統計學 (一)

Anderson's Statistics for Business & Economics (14/E)

Chapter 5: Discrete Probability Distributions

上課時間地點: 二 D56, 學思 040103

授課教師: 吳漢銘 (國立政治大學統計學系副教授)

教學網站: http://www.hmwu.idv.tw

5.1 Random Variables

1. A random variable is a ____________ description of the ____________ of an experiment.
2. A random variable can be classified as being either ____________ or ____________ depending on the numerical values it assumes.
3. A random variable, often denoted by X (or some other capital letter), is a ____________ that maps ____________ to ____________ on the real line.

Discrete Random Variables

1. A discrete random variable assumes either a ____________ of values or an ____________ of values such as $0, 1, 2, \cdots$.

TABLE 5.1 Examples of Discrete Random Variables

Random Experiment	Random Variable (x)	Possible Values for the Random Variable
Flip a coin	Face of coin showing	1 if heads; 0 if tails
Roll a die	Number of dots showing on top of die	1, 2, 3, 4, 5, 6
Contact five customers	Number of customers who place an order	0, 1, 2, 3, 4, 5
Operate a health care clinic for one day	Number of patients who arrive	0, 1, 2, 3, . . .
Offer a customer the choice of two products	Product chosen by customer	0 if none; 1 if choose product A; 2 if choose product B

Continuous Random Variables

1. A continuous random variable assumes any ____________ value in an __________ or collection of intervals.

2. Experimental outcomes based on measurement scales such as time, weight, distance, and temperature can be described by _____________ random variables.

TABLE 5.2 Examples of Continuous Random Variables

Random Experiment	Random Variable (x)	Possible Values for the Random Variable
Customer visits a web page	Time customer spends on web page in minutes	$x \geq 0$
Fill a soft drink can (max capacity = 360 milliliters)	Number of milliliters	$0 \leq x \leq 360$
Test a new chemical process	Temperature when the desired reaction takes place (min temperature = 65°C; max temperature = 100°C)	$65 \leq x \leq 100$
Invest $10,000 in the stock market	Value of investment after one year	$x \geq 0$

☺ **EXERCISES 5.1**: 1, 3, 5

5.2 Developing Discrete Probability Distributions

1. The **probability distribution** for a random variable describes how ______________ are distributed over the _________ of the random variable.

2. For a discrete random variable X, a probability function, denoted by _________________, provides the probability for each value of the random variable.

3. The __________, __________, and ______________ methods of assigning probabilities can be used to develop discrete probability distributions.

4. *The classical method*:

   (a) The classical method of assigning probabilities to values of a random variable is applicable when the experimental outcomes generate values of the random variable that are ____________.

   (b) Example The experiment of rolling a die.

      i. Each of th outcomes (the numbers 1, 2, 3, 4, 5, or 6) is equally likely.

      ii. Let ______________________________.

      iii. (Table 5.3) the probability distribution function $f(x)$ of X.

TABLE 5.3 Probability Distribution for Number Obtained on One Roll of a Die

Number Obtained x	Probability of x $f(x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

5. *The subjective method*:

   (a) Each probability is assigned by user's best __________. Different people can be expected to obtain different probability distributions.

6. *The relative frequency method*:

   (a) We treat the (large amounts of) data as if they were the __________ and use the ______________ method to assign probabilities to the experimental outcomes.

   (b) The use of the relative frequency method to develop discrete probability distributions leads to what is called an __________ discrete distribution. (more widely used in practice)

✎ Question .... (p229)

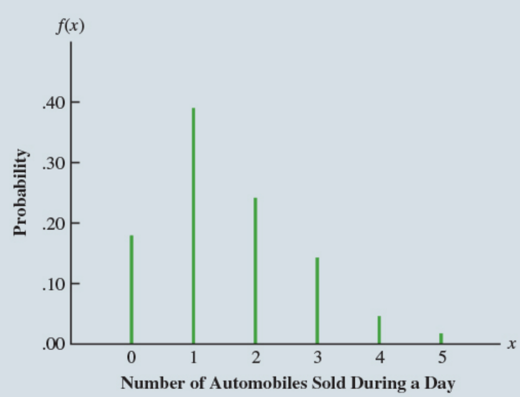
Over the past 300 days, DiCarlo DiCarlo Motors in Saratoga, New York, has experienced 54 (117, 72, 42, 12, 3) days with no (1, 2, 3, 4, 5) automobiles sold. Suppose we consider the experiment of observing a day of operations at DiCarlo Motors and define the random variable of interest as X = the number of automobiles sold during a day. (a) Use the relative frequency method to develop a probability distribution for the number of cars sold per day. (b) Graph the probability distribution such that the values of the random variable X for DiCarlo Motors are shown on the horizontal axis and the probability associated with these values is shown on the vertical axis.

