STUDY SET 1

Discrete Probability Distributions

1. Consider the following probability distribution function. Compute the mean and standard deviation of X.

x	0	1	2	3	4	5	6	7
$P(x)$	0.05	0.16	0.19	0.24	0.18	0.11	0.05	0.02

ANSWER: $\mu_X = \sum_x xP(x) =$ 2.97 and $\sigma_x = 1.6337$

2. Suppose you know that the number of complaints coming into a phone center averages 3 every ten minutes. Assume that the number of calls follows the Poisson distribution. What is the probability that there are exactly three calls during the next ten minutes?

ANSWER: 0.224

3. The following table lists the relative frequency distribution of the number of calls coming into a call center each hour. If each call takes five minutes to process, what is the mean and standard deviation of the number of minutes the operators are answering questions each hour?

x	2	3	4	5	6	7
$P(x)$	0.07	0.11	0.23	0.31	0.18	0.1

ANSWER: $\mu_Y = 5\mu_x = 23.6$

$\sigma_X = 1.3422$, and $\sigma_Y = 5\sigma_X$ = 6.7112

4. In a box of 20 chocolates, there are four chocolates with coconut filling. What is the probability of choosing four chocolates, one or more of which have coconut filling?

ANSWER: 0.624

<u>THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
Consider the following probability distribution function.

x	0	1	2	3	4	5	6
$P(x)$	0.07	0.19	0.23	0.17	0.16	0.14	0.04

5. What is $P(X > 3)$?

ANSWER: 0.34

6. What is $P(2 < X < 5)$?

ANSWER: 0.33

7. What is $P(X \geq 2)$?

ANSWER: 0.74

8. What is $P(X < 6)$?

ANSWER: 0.96

9. An auditor reviewing the invoices of a small company finds that there are errors in 1.5% of them. If the auditor looks at 500 invoices, what is the probability that he finds more than 3 invoices with errors? Use the Poisson approximation to the binomial distribution.

ANSWER: 0.9408

10. The following table displays the joint probability distribution of X and Y. What is the covariance between the X and Y?

		X = 1	2	3
Y	1	0.10	0.08	0.06
	2	0.16	0.1	0.11
	3	0.02	0.16	0.21

ANSWER: 0.215

11. In basketball, one of the possible penalties after a foul is called a free throw. The fouled player gets one shot at the basket from the foul line. If he misses this shot, he is awarded no points and the penalty is over. If he makes the shot, then he gets one point and gets to take another shot. If he makes the second shot he gets an additional point, and the ball is turned over to the other team. If he misses the second shot, the penalty is over. Suppose that the probability of making a shot is 80%, and the likelihood of making one shot is independent of the outcome of any other shot. What is the probability distribution function for the points from the penalty? On average, how many points would you expect to get from the penalty?

ANSWER:

x	0	1	2
$P(x)$	0.20	0.16	0.64
$x.P(x)$	0	0.16	1.28

1.44 points

12. Consider the following game. You choose a number from 1 to 6 and pay \$1 to roll three dice. If your number is rolled on any of the dice, you get your dollar back and one dollar for each of the times your number came up. For instance, if your number comes up on two of the three dice, than you get a total of three dollars back. On average, how much would you expect to win at this game?

ANSWER:
On average, you would expect to lose 8 cents at this game.

13. Thirty percent of all households have a DVD player. Suppose you select 20 houses at random. On average, how many houses of the twenty would you expect to have a DVD? What is the standard deviation of the number of houses with a DVD?

ANSWER:
$\mu_X = nP = 6$,
$\sigma_X = 2.0494$.

14. In a class of 18 students, there are 7 males and 11 females. A sub-committee of five students will be formed from this class. What is the probability that the sub-committee contains more males than females?

ANSWER: 0.272

<u>THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
In a recent survey of 300 teenagers, 62% of the teenagers indicated that they had been to a movie within the past month. 75% of those teenagers who had seen a movie also had gone out to dinner in the past month, while only 64% of the teenagers who had not seen a movie had been out to dinner in the past month. Define the random variables as follows:
$X = 1$ if teenager had been to movie; $X = 0$ otherwise
$Y = 1$ if teenager had been out to dinner; $Y = 0$ otherwise

15. Find the joint probability function of X and Y.

ANSWER:

	$X = 0$	$X = 1$	**Total**
$Y = 0$	0.1368	0.155	0.2918
$Y = 1$	0.2432	0.465	0.7082
Total	0.38	0.62	

16. Find the conditional probability function of X, given $Y = 1$.

ANSWER: 0.3434

17. Find and interpret the covariance between X and Y.

ANSWER: 0.0259

18. A cooler contains 14 drinks: 8 soft drinks and 6 beers. You select 3 drinks from the cooler. Let the random variable *X* be the number of soft drinks you get out of 3. Develop the probability distribution function of *X*.

ANSWER:

x	0	1	2	3
$P(x)$	0.0550	0.3297	0.4615	0.1538

19. The number of accidents on a US -131 highway average 4.4 per year. Assuming that the number of accidents follows a Poisson distribution, what is the probability that there are more than three accidents next year on US -131?

ANSWER: 0.6406

20. There are 20 professors in the School of Business Administration. 15 of them have received good evaluations from students, while 5 received poor evaluations. You will take four courses in the School of Business Administration next semester. What is the probability that the majority of your professors next semester have received good evaluations?

ANSWER: 0.751.

THE NEXT ELEVEN QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
The following table displays the joint probability distribution of two discrete random variables *X* and *Y*.

		X		
		1	2	3
	0	0.10	0.12	0.06
Y	1	0.05	0.10	0.11
	2	0.02	0.16	0.28

21. Determine the marginal probability distribution for *X*.

ANSWER:

x	1	2	3
$P(x)$	0.17	0.38	0.45

22. Compute the expected value for *X*.

ANSWER: 2.28

23. Compute the standard deviation for *X*.

ANSWER: 0.7359

24. Determine the marginal probability distribution for *Y*.

ANSWER:

y	0	1	2
P(*y*)	0.28	0.26	0.46

25. Compute the expected value for *Y*.

ANSWER: 1.18

26. Compute the standard deviation for *Y*.

ANSWER: 0.8412

27. Compute the covariance between *X* and *Y*.

ANSWER: 0.2496

28. Compute the correlation between *X* and *Y*.

ANSWER: 0.4032

29. Compute the mean for the linear function $W = 2X + Y$.

ANSWER: 5.74

30. Compute the variance for the linear function $W = 2X + Y$.

ANSWER: 3.8724

31. Are *X* and *Y* statistically independent? Explain.

ANSWER: The variables *X* and *Y* are not statistically independent

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:

The finishing process on new furniture leaves slight blemishes. The following table displays a manager's probability assessment of the number of blemishes in the finish of new furniture. For each piece of furniture coming off the line it takes 15 minutes to pack for shipping. In addition, fixing the blemishes take another 2 minutes for each blemish.

Number of Blemishes	0	1	2	3	4	5
Probability	0.34	0.25	0.19	0.11	0.07	0.04

32. On average, how long should it take between the time a piece of furniture comes off the line and is ready to ship?

ANSWER: 17.88 minutes

33. What is the variance of the time it takes between the time a piece of furniture comes off the line and is ready to ship?

ANSWER: $\sigma_X^2 = \sum x^2 P(x) - \mu_X^2 = 4.12 - (1.44)^2 = 2.0464$

$\sigma_Y^2 = (2)^2 \sigma_X^2 = 4(2.0464) = 8.1856$

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:

A bank grants mortgages to 87% of all applicants. After the applicant gets approval, the bank sends an appraiser to evaluate the value of the property. The bank pays an appraiser a salary of $2000 a month plus $200 for each appraisal. Assume the bank gets 10 loan applications next month.

34. What is the variance of the amount of money the bank will have to pay its appraiser?

ANSWER: 45,240

35. What is the expected value of the amount of money the bank will have to pay its appraiser?

ANSWER: 3,740

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:

As a purchasing agent, you are responsible for selecting sources of supply for manufactured components to use in your firm's production process. The salesman for a certain supplier has indicated that they can supply an electronic sub-assembly that has a defect rate of 1.1%—well below your current supplier's defect rate. You accept 100 sub-assemblies for evaluation, and find that there were four defects.

36. Using the Poisson approximation to the binomial, how likely is it to get four or less defects out of 100?

ANSWER: 0.9946

37. Using the Poisson approximation to the binomial, how likely is it to get exactly four defects out of 100?

ANSWER: 0.0203

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:

It has been reported that 1.7% of the work force will retire this year. Consider a random sample of 200 workers.

38. What is the probability that more than three of them will retire this year? Use the Poisson approximation to the binomial.

ANSWER: 0.4416

39. What is an estimate of the standard deviation of the number of people who will retire this year? Use the Poisson approximation to the binomial.

ANSWER: 1.8439

THE NEXT FIVE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
The manager of a movie rental store was interested in examining the relationship between the weekly take-home pay for a family and the amount that family spends weekly on recreational activities. The following output was generated using Minitab:

Covariances

	Take-home pay	Amount spent
Take-home pay	4413.84	
Amount spent	2419.64	1364.29

Let X = weekly take-home pay, and Y = amount spent weekly on recreational activities

40. Identify the covariance between X and Y.

ANSWER:
$Cov(X, Y) = 2{,}419.64$

41. Identify the variance of weekly take-home pay.

ANSWER:
$s_x^2 = 4{,}413.84$

42. Identify the variance of amount spent weekly on recreational activities.

ANSWER:
$s_y^2 = 1{,}364.29$

43. Calculate the correlation between weekly take-home pay and amount spent weekly on recreational activities.

ANSWER:
$$r = \frac{Cov(X,Y)}{s_x s_y} = \frac{2419.64}{\sqrt{(4413.84)(1364.29)}} = 0.986$$

44. Interpret the correlation coefficient found in the previous question.

ANSWER:

There is a strong positive linear relationship between weekly take-home pay and amount spent weekly on recreational activities.

45. Develop a realistic example of a pair of random variables for which you would expect to find negative covariance.

ANSWER
Price of cars and number of cars sold

46. Develop a realistic example of a pair of random variables for which you would expect to find zero covariance.

ANSWER
Dow Jones stock market average and rainfall in Egypt

<u>THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
Let X be a random variable with the following distribution:

x	-10	-5	0	5	10
$P(x)$	0.1	0.3	0.1	0.3	0.2

47. Find the expected value of X.

ANSWER:
$E(X) = \mu_X = \sum x \cdot P(x) = 1.0$

48. Find the standard deviation of X.

ANSWER:
$\sigma_X = \sqrt{\sum (x-\mu)^2 \cdot P(x)} = 6.633$

49. What is the probability that X is farther than one standard deviation from the mean?

ANSWER: 0.3

<u>THE NEXT FIVE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
From past experience, it is known 90% of one-year-old children can distinguish their mother's voice from the voice of a similar sounding female. A random sample of 20 one-year-olds is given this voice recognition test.

50. Find the probability at least 3 children do not recognize their mother's voice.

ANSWER:
$P(X \geq 3) = 0.323$ (Here $P = 0.10$)

51. Find the probability all 20 children recognize their mother's voice.

ANSWER:
$P(X = 20) = 0.122$ (Here $P = 0.90$)

52. Let the random variable X denote the number of children who do not recognize their mother's voice. Find the mean of X.

ANSWER:
$\mu = nP = 2$

53. Let the random variable X denote the number of children who do not recognize their mother's voice. Find the variance of X.

ANSWER:
$\sigma^2 = nP(1-P) = 1.8$

54. Find the probability that at most 4 children do not recognize their mother's voice?

ANSWER:
$P(X \leq 4) = 0.957$ (Here $P = 0.10$)

55. A shipment of six parrots from Brazil includes two parrots with a potentially fatal disease. As usual, the U. S. Customs Office at the shipment's point of entry randomly samples two parrots and tests them for disease. Let the random variable X be the number of healthy parrots in the sample. Find the probability distribution of X.

ANSWER:
X has a hypergeometric distribution with 2 diseased parrots and 4 healthy parrots, for a total of six parrots. Then, the probability distribution of X is given by:

$$p(x) = \frac{C_x^2 C_{2-x}^4}{C_2^6} \quad \text{for } x = 0, 1, 2.$$

56. A package of six light bulbs contains 2 defective bulbs. If three bulbs are selected for use, find the probability none are defective.

ANSWER:
The random variable X = number of defective bulbs has a hypergeometric distribution. Then,

$$P(X = 0) = \frac{C_3^4 C_0^2}{C_3^6} = 1 / 5 = 0.20$$

57. Three yellow and two blue pencils are in a drawer. If we randomly select two pencils from the drawer, find the probability distribution of X, the number of yellow pencils selected.

ANSWER:

x	$P(x)$
0	0.10
1	0.60
2	0.30

58. From a group of 10 bank officers, 3 are selected at random to be relocated and supervise new branch offices. If two of the 10 officers are women and 8 are men, what is the probability exactly one of the officers to be relocated will be a woman?

ANSWER: 0.4667

THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
The quality of computer disks is measured by sending the disks through a certifier which counts the number of missing pulses. A certain brand of computer disks averages 0.1 missing pulse per disk. Let the random variable X denote the number of missing pulses.

59. What is the distribution of X?

ANSWER:
X has a Poisson distribution with mean 0.1.

60. Find the probability the next inspected disk will have no missing pulse.

ANSWER:
$P(X = 0) = 0.905$

61. Find the probability the next disk inspected will have more than one missing pulse.

ANSWER:
$P(X > 1) = 0.005$

62. Find the probability neither of the next two disks inspected will contain any missing pulse.

ANSWER: 0.819

THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
It is known that 70% of the customers in a sporting goods store purchase a pair of running shoes. A random sample of 25 customers is selected. Assume that customers' purchases are made independently, and let X represent the number of customers who purchase running shoes. (Hint: Solve using Excel.)

63. What is the probability that exactly 18 customers purchase running shoes?

ANSWER: 0.171

64. What is the probability that no more than 19 customers purchase running shoes?

ANSWER: $P(X \leq 19) = 0.807$

65. What is the probability that at least 17 customers purchase running shoes?

ANSWER: 0.677

66. What is the probability that between 17 and 21 customers, inclusively, purchase running shoes?

ANSWER: 0.644

THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
The number of people arriving at a bicycle repair shop follows a Poisson distribution with an average of 5 arrivals per hour. Let X represent the number of people arriving per hour.

67. What is the probability that seven people arrive at the bike repair shop in a one hour period of time?

ANSWER: 0.105

68. What is the probability that at most seven people arrive at the bike repair shop in a one hour period of time?

ANSWER: $P(X \leq 7) = 0.867$

69. What is the probability that more than seven people arrive at the bike repair shop in a one hour period of time?

ANSWER: 0.133

70. What is the probability that between 4 and 9 people, inclusively, arrive at the bike repair shop in a one hour period of time?

ANSWER: 0.703

THE NEXT EIGHT QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A small grocery store has two checkout lines available to its customers: a regular checkout line and an express checkout line. Customers with 5 or fewer items are expected to use the express line. Let X and Y be the number of customers in the regular checkout line and the express checkout line, respectively. Note that these numbers include the customers being served, if any. The joint probability distribution is given in the table below.

	$Y = 0$	$Y = 1$	$Y = 2$	$Y \geq 3$
$X = 0$	0.06	0.04	0.03	0.15
$X = 1$	0.09	0.06	0.03	0.04
$X = 2$	0.08	0.05	0.01	0.12
$X \geq 3$	0.07	0.05	0.03	0.09

71. Find the marginal distribution of X. What does this distribution tell you?

ANSWER: 0.24

This distribution indicates the likelihood of observing a particular number of customers in the regular checkout line.

72. Find the marginal distribution of *Y*. What does this distribution tell you?

ANSWER: 0.40
This distribution indicates the likelihood of observing a particular number of customers in the express checkout line.

73. Calculate the conditional distribution of *X* given *Y*.

ANSWER:
The conditional distribution of *X* given *Y* is:

	$Y = 0$	$Y = 1$	$Y = 2$	$Y \geq 3$
$X = 0$	0.200	0.200	0.300	0.375
$X = 1$	0.300	0.300	0.300	0.100
$X = 2$	0.267	0.250	0.100	0.300
$X \geq 3$	0.233	0.250	0.300	0.225
	1.00	1.00	1.00	1.00

74. What is the practical benefit of knowing the conditional distribution in the previous question?

ANSWER:
If we find that the probability that customers are waiting in the regular line when the express line is empty is relatively large, we might permit some customers in the regular line to switch to the express line when it is empty. Conversely, if we learn that the probability that no customers are waiting in the regular line when the express line is busy is relatively large, we might then encourage express line customers to switch to the idle regular line. The idea here is to reduce the average waiting time of the customers.

75. What is the probability that no one is waiting or being served in the regular checkout line?

ANSWER: 0.28

76. What is the probability that no one is waiting or being served in the express checkout line?

ANSWER: 0.30

77. What is the probability that no more than two customers are waiting in both lines combined?

ANSWER: 0.36

78. On average, how many customers would you expect to see in each of these two lines at the grocery store?

ANSWER:
Expected number of customers in regular line = $E(X) = 1.60$
Expected number of customers in express line = $E(Y) = 1.60$

<u>THE NEXT TEN QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
The joint probability distribution of variables X and Y is shown in the table below, where X is the number of tennis racquets and Y is the number of golf clubs sold daily in a small sports store.

Y/X	1	2	3
1	0.30	0.18	0.12
2	0.15	0.09	0.06
3	0.05	0.03	0.02

79. Calculate $E(XY)$.

ANSWER:
$E(XY) = 2.55$

80. Determine the marginal probability distributions of X and Y.

ANSWER:

x	1	2	3
$P(x)$	0.5	0.3	0.2

y	1	2	3
$P(y)$	0.6	0.3	0.1

81. Are X and Y independent? Explain.

ANSWER:
Clearly X and Y are independent since $P(x,y) = P(x).P(y)$ for all pairs (x,y).
Difficulty: 2 Moderate

82. Calculate the conditional probability $P(Y = 2|X = 1)$.

ANSWER:
$P(Y = 2| X = 1) = P(X = 1 \text{ and } Y = 2) / P(X = 1) = 0.15 / 0.50 = 0.30$

83. Calculate the expected values of X and Y.

ANSWER:
$E(X) = 1.7$ and $E(Y) = 1.5$

84. Calculate the variances of X and Y.

ANSWER:
Var(X) = 0.61 and Var(Y) = 0.45

85. Calculate Cov(X,Y). Did you expect this answer? Why?

ANSWER:
Cov(X,Y) = E(XY) – E(X).E(Y) = 2.55 – (1.70)(1.50) = 0.0.
Yes, since X and Y are independent.

86. Find the probability distribution of the random variable $X + Y$.

ANSWER:

$x + y$	2	3	4	5	6
$P(x + y)$	0.30	0.33	0.26	0.09	0.02

87. Calculate $E(X + Y)$ and Var($X + Y$) directly by using the probability distribution of $X + Y$ in the previous question.

ANSWER:
$E(X + Y)$ = 3.2, and Var($X + Y$) = 1.06

88. Show that Var($X + Y$) = Var(X) + Var(Y). Did you expect this result? Why?

ANSWER:
Var(X) + Var(Y) = 0.61 + 0.45 = 1.06 = Var($X + Y$). Yes, since X and Y are independent random variables.

89. The binomial distribution is widely used in business applications. Why do you think this is the case?

ANSWER:
The conditions necessary for applying the binomial distribution are not as restrictive as they may seem. The situation of repeated independent trials with only two possible outcomes occurs quite frequently: sales calls achieve the sale or not, investments are profitable or not, productions achieve quotas or not, etc. Further, a number of situations, which might have a range of outcomes can be classified into two categories-made into categorical variables-and, can thus also be represented by the binomial distribution

THE NEXT FIVE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
An investor puts $15,000 into each of four stocks, labeled *A*, *B*, *C*, and *D*. The table shown below contains the means and standard deviations of the annual returns of these four stocks.

Stock	*Mean Annual Return*	*Standard Deviation of Annual Return*
A	0.15	0.05
B	0.18	0.07
C	0.14	0.03
D	0.17	0.06

90. Assuming that the returns of these four stocks are independent of each other, find the mean, and standard deviation of the total amount that this investor earns in one year from these four investments.

ANSWER: $9,600 and $1,636.307

91. Now, assume that the returns of the four stocks are no longer independent of one another. Specifically, the correlations between all pairs of stock returns are given in the table below.

Correlations	Stock A	Stock B	Stock C	Stock D
Stock A	1.00	0.50	0.80	-0.55
Stock B	0.50	1.00	0.60	-0.30
Stock C	0.80	0.60	1.00	-0.75
Stock D	-0.55	-0.30	-0.75	1.00

Find the mean and standard deviation of the total amount that this investor earns in one year from these four investments.

ANSWER:

Summary measures of portfolio

Mean	$9,600.00
Variance	2655000
Stdev	$1,629.42

92. Compare the results of the previous two questions. Explain the differences in your answers.

ANSWER:
The mean total return in the previous two questions are the same, as the mean annual rates have not changed. The standard deviation of the total return in the second question is less than the value computed in the first question. This difference is explained by the fact that the stock returns are no longer independent; in fact, some pairs of stocks have negative correlations. Investing in stocks with negative correlations tends to reduce the overall risk associated with the portfolio.

93. Continue to assume that the returns of the four stocks are no longer independent of one another, and the correlations between all pairs of stock returns are as given in the second question. Now, suppose that this investor decides to place $20,000 each in stocks *B* and *D*, and $10,000 each in stocks *A* and *C*. Find the mean and standard deviation of the total amount that this investor earns in one year from these four investments.

ANSWER:

Summary measures of portfolio

Mean	$9,900.00
Variance	2976000
StDev	$1,725.11

94. How do the mean and standard deviation of the total amount that this investor earns in one year change as the $60,000 in cash available is reallocated for investment? Provide an intuitive explanation for the changes you observed here.

ANSWER:
The mean total return in the fourth question is now higher than that in the second question and the standard deviation in the fourth question is also higher than that in the second question. Higher proportions of the available cash are now being invested in the stocks with higher mean annual rates of return and higher standard deviations of annual returns; that is, stocks B and D.