

Freeport AP Statistics

Chapter 6: Random Variables

6.1 Discrete and Continuous Random Variables

Chapter 6 Big Ideas:

- Random variables are divided into two categories: discrete (typically resulting from counting) and continuous (typically resulting from measuring).
- We are interested in the shape of probability distributions – much like ch. 1.
- Probabilities are determined by finding the area under the density curve and above the value of interest.
- Linear transformations have the same effect on probability distributions of a random variable as they have on a distribution of data.
- Need to distinguish linear transformations from combinations of random variables.
- Two common types of discrete random variables: binomial random variables and geometric random variables.

6.1 OBJECTIVE(S):

- *Students will learn how to use a probability distribution to answer questions about possible values of a random variable.*
- *Students will learn how to calculate the mean and standard deviation of a discrete random variable.*
- *Students will learn how to interpret the mean and standard deviation of a random variable.*

Random Variable –

Probability Distribution -

Discrete Random Variable -

The probabilities of a probability distribution must satisfy two requirements:

a.

b.

Mean (expected value) of a discrete random variable

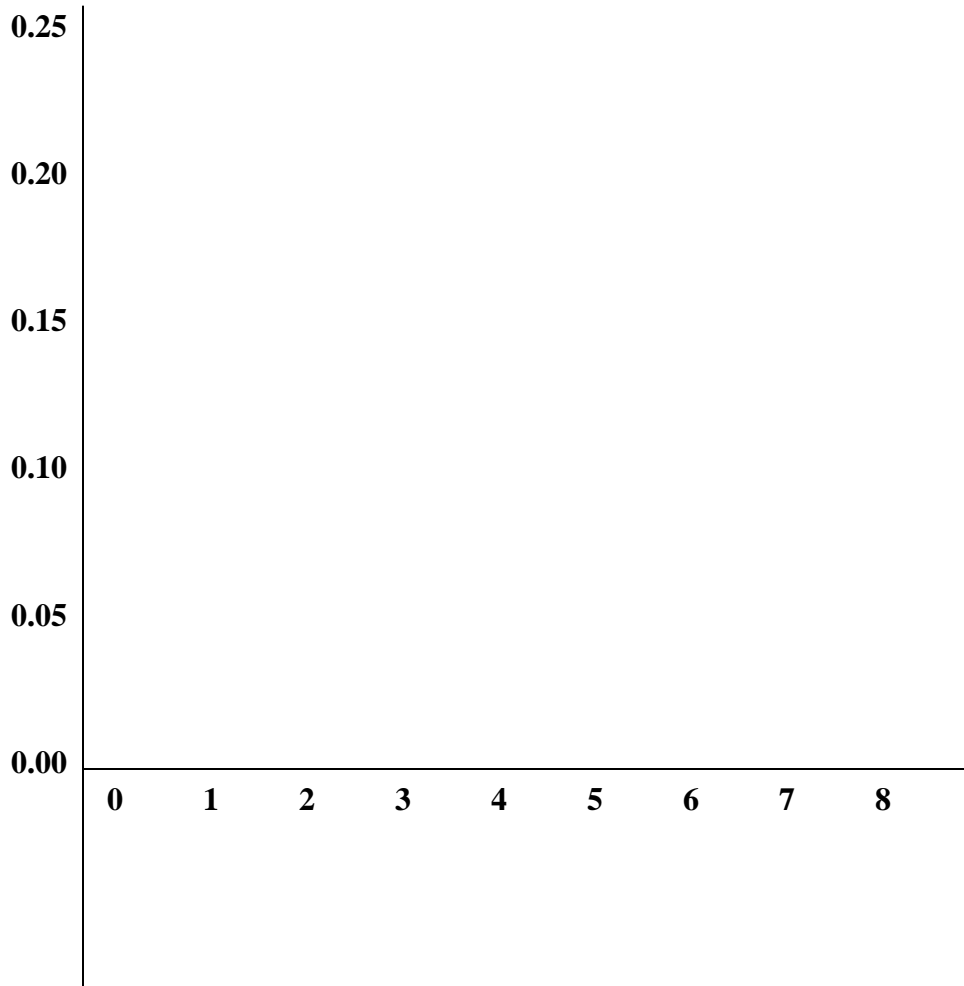
$$\mu_x = E(X) =$$

=

- In 2010, there were 1319 games played in the National Hockey League’s regular season. Imagine selecting one of these games at random and then randomly selecting one of the two teams that played in the game. Define the random variable X = number of goals scored by a randomly selected team in a randomly selected game. The table below gives the probability distribution of X :

Goals:	0	1	2	3	4	5	6	7	8	9
Probability:	0.061	0.154	0.228	0.229	0.173	0.094	0.041	0.015	0.004	0.001

- Show that the probability distribution for X is legitimate.
- Make a histogram of the probability distribution. Describe what you see.



- c. What is the probability that the number of goals scored by a randomly selected team in a randomly selected game is at least 6?
- d. Compute the mean of the random variable X and interpret this value in context.

NOTE: If the mean of a random variable has a non-integer value, but you report it as an integer (i.e. due to rounding), your answer will be marked as incorrect.

- 2. Suppose you roll a pair of fair, six-sided dice. Let T = the sum of the spots showing on the up-faces.
 - a. Find the probability distribution of T .

- b. Find $P(T \geq 5)$ and interpret the result.

- 3. Choose a person aged 19 to 25 years at random and ask, “In the past seven days, how many times did you go to an exercise or fitness center or work out?” Call the response Y for short. Based on a large sample survey, here is a probability model for the answer you will get:

Days:	0	1	2	3	4	5	6	7
Probability:	0.68	0.05	0.07	0.08	0.05	0.04	0.01	0.02

- a. Show that this is a legitimate probability distribution.

- b. Describe the event $Y < 7$ in words. What is $P(Y < 7)$?
- c. Express the event “worked out at least once.” in terms of Y . What is the probability of this event?
- d. Consider the events $A =$ works out at least once and $B =$ works out less than 5 times per week.
- i. What outcomes make up the event A ? What is the $P(A)$?
 - ii. What outcomes make up the event B ? What is the $P(B)$?
 - iii. What outcomes make up the event “ A and B ”? What is $P(A \text{ and } B)$? Why is this probability not equal to $P(A) \cdot P(B)$?

CHAPTER 6

4. Suppose a homeowner spends \$300 for a home insurance policy that will payout \$200,000 if the home is destroyed by fire. Let Y = the profit made by the company on a single policy. From previous data, the probability that a home in this area will be destroyed by fire is 0.0002.
- a. Make a table that shows the probability distribution of Y .

 - b. Compute the expected value of Y . Explain what this result means for the insurance company.

Variance and Standard Deviation of a Discrete Random Variable

$$\text{Var}(X) = \sigma_x^2 =$$

=

The standard deviation of X , _____, is

5. The mean number of goals for a randomly selected team in a randomly selected game to be $\mu_x = 2.851$. The table below gives the probability distribution for X one more time. Compute and interpret the standard deviation of the random variable X .

Goals:	0	1	2	3	4	5	6	7	8	9
Probability:	0.061	0.154	0.228	0.229	0.173	0.094	0.041	0.015	0.004	0.001

Continuous Random Variable -

What is the difference between Discrete Random Variables and Continuous Random Variables?

6. A life insurance company sells a term insurance policy to a 21-year-old male that pays \$100,000 if the insured dies within the next 5 years. The probability that a randomly chosen male will die each year can be found in modality tables. The company collects a premium of \$250 each year as payment for the insurance. The amount Y that the company earns on this policy is \$250 per year, less the \$100,000 that it must pay if the insured dies. Here is a partially completed table that shows information about risk of mortality and the values of $Y = \text{profit earned by the company}$:

Age at death:	21	22	23	24	25	26 or more
Profit:	-\$99,750	-\$99,500				
Probability:	0.00183	0.00186	0.00189	0.00191	0.00193	

- a. Fill in the missing values of Y and find the missing probability.

Age at death:	21	22	23	24	25	26 or more
Profit:	-\$99,750	-\$99,500				
Probability:	0.00183	0.00186	0.00189	0.00191	0.00193	

- b. Calculate the mean μ_Y . Interpret this value in context.

- c. It would be quite risky for you to insure the life of a 21-year-old friend under the terms of the insurance policy. There is a high probability that your friend would live and you would gain \$1,250 in premiums. But if he were to die, you would lose almost \$100,000. Explain carefully why selling insurance is not risky for an insurance company that insures many thousands of 21-year-old men.

- d. The risk of an investment is often measured by the standard deviation of the return on the investment. The more variable the return is, the riskier the investment. We can measure the great risk of insuring a single person's life by computing the standard deviation of the income Y that the insurer will receive. Find σ_Y using the distribution and mean.

7. In government data, a household consists of all occupants of a dwelling unit, while a family consists of two or more persons who live together and are related by blood or marriage. So all families form household, but some households are not families. Here are the distributions of household size and family size in the United States:

	Number of Persons						
	1	2	3	4	5	6	7
Household probability	0.25	0.32	0.17	0.15	0.07	0.03	0.01
Family probability	0	0.42	0.23	0.21	0.09	0.03	0.02

Let X = the number of people in a randomly selected U.S. household and Y = the number of people in a randomly chosen U.S. family.

- a. Find the mean for each random variable. Explain why this difference makes sense.
- b. Find the standard deviation of both X and Y . Explain why this difference makes sense.

CHAPTER 6

8. A study of 12,000 able-bodied male students at the University of Illinois found that their times for the mile run were approximately Normal with mean 7.11 minutes and standard deviation 0.74 minutes.
- Choose a student at random from this group and call his time for the mile Y . Find $P(Y < 6)$ and interpret the result.

- Find the $P(Y \leq 6)$.

Freeport AP Statistics

Chapter 6: Random Variables

6.2 Transforming and Combining Random Variables

OBJECTIVE(S):

- Students will learn how to describe the effects of transforming a random variable by adding or subtracting a constant and multiplying or dividing by a constant.
- Students will learn how to find the mean and standard deviation of the sum or difference of independent random variables.
- Students will learn how to determine whether two random variables are independent.
- Students will learn how to find the probabilities involving the sum or difference of independent Normal random variables.

Effect on a Random Variable of Multiplying (or Dividing) by a Constant

-
-
-

Effect on a Random Variable of Adding (or Subtracting) by a Constant

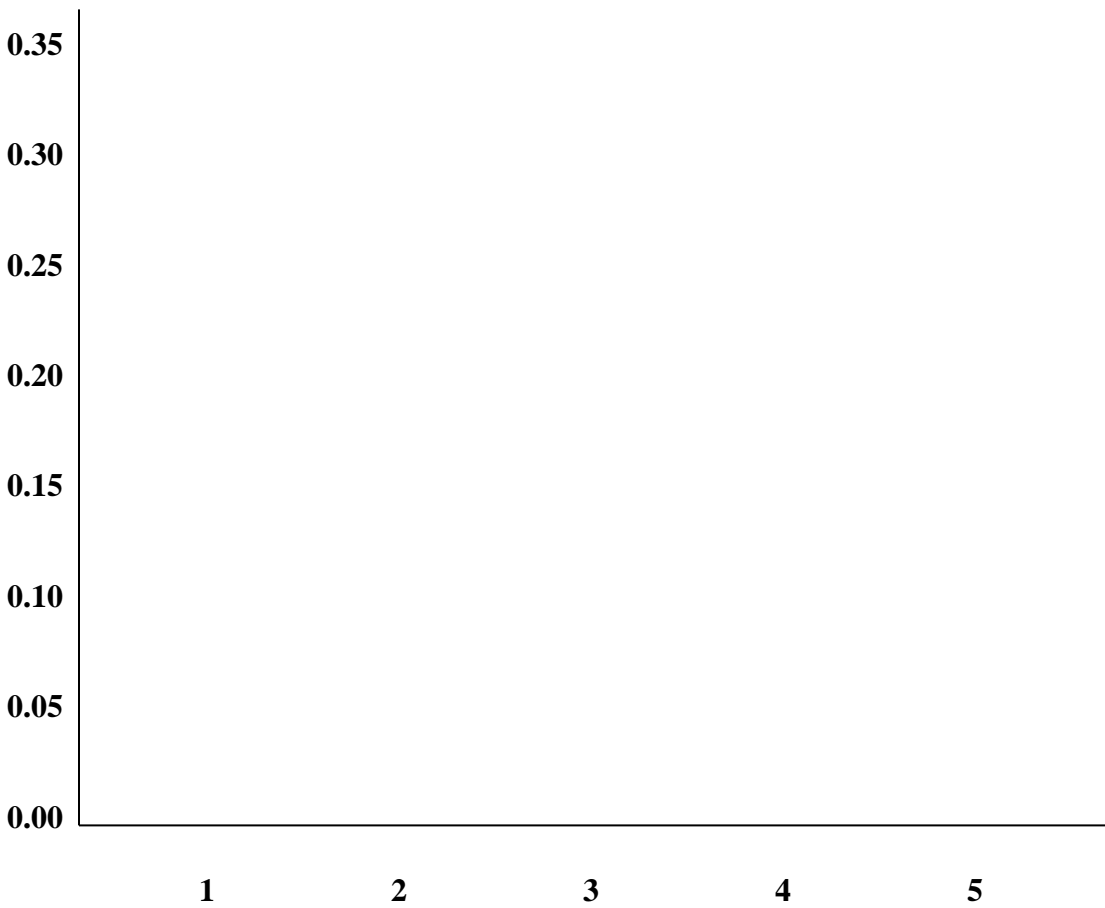
-
-

9. A report of the National Center for Health Statistics says that the height of a 20-year-old man chosen at random is a random variable H with mean 5.8 feet (ft) and standard deviation 0.24 ft. Find the mean and standard deviation of the height J of a randomly selected 20-year-old man in inches. There are 12 inches in a foot.

10. Ana is a dedicated Skee Ball player who always rolls for the 50-point slot. The probability distribution of Ana's score X on a single roll of the ball is shown below. You can check that $\mu_x = 23.8$ and $\sigma_x = 12.63$.

Score:	10	20	30	40	50
Probability:	0.32	0.27	0.19	0.15	0.07

- a. A player receives one ticket from the game for ever 10 points scored. Make a graph of the probability distribution for the random variable T = number of tickets Ana gets on a randomly selected throw.



- b. Find and interpret μ_T .

- c. Compute and interpret σ_T .

11. Most states and Canadian provinces have government-sponsored lotteries. Here is a simple lottery wager, from the Tri-State Pick 3 game that New Hampshire shares with Maine and Vermont. You choose a number with 3 digits from 0 to 9; the state chooses a three-digit winning number at random and pays you \$500 if your number is chosen. Because there are 1000 numbers with three digits, you have probability $1/1000$ of winning. Taking X to be the amount your ticket pays you, the probability distribution of X is

Payoff X:	\$0	\$500
Probability:	0.999	0.001

- a. Calculate the μ_X and σ_X .
- b. If you buy a Pick 3 ticket, your winnings are $W = X - 1$, because it costs \$1 to play. Find the mean and standard deviation of W . Interpret each of these values in context?
12. Rotter Partners is planning a major investment. The amount of profit X (in millions of dollars) is uncertain, but an estimate gives the following probability distribution:

Profit:	1	1.5	2	4	10
Probability:	0.1	0.2	0.4	0.2	0.1

Based on this estimate, $\mu_X = 3$ and $\sigma_X = 2.52$. Rotter Partners owes its lender a fee of \$200,000 plus 10% of the profits X . So the firm actually retains $Y = 0.9X - 0.2$ from the investment. Find the mean and standard deviation of the amount Y that the firm actually retains from the investment.

Effects of a Linear Transformation on the Mean and Standard Deviation

-
-
-

Mean of the Sum of Random Variables

Independent Random Variables -

Variance of the Sum of Random Variables

Mean of the Difference of Random Variables

Variance of the Difference of Random Variables

13. The academic motivation and study habits of female students as a group are better than those of males. The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures these factors. The distribution of SSHA scores among the women at a college has mean 120 and standard deviation 28, and the distribution of scores among male students has mean 105 and standard deviation 35. You select a single male student and a single female student at random and give them the SSHA test.
- Explain why it is reasonable to assume that the scores of the two students are independent.
 - What are the expected value and standard deviation of the difference (female minus male) between their scores?
 - From the information given, can you find the probability that the woman chosen scores higher than the man? If so, find this probability. If not, explain why you cannot.

14. Typographical and spelling errors can be either “nonword errors” or “word errors.” A nonword error is not a real word, as when “the” is typed as “teh.” A word error is a real word, but not the right word, as when “lose” is typed as “loose.” When students are asked to write a 250-word essay (without spell-checking), the number of nonword errors X has the following probability distribution:

Value of X:	0	1	2	3	4
Probability:	0.1	0.2	0.3	0.3	0.1

$$\mu_x = 2.1 \quad \sigma_x = 1.136$$

The number of word errors Y has this probability distribution:

Value of Y:	0	1	2	3
Probability:	0.4	0.3	0.2	0.1

$$\mu_y = 1.0 \quad \sigma_y = 1.0$$

- a. Find the mean and standard deviation of the difference $Y - X$ in the number of errors made by a randomly selected student. Interpret each value in context.

- b. Find the probability that a randomly selected student makes more word errors than nonword errors.

15. The design of an electronic circuit for a toaster calls for a 100-ohm resistor and a 250-ohm resistor connected in series so that their resistances add. The components used are not perfectly uniform, so that the actual resistances vary independently according to Normal distributions. The resistance of 100-ohm resistors has mean 100 ohms and standard deviation 2.5 ohms, while that of 250 ohm resistors has mean 250 ohms and standard deviation 2.8 ohms.
- What is the distribution of the total resistance of the two components in series?

 - What is the probability that the total resistance lies between 345 and 355 ohms? Show your work.
16. Leona and Fred are friendly competitors in high school. Both are about to take the ACT college entrance examination. They agree that if one of them scores 5 or more points better than the other, the loser will buy the winner a pizza. Suppose that in fact Fred and Leona have equal ability, so that each score varies Normally with mean 24 and standard deviation 2. (The variation is due to luck in guessing and the accident of the specific questions being familiar to the student.) The two scores are independent. What is the probability that the scores differ by 5 or more points in either direction?

Freeport AP Statistics

Chapter 6: Random Variables

6.3 Binomial and Geometric Random Variables

OBJECTIVE(S):

- Students will learn how to determine whether the conditions for a binomial random variable are met.
- Students will learn how to calculate the mean and standard deviation of a binomial random variable and interpret these values in context.
- Students will learn how to compute and interpret probabilities involving binomial distributions.

Binomial Setting -

What are the four conditions for a binomial setting?

- **B**
- **I**
- **N**
- **S**

Binomial Random Variable -

Binomial Distribution -

17. Determine whether the random variables below have a binomial distribution. Justify your answer.

a. Roll a fair die 10 times and let X = the number of sixes.

b. Shoot a basketball 20 times from various distances on the court. Let Y = number of shots made.

c. Observe the next 100 cars that go by and let C = color.

Mean and Standard Deviation of a Binomial Random Variable

$$\mu_X =$$

$$\sigma_X =$$

20. The makers of a diet cola claim that its taste is indistinguishable from the full-calorie version of the same cola. To investigate, an AP Statistics student named Maurice prepared small samples of each type of soda in identical cups. Then she had volunteers taste each cola in a random order and try to identify which was the diet cola and which was the regular cola. Overall, 23 of the 30 subjects made the correct identification. If we assume that the volunteers really couldn't tell the difference, then each one was guessing with a $\frac{1}{2}$ chance of being correct. Let X = the number of volunteers who correctly identify the colas.
- Explain why X is a binomial random variable.
 - Find the mean and the standard deviation of X . Interpret each value in context.
 - Of the 30 volunteers, 23 made correct identifications. Does this give convincing evidence that the volunteers can taste the difference between the diet and regular colas?

Sampling Without Replacement Condition

21. A federal report finds that lie detector tests given to truthful persons have probability about 0.2 of suggesting that the person is deceptive. A company asks 12 job applicants about thefts from previous employers, using a lie detector to assess their truthfulness. Suppose that all 12 answer truthfully. Let X = the number of people who the lie detector says are being deceptive.
- Find and interpret μ_X .
 - Find and interpret σ_X .
 - Let Y = the number of people who the lie detector says are telling the truth.
 - Find $P(Y \geq 10)$. How is this related to $P(X \leq 2)$? Explain.
 - Calculate μ_Y and σ_Y . How do they compare with μ_X and σ_X ? Explain why this makes sense.

22. As a special promotion for its 20-ounce bottles of soda, a soft drink company printed a message on the inside of each cap. Some of the caps said, "Please try again," while others said, "You're a winner!" The company advertised the promotion with the slogan "1 in 6 wins a prize." Suppose the company is telling the truth and that every 20-ounce bottle of soda it fills has a 1-in-6 chance of being a winner. Seven friends each buy one 20-ounce bottle of the soda at a local convenience store. Let X = the number who win a prize.
- Explain why X is a binomial random variable.

- Find the mean and standard deviation of X . Interpret each value in context.

CHAPTