the handling of taxpayers' money as in Example 3. What is the expected number of adult Canadians sampled until one is found who agrees that governments are inefficient in handling taxpayers' money? What is the expected number of adult Canadians sampled until one is found who does not agree that governments are inefficient in handling taxpayers' money? What is the probability that the first person who agrees that governments are inefficient in handling taxpayers' money is the third person sampled?

(3). Suppose a telephone banking system is busy 20% of the time. That is, the probability that a person finds the service busy when phoning in is 0.20. What is the probability that you finally get through to the banking system on your fourth attempt? What is the mean number of attempts that are needed in order to finally get through to the banking system?

**C. The NEGATIVE BINOMIAL DISTRIBUTION**

The Geometric distribution is the appropriate probability distribution in a sequence of Bernoulli trials when waiting for the first Success to occur. What is the probability distribution in a sequence of Bernoulli trials when waiting for two successes, or three successes, etc.? In general, what is the waiting time distribution to the $k^{th}$ success in a sequence of Bernoulli trials, for $k = 2, 3, 4, \ldots$? The answer is the Negative Binomial distribution.

Suppose that 8% of the automobile engines produced on one assembly line are defective. If engines are selected at random from the assembly line, what is the probability that the second
defective engine is the fourth engine selected? Letting $D$ denote defective and $N$ denote non-defective, the different sequences that lead to finding a second defective as the fourth engine selected are $DNND$ and $NDND$ and $NNDD$. Note the following two properties of each of these sequences: the fourth engine is Defective, and there is one Defective engine amongst the first three selected. The probability of a Defective engine is 0.08 and of a non-defective engine 0.92. Thus the probability of each of the three sequences is $q^2p^2 = (0.92)^2(0.08)^2 = 0.00541696$ and the required probability is $3q^2p^2 = 3(0.92)^2(0.08)^2 = 0.01625088$.

How can the event "second defective as fourth engine selected" occur? The answer is: there must be exactly one defective in the first three selected, and the fourth one must then be another defective. What is the probability of exactly one defective in the first three selected? This is just a Binomial distribution question, say $Y \sim B(3,0.08)$, and

$$P[Y = 1] = \binom{3}{1}(0.08)^1(0.92)^2 = 0.203136.$$ 

Multiplying this answer by $p = 0.08$ to account for the second defective at the fourth draw gives the required answer as above.

This is the essential nature of the Negative Binomial distribution. In waiting for the $k^{th}$ success in a sequence of Bernoulli trials with parameter $p$, exactly $x$ trials will be required if the first $x - 1$ trials result in $k - 1$ successes, and then the $x^{th}$ trial results in the $k^{th}$ success. Note that at least $k$ trials will be required to achieve $k$ successes so that the value space of the random variable is $\{k,k+1,k+2,\cdots\}$.

**Definition:** A random variable $Y$ is said to have a **Negative Binomial** distribution with parameters $k$ and $p$ - that is, $X \sim NB(k,p)$ - if its value space is $V = \{k,k+1,k+2,\cdots\}$ and its probability function is given by

$$p_X(x) \equiv P[X = x] = \left(\begin{array}{c}x-1 \\ k-1 \end{array}\right)p^{k-1}q^{(x-1)-(k-1)} = \left(\begin{array}{c}x-1 \\ k-1 \end{array}\right)p^kq^{x-k}$$

for $x \in V = \{k,k+1,k+2,\cdots\}$.

In waiting for the $k^{th}$ success, it is necessary to first wait for the first success, then for the second success, and so on, and then for the $k^{th}$ success. This results in the mean and variance of a Negative Binomial random variable each being $k$ times the mean and variance of a Geometric random variable with the same probability parameter $p$.

**Theorem 3:**

Suppose $X$ is a Negative Binomial random variable with parameters $k$ and $p$; that is, $X \sim NB(k,p)$. Then

the mean or expected value of $X$ is $\mu = \frac{k}{p}$, 

the variance of $X$ is $\sigma^2 = \frac{kq}{p^2}$, and
Bernoulli Trials and Related Probability Distributions

the standard deviation of $X$ is $\sigma = \sqrt{\frac{kq}{p}}$.

Example 12:
A wildcat oil well drilling company successfully strikes oil on 20% of the wells it drills. The company has a new field to explore and begins drilling wells one after another. What is the probability that the third successful well will be the eighth one drilled? What is the expected number of wells that must be drilled in order to find oil three times?

A Negative Binomial random variable with $k = 3$ and $p = 0.2$ is called for. Let $X \sim NB(3,0.2)$. The probability that the third successful well is the eighth one drilled is

$$p_X(8) = P[X = 8] = \binom{8-1}{3-1} (0.2)^3 (0.8)^{8-3} = \binom{7}{2} (0.2)^3 (0.8)^5 = 21 \times 0.008 \times 0.32768.$$ 

$$= 0.05505024$$

The expected number of wells to drill in order to find a third successful one is 

$$\mu = \frac{k}{p} = \frac{3}{0.2} = 15.0.$$ 

[Note: This expected value is reasonable. The company is successful on one-fifth of its attempts suggesting an average of one success in every five attempts. Thus 15 attempts would be required - on average - to obtain three successes.]

Example 13: Suppose that 52 female children are born for every 48 male children. A couple decides that it wants to have two sons. What is the probability that, if they keep having children until they have their second son, they will end up with five children? Assume the sexes of the children are independent of each other. What is the expected number of children in such a family wanting to have two sons?

Let $Y$ be a random variable counting the number of children in this family. Since the family will stop having children once a second son is born, $Y$ is a Negative Binomial random variable with parameters $k = 2$ and $p = 48/(52+48) = 0.48$. That is, $Y \sim NB(2,0.48)$.

The probability that the family has five children is

$$P[Y = 5] = \binom{5-1}{2-1} (0.48)^2 (0.52)^{5-2} = \binom{4}{1} (0.48)^2 (0.52)^3 = 0.129584333.$$ 

The expected number of children in the family is 

$$\mu = \frac{k}{p} = \frac{2}{0.48} = 4.166667.$$ 

Exercises 4:
(1). A die is rolled until a "5" is observed for the third time. What is the probability that exactly seven rolls are required? What is the probability that at most four rolls are required? What is the expected number of rolls required?
(2). In the NHL's Stanley Cup Playoffs, two teams play each other until one team has won four games. Suppose in one series team $A$ has a probability of 0.55 of winning any given game, and that the outcomes from game to game are independent events. What is the probability that team $A$ wins the series in exactly 6 games? What is the probability that team $A$ wins the series in less than 6 games? What is the probability that team $A$ wins the series?

Final Comments:

(1). Probability distributions are used as models in finding probabilities of events. Each model is built upon certain conditions which should hold in order to use that model. Sometimes a model will be used even when the conditions are not exactly satisfied but are approximately so.

(2). Remember that the Binomial, Geometric and Negative Binomial distributions are all built upon Bernoulli trials. What are the three properties of Bernoulli trials? What kinds of situations lead to Bernoulli trials?

Additional Exercises:

(1). The University of Saskatchewan reported student enrollment of 19,717 students in the Fall of 2003. Of this total, 80% are full-time students while the other 20% are part-time students. Suppose students are randomly selected from this large group.

a). If 12 students are selected at random, what is the probability that more than two of them are full-time students? How many students in this sample would be expected to be full-time?

b). If students are selected one after another, what is the probability that a part-time student will be selected for the first time as the seventh student selected? How many students would one expect to have to select in order to obtain a part-time student?

c). Suppose students are selected until a third part-time student is selected. What is the probability that 12 students will be selected? How many students would one expect to have to select in order to obtain three part-time students.

(2). A major league baseball player has a .300 batting average. This means that he has successfully “hit” on 30% of his times “at-bat” (in the past). Suppose that he enters one game with this batting average 0.300 as his probability of getting a hit the first time he comes to bat. As the game progresses, his probability of getting a hit changes as follows. If he does get a hit during one “at-bat”, his probability of getting a hit the next time up increases by 0.100. If he does not get a hit during one “at-bat”, his probability of getting a hit the next time up decreases by 0.050. Suppose this player comes to bat five times during one game.

a). What is the probability that he gets a hit all five times he is at bat?

b). What is the probability that he gets no hits during this game?

c). What is the probability that he gets exactly one hit during this game?