sol:

TABLE 5.4 Probability Distribution for the Number of Automobiles Sold During a Day at Dicarlo Motors

x	*f*(*x*)
0	.18
1	.39
2	.24
3	.14
4	.04
5	.01
Total	1.00

FIGURE 5.1 Graphical Representation of the Probability Distribution for the Number of Automobiles Sold During a Day at Dicarlo Motors



7. A primary advantage of defining a random variable and its probability distribution is that once the probability distribution is known, it is relatively easy to determine the probability of a variety of __________ that may be of interest to a decision maker.

8. *Required Conditions for a Discrete Probability Function*:

   ____________ and ______________

9. **Discrete Uniform Probability Distribution**: ______________ where n is the number of values the random variable may assume.

✎ Question .... (p230)

For the experiment of rolling a fair die, we define the random variable X to be the number of dots on the upward face. Find the probability function for this discrete random variable.

sol:

☺ **EXERCISES 5.2**: 9, 10

5.3 Expected Value and Variance

Expected Value

1. The expected value, or __________, of a random variable X is a measure of the ____________ location for the discrete random variable:

$$E(X) = \underline{\qquad\quad} = \underline{\qquad\qquad\quad}.$$

2. The expected value is a ______________________ of the values of the random variable where the weights are the __________________.

Variance

1. Use variance to summarize the _______________ in the values of a discrete random variable:

$$Var(X) = \underline{\qquad\quad} = \underline{\qquad\qquad\qquad}$$

2. The variance is a weighted average of the _______________________ of a random variable from its mean. The weights are the probabilities.

Question .... (p234)

Let the random variable X be the number of auto mobiles sold during a day in the DiCarlo Motors automobile sales example from Section 5.2, find the expected value and the variance of this discrete random variable.

sol:

TABLE 5.5 Calculation of the Expected Value for the Number of Automobiles Sold During a Day at Dicarlo Motors

x	f(x)	xf(x)
0	.18	0(.18) = .00
1	.39	1(.39) = .39
2	.24	2(.24) = .48
3	.14	3(.14) = .42
4	.04	4(.04) = .16
5	.01	5(.01) = .05
		1.50

$E(x) = \mu = \Sigma x f(x)$

TABLE 5.6 Calculation of the Variance for the Number of Automobiles Sold During a Day at Dicarlo Motors

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	0 − 1.50 = −1.50	2.25	.18	2.25(.18) = .4050
1	1 − 1.50 = −.50	.25	.39	.25(.39) = .0975
2	2 − 1.50 = .50	.25	.24	.25(.24) = .0600
3	3 − 1.50 = 1.50	2.25	.14	2.25(.14) = .3150
4	4 − 1.50 = 2.50	6.25	.04	6.25(.04) = .2500
5	5 − 1.50 = 3.50	12.25	.01	12.25(.01) = .1225
				1.2500

$\sigma^2 = \Sigma(x - \mu)^2 f(x)$

😍 **EXERCISES 5.3**: 16, 21, 23

5.4 Bivariate Distributions, Covariance, and Financial Portfolios

1. A probability distribution involving two random variables is called a ____________ probability distribution.

2. **Bivariate experiment**: Each outcome for a bivariate experiment consists of ______________, one for each random variable.

3. Example Consider the bivariate experiment of rolling a pair of dice. The outcome consists of two values, the number obtained with the first die and the number obtained with the second die.

4. Example Consider the experiment of observing the financial markets for a year and recording the percentage gain for a stock fund and a bond fund. Again, the

experimental outcome provides a value for two random variables, the percent gain in the stock fund and the percent gain in the bond fund.

A Bivariate Empirical Discrete Probability Distribution

1. Example (Section 5.2) DiCarlo Motors automobile dealership in Saratoga, New York.

   (a) (Table 5.7) DiCarlo has another dealership in Geneva, New York. Table 5.7 shows the number of cars sold at each of the dealerships over a 300-day period.

   (b) The numbers in the bottom (total) row are the frequencies we used to develop an empirical probability distribution for daily sales at DiCarlo's Saratoga dealership. The numbers in the rightmost (total) column are the frequencies of daily sales for the Geneva dealership.

   (c) Entries in the body of the table give the number of days the Geneva dealership had a level of sales indicated by the row, when the Saratoga dealership had the level of sales indicated by the column.

   TABLE 5.7 Number of Automobiles Sold at DiCarlo's Saratoga and Geneva Dealerships Over 300 Days

	Saratoga Dealership						
Geneva Dealership	**0**	**1**	**2**	**3**	**4**	**5**	**Total**
0	21	30	24	9	2	0	86
1	21	36	33	18	2	1	111
2	9	42	9	12	3	2	77
3	3	9	6	3	5	0	26
Total	54	117	72	42	12	3	300

   (d) Consider the bivariate experiment of observing a day of operations at DiCarlo Motors and recording the number of cars sold. Define

   $$X = \text{number of cars sold at the Geneva dealership}$$
   $$Y = \text{the number of cars sold at the Saratoga dealership}$$

   (e) (Table 5.8) divide all of the frequencies in Table 5.7 by the number of observations (300) to develop a ____________________ probability distribution for automobile sales at the two DiCarlo dealerships.

TABLE 5.8 Bivariate Empirical Discrete Probability Distribution for Daily Sales at DiCarlo Dealerships in Saratoga and Geneva, New York

	Saratoga Dealership						
Geneva Dealership	0	1	2	3	4	5	Total
0	.0700	.1000	.0800	.0300	.0067	.0000	.2867
1	.0700	.1200	.1100	.0600	.0067	.0033	.3700
2	.0300	.1400	.0300	.0400	.0100	.0067	.2567
3	.0100	.0300	.0200	.0100	.0167	.0000	.0867
Total	.18	.39	.24	.14	.04	.01	1.0000

(f) The probabilities in the lower (right) margin provide the ____________ distribution for the DiCarlo Motors Saratoga (Geneva) dealership.

(g) The probabilities in the body of the table provide the ____________ probability distribution for sales at both dealerships.

2. Bivariate probabilities are often called ____________________.

3. Note that there is one bivariate probability for each experimental outcome.

4. With _______ possible values for X and _______ possible values for Y, there are ____________ experimental outcomes and bivariate probabilities.

Question .... (p240)

Refer to the DiCarlo Motors automobile dealership data, find the probability distribution for total sales at both DiCarlo dealerships and the expected value and variance of total sales.

sol:

TABLE 5.9 Calculation of the Expected Value and Variance for Total Daily Sales at DiCarlo Motors

s	$f(s)$	$sf(s)$	$s - E(s)$	$(s - E(s))^2$	$(s - E(s))^2 f(s)$
0	.0700	.0000	−2.6433	6.9872	.4891
1	.1700	.1700	−1.6433	2.7005	.4591
2	.2300	.4600	−.6433	.4139	.0952
3	.2900	.8700	.3567	.1272	.0369
4	.1267	.5067	1.3567	1.8405	.2331
5	.0667	.3333	2.3567	5.5539	.3703
6	.0233	.1400	3.3567	11.2672	.2629
7	.0233	.1633	4.3567	18.9805	.4429
8	.0000	.0000	5.3567	28.6939	.0000
		$E(s) = 2.6433$			$Var(s) = 2.3895$

5. The covariance between two random variables X and Y is:

$$\sigma_{xy} = Cov(X, Y) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad} \quad \text{or}$$

$$\sigma_{xy} = Cov(X, Y) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}.$$

6. The correlation between two random variables X and Y is the covariance divided by the product of the standard deviations for the two random variables:

$$\rho_{xy} = \frac{\qquad\qquad}{\underline{\qquad\qquad}}.$$

7. *Expected Value of a Linear Combination of Random Variables X and Y*:

$$E(aX + bY) = \underline{\qquad\qquad\qquad\qquad}.$$

8. *Variance of a Linear Combination of Two Random Variables*:

$$Var(aX + bY) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}$$

where $Cov(X, Y) = \sigma_{xy}$ is the covariance of X and Y.

Question .... (p241)

Refer to the DiCarlo Motors automobile dealership data, compute the covariance and the correlation coefficient between daily sales at the two DiCarlo dealerships.

sol:

TABLE 5.10 Calculation of the Expected Value and Variance of Daily Automobile Sales at DiCarlo Motors' Geneva Dealership

x	$f(x)$	$xf(x)$	$x - E(x)$	$[(x - E(x)]^2$	$[x - E(x)]^2 f(x)$
0	.2867	.0000	−1.1435	1.3076	.3749
1	.3700	.3700	−.1435	.0206	.0076
2	.2567	.5134	.8565	.7336	.1883
3	.0867	.2601	1.8565	3.447	.2988
		$E(x) = 1.1435$			$Var(x) = .8696$

Financial Applications

1. To see how what we have learned can be useful in constructing financial portfolios that provide a good balance of ________ and __________.
2. The ____________________ of percent return is often used as a measure of risk associated with an investment.

✎ Question .... (p242)

(Table 5.11) A financial advisor is considering four possible economic scenarios for the coming year and has developed a probability distribution showing the percent return, X, for investing in a largecapstock fund and the percent return, Y, for investing in a long-term government bond fund given each of the scenarios. The bivariate probability distribution for X and Y is shown in Table 5.11.

TABLE 5.11 Probability Distribution of Percent Returns for Investing in a Large-Cap Stock Fund, *x*, and Investing in a Long-Term Government Bond Fund, *y*

Economic Scenario	Probability *f(x, y)*	Large-Cap Stock Fund (*x*)	Long-Term Government Bond Fund (*y*)
Recession	.10	−40	30
Weak Growth	.25	5	5
Stable Growth	.50	15	4
Strong Growth	.15	30	2

Compute the expected percent return for investing in the stock fund, $E(X)$, and the expected percent return for investing in the bond fund, $E(Y)$. Draw the conclusion.

sol:

✎ Question .... (p242)

Refer to the financial advisor example. Compute the standard deviation of the percent returns for the stock and bond fund investments.

sol:

3. We have already seen that the stock fund offers a greater expected return, so if we want to choose between investing in either the ________ fund or the ________ fund it depends on our attitude toward ________________.

4. An ____________ investor might choose the stock fund because of the higher expected return; a ______________ investor might choose the bond fund because of the lower risk. But, there are other options.

5. What about the possibility of investing in a portfolio consisting of both an investment in the stock fund and an investment in the bond fund?

✎ Question .... (p242)

Refer to the financial advisor example. Suppose we would like to consider a portfolio by investing equal amounts in the largecap stock fund and in the longterm government bond fund. Evaluate this portfolio by computing its expected return and risk.

sol:

1.

2.

3.

✎ Question .... (p242)

Refer to the financial advisor example. Compare the three investment alternatives according to their expected returns, variances, and standard deviations: investing solely in the stock fund, investing solely in the bond fund, or creating a portfolio by dividing our investment amount equally between the stock and bond funds. Which of these alternatives would you prefer?

TABLE 5.12 Expected Values, Variances, and Standard Deviations for Three Investment Alternatives

Investment Alternative	Expected Return (%)	Variance of Return	Standard Deviation of Return (%)
100% in Stock Fund	9.25	328.1875	18.1159
100% in Bond Fund	6.55	61.9475	7.8707
Portfolio (50% in stock fund, 50% in bond fund)	7.90	29.865	5.4650

sol:

1. The expected return is highest for investing 100% in the stock fund, but the risk is also highest. The standard deviation is 18.1159%.
2. Investing 100% in the bond fund has a lower expected return, but a significantly smaller risk.
3. Investing 50% in the stock fund and 50% in the bond fund (the portfolio) has an expected return that is halfway between that of the stock fund alone and the bond fund alone. But note that it has less risk than investing 100% in either of the individual funds. It has both a higher return and less risk (smaller standard deviation) than investing solely in the bond fund.
4. Whether you would choose to invest in the stock fund or the portfolio depends on your attitude toward risk.
5. The stock fund has a higher expected return. But the portfolio has significantly less risk and also provides a fairly good return. Many would choose it.
6. It is the negative covariance between the stock and bond funds that has caused the portfolio risk to be so much smaller than the risk of investing solely in either of the individual funds.

☺ **EXERCISES 5.4**: 26, 28, 29

5.5 Binomial Probability Distribution

A Binomial Experiment

1. A binomial experiment exhibits the following four properties.

   (a) The experiment consists of a sequence of ________________.

   (b) Two outcomes are possible on each trial. We refer to one outcome as a __________ and the other outcome as a __________.

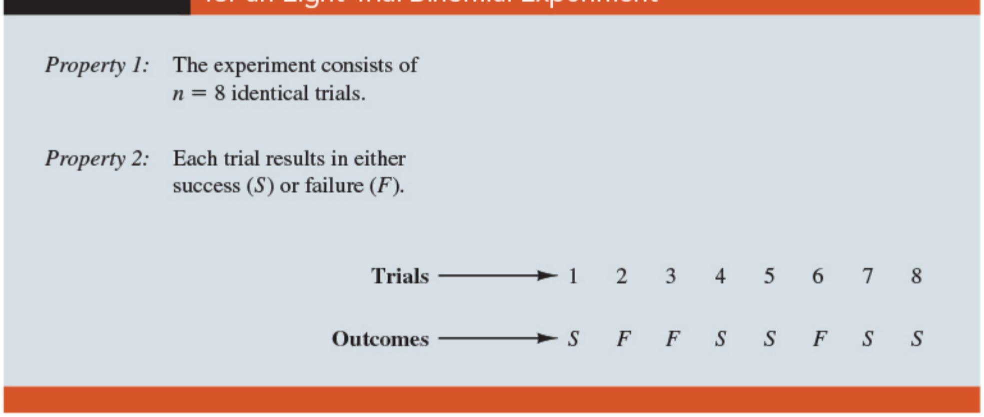
   (c) The probability of a success, denoted by _____, does not change from trial to trial. Consequently, the probability of a failure, denoted by ________, does not change from trial to trial. (stationarity assumption)

   (d) The trials are ______________.

2. If properties (b), (c), and (d) are present, we say the trials are generated by a ____________ process.

3. (Figure 5.2) depicts one possible sequence of successes and failures for a binomial experiment involving eight trials.

FIGURE 5.2 One Possible Sequence of Successes and Failures for an Eight-Trial Binomial Experiment



4. In a binomial experiment, our interest is in the ______________________ occurring in the ___________. If we let X denote the number of successes occurring in the n trials, we see that X can assume the values of _________________.

5. Because the number of values is _________, X is a discrete random variable. The probability distribution associated with this random variable is called the ____________ probability distribution. Denote by _______________.

✎ Question .... (p248)

Consider the experiment of tossing a coin five times and on each toss observing whether the coin lands with a head or a tail on its upward face. Suppose we want to count the number of heads appearing over the five tosses. Does this experiment show the properties of a binomial experiment? What is the random variable of interest?

sol:

Four properties of a binomial experiment are satisfied:

1. The experiment consists of five identical trials; each trial involves the tossing of one coin.
2. Two outcomes are possible for each trial: a head or a tail. We can designate head a success and tail a failure.
3. The probability of a head and the probability of a tail are the same for each trial, with $p = 0.5$ and $1 - p = 0.5$.
4. The trials or tosses are independent because the outcome on any one trial is not affected by what happens on other trials or tosses.

The random variable of interest is X = the number of heads appearing in the five trials. In this case, X can assume the values of 0, 1, 2, 3, 4, or 5.

✎ Question .... (p249)

Consider an insurance salesperson who visits 10 randomly selected families. The outcome associated with each visit is classified as a success if the family purchases an insurance policy and a failure if the family does not. From past experience, the salesperson knows the probability that a randomly selected family will purchase an insurance policy is 0.10. Checking the properties of a binomial experiment.

sol:

Four assumptions are satisfied, this example is a binomial experiment:

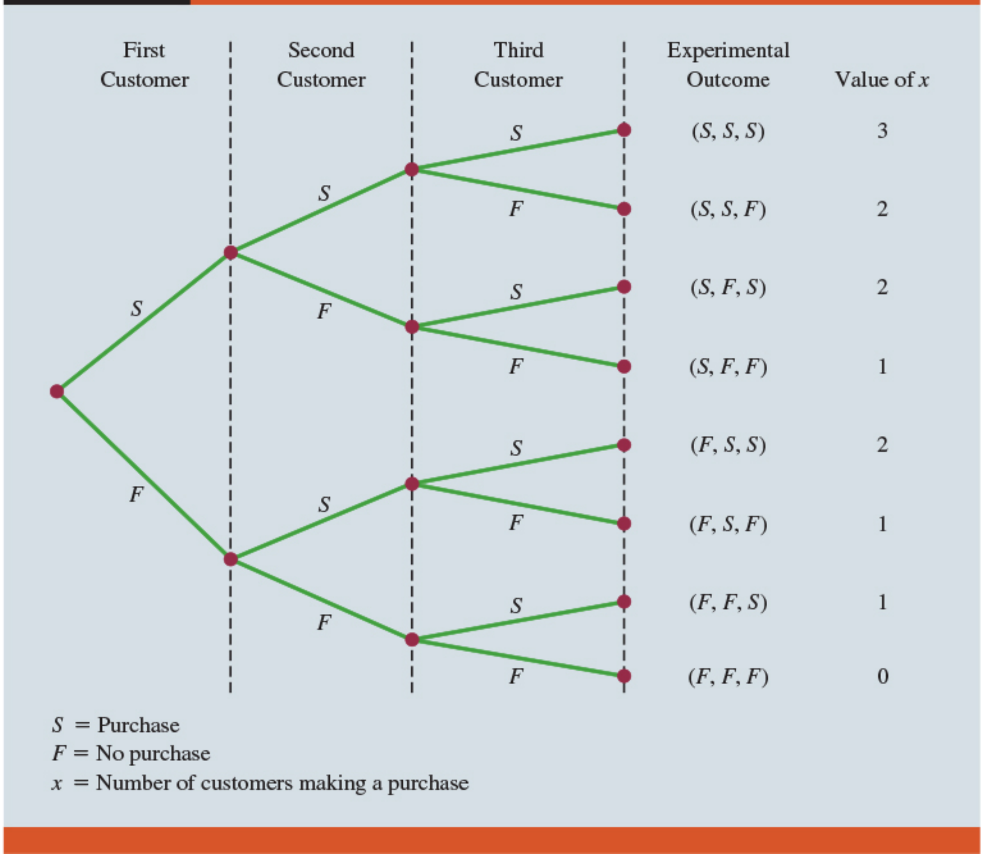
1. The experiment consists of 10 identical trials; each trial involves contacting one family.
2. Two outcomes are possible on each trial: the family purchases a policy (success) or the family does not purchase a policy (failure).
3. The probabilities of a purchase and a nonpurchase are assumed to be the same for each sales call, with $p = 0.10$ and $1-p = 0.90$.
4. The trials are independent because the families are randomly selected.

The random variable of interest is the number of sales obtained in contacting the 10 families. In this case, X can assume the values of $0, 1, \cdots, 10$.

Martin Clothing Store Problem

1. Let us consider the purchase decisions of the next three customers who enter the Martin Clothing Store. On the basis of past experience, the store manager estimates the probability that any one customer will make a purchase is 0.30. What is the probability that two of the next three customers will make a purchase?
2. (Figure 5.3) tree diagram shows that the experiment of observing the three customers each making a purchase decision has eight possible outcomes.

FIGURE 5.3 Tree Diagram for the Martin Clothing Store Problem



3. Using S to denote success (____________) and F to denote failure (_____________), we are interested in experimental outcomes involving two successes in the three trials (purchase decisions).

4. Verify that the experiment involving the sequence of three purchase decisions can be viewed as a binomial experiment:

   (a) The experiment can be described as a sequence of three identical trials, one trial for each of the three customers who will enter the store.

   (b) Two outcomes - the customer makes a purchase (success) or the customer does not make a purchase (failure) - are possible for each trial.

   (c) The probability that the customer will make a purchase (0.30) or will not make a purchase (0.70) is assumed to be the same for all customers.

   (d) The purchase decision of each customer is independent of the decisions of the other customers.

5. The number of experimental outcomes resulting in exactly x successes in n trials can be computed using:

$$\underline{\hspace{4cm}},$$

where $n! = n(n-1)(n-2)\cdots(2)(1)$ and $0! = 1$.

6. Because the trials of a binomial experiment are independent, we can simply ____________ the probabilities associated with each trial outcome to find the probability of a particular sequence of successes and failures.

7. The probability of purchases by the first two customers and no purchase by the third customer, denoted (S, S, F), is given by ____________. With a 0.30 probability of a purchase on any one trial, the probability of a purchase on the first two trials and no purchase on the third is given by __________________________.

8. In any binomial experiment, all sequences of trial outcomes yielding x successes in n trials have the same probability of occurrence. The probability of each sequence of trials yielding x successes in n trials is ________________.

9. **The binomial probability function**

$$f(x) = \underline{\hspace{4cm}},$$

where

$$\begin{aligned} x &= \text{the number of successes}, x = 0, 1, 2, \cdots, n \\ p &= \text{the probability of a success on one trial} \\ n &= \text{the number of trials} \\ f(x) &= \text{the probability of } x \text{ successes in } n \text{ trials} \end{aligned}$$

10. For the binomial probability distribution, X is a discrete random variable with the probability function __________________ applicable for values of $x = 0, 1, 2, \cdots, n$. (________________)

✐ Question .... (p113)

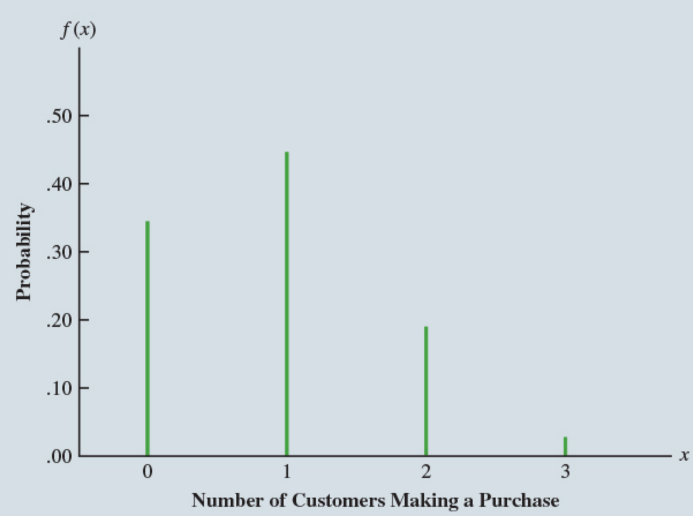
Refer to the Martin Clothing Store example, compute the probability that no customer makes a purchase, exactly one customer makes a purchase, exactly two customers make a purchase, and all three customers make a purchase.

sol: The calculations are summarized in Table 5.13, which gives the probability distribution of the number of customers making a purchase. Figure 5.4 is a graph of this probability distribution.

TABLE 5.13 Probability Distribution for the Number of Customers Making a Purchase

x	f(x)
0	$\frac{3!}{0!3!}(.30)^0(.70)^3 = .343$
1	$\frac{3!}{1!2!}(.30)^1(.70)^2 = .441$
2	$\frac{3!}{2!1!}(.30)^2(.70)^1 = .189$
3	$\frac{3!}{3!0!}(.30)^3(.70)^0 = .027$
	1.000

FIGURE 5.4 Graphical Representation of the Probability Distribution for the Number of Customers Making a Purchase



Using Tables of Binomial Probabilities

1. (Table 5.14) Table 5 of Appendix B provides a table of binomial probabilities. To use this table, we must specify the values of n, p, and x for the binomial experiment of interest.

TABLE 5.14 Selected Values from the Binomial Probability Table
Example: $n = 10$, $x = 3$, $P = .40$; $f(3) = .2150$

						p					
n	***x***	**.05**	**.10**	**.15**	**.20**	**.25**	**.30**	**.35**	**.40**	**.45**	**.50**
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	**.2150**	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010

Expected Value and Variance for the Binomial Distribution

1. Expected Value and Variance for the Binomial Distribution:

$$
\begin{aligned}
E(X) &= \mu = \underline{\hspace{2cm}} \\
Var(X) &= \sigma^2 = \underline{\hspace{4cm}}
\end{aligned}
$$

✎ Question .... (p255)

Refer to the Martin Clothing Store problem with three customers, (a) compute the expected number, the variance and standard deviation of customers who will make a purchase. (b) Suppose that for the next month the Martin Clothing Store forecasts 1000 customers will enter the store. (c) What is the expected number, the variance and standard deviation of customers who will make a purchase?

sol:

☺ **EXERCISES 5.5**: 32, 35, 39, 42

5.6 Poisson Probability Distribution

1. The probability distribution of __________ random variable is called a Poisson distribution. It is a discrete random variable that is often useful in estimating the ______________________ over a specified interval of time or space.

2. Example The random variable of interest might be the number of arrivals at a car wash in one hour, the number of repairs needed in 10 miles of highway, or the number of leaks in 100 miles of pipeline.

3. Two properties must be satisfied, the number of occurrences is a random variable described by the Poisson probability distribution:

   (a) The probability of an occurrence is the ________ for any two intervals of equal length.

   (b) The occurrence or nonoccurrence in any interval is ______________ of the occurrence or nonoccurrence in any other interval.

4. The **Poisson probability function** is defined by

$$f(x) = \frac{\qquad}{\qquad},$$

   where

$$\begin{aligned} x &= \text{the number of occurrences in an interval}, x = 0, 1, 2, \cdots \\ f(x) &= \text{the probability of } X \text{ occurrences in an interval} \\ \mu &= \text{expected value or mean number of occurrences} \\ e &= 2.71828. \end{aligned}$$

5. For the Poisson probability distribution, X is a discrete random variable indicating the number of occurrences in the interval. (Denote ____________ or ______________)

6. Since there is no stated upper limit for the number of occurrences, the probability function $f(x)$ is applicable for values _______________ without limit.

7. In practical applications, X will eventually become large enough so that $f(x)$ is approximately _______ and the probability of any larger values of X becomes ____________.

An Example Involving Time Intervals

1. Suppose that we are interested in the number of patients who arrive at the emergency room of a large hospital during a 15-minute period on weekday mornings. If we can assume that the probability of a patient arriving is the same for any two time periods of equal length and that the arrival or nonarrival of a patient in any time period is independent of the arrival or nonarrival in any other time period, the Poisson probability function is applicable.

2. Suppose these assumptions are satisfied and an analysis of historical data shows that the average number of patients arriving in a 15-minute period of time is 10; in this case, the following probability function applies:

   $f(x) =$ ____________

   The random variable here is $X =$ ______________________________ in any 15-minute period.

3. If management wanted to know the probability of exactly five arrivals in 15 minutes, we would set $X = 5$ and thus obtain

   Probability of exactly 5 arrivals in 15-minutes $= f(5) =$ ____________________

4. (Table 5.15) Table 7 of Appendix B provides probabilities for specific values of x and μ.

5. In the preceding example, the mean of the Poisson distribution is $\mu = 10$ arrivals per 15-minute period.

6. A property of the Poisson distribution is that the mean of the distribution and the variance of the distribution are __________. Thus, the variance for the number of arrivals during 15-minute periods is $\sigma^2 = 10$. The standard deviation is $\sigma = \sqrt{10} = 3.16$.

7. When computing a Poisson probability for a different time interval, we must first ____________ the ____________________ to the time period of interest and then compute the probability.

Question .... (p260)

Suppose the number of patients who arrive at the emergency room of a large hospital during a 15minute period on weekday morning is the Poisson distribution with the average number of patients arriving in a 15-minute period of time is 10, compute the probability of one arrival in a 3-minute period.

sol:

An Example Involving Length or Distance Intervals

✎ Question .... (p260)

Suppose we are concerned with the occurrence of major defects in a highway one month after resurfacing. We will assume that the probability of a defect is the same for any two highway intervals of equal length and that the occurrence or nonoccurrence of a defect in any one interval is independent of the occurrence or nonoccurrence of a defect in any other interval. Hence, the Poisson distribution can be applied. Suppose we learn that major defects one month after resurfacing occur at the average rate of two per mile. Find the probability of no major defects in a particular three mile section of the highway.

sol:

☺ **EXERCISES 5.6**: 45, 49, 50

5.7 Hypergeometric Probability Distribution

1. The hypergeometric probability distribution is closely related to the ____________ distribution.

2. The two probability distributions differ in two key ways. With the hypergeometric distribution, the trials are not _______________; and the probability of success ___________ from trial to trial.

3. Notation for the hypergeometric distribution, ______ denotes the number of elements in the population of size ______ labeled ___________, and _________ denotes the number of elements in the population labeled __________.

4. The hypergeometric probability function is used to compute the probability that in a random selection of n elements, selected ______________________, we obtain x elements labeled success and $n-x$ elements labeled failure.

5. For this outcome to occur, we must obtain x successes from the r successes in the population and $n-x$ failures from the $N-r$ failures.

6. **Hypergeometric Probability Function** (Denote by ____________________):

$$f(x) = \frac{\qquad}{\qquad}$$

where:

$$\begin{aligned} x &= \text{number of successes}, x = 0, 1, 2, \cdots, n \\ n &= \text{number of trials} \\ f(x) &= \text{probability of } x \text{ successes in } n \text{ trials} \\ N &= \text{number of elements in the population} \\ r &= \text{number of elements in the population labeled success}, x \le r \end{aligned}$$

$\binom{N}{n}$: the number of ways n elements can be selected from a population of size N

$\binom{r}{x}$: the number of ways that x successes can be selected from a total of r successes in the

$\binom{N-r}{n-x}$: the number of ways that $n-x$ failures can be selected from a total of $N-r$ failures in

7. The mean and variance of a hypergeometric distribution are

$$
\begin{aligned}
E(X) &= \mu = \underline{\hspace{4cm}} \\
Var(X) &= \sigma^2 = \\
& \qquad \underline{\hspace{12cm}}
\end{aligned}
$$

✎ Question .... (p263)

Electric fuses produced by Ontario Electric are packaged in boxes of 12 units each. Suppose an inspector randomly selects three of the 12 fuses in a box for testing. (a) If the box contains exactly five defective fuses, what is the probability that the inspector will find exactly one of the three fuses defective? (b) What is the probability of finding at least one defective fuse. (c) Find the mean and variance for the number of defective fuses.

sol:

☺ **EXERCISES 5.7**: 55, 57, 58

☺ See Also: SUMMARY, GLOSSARY, KEY FORMULAS

☺ **SUPPLEMENTARY EXERCISES**: 59, 62, 66, 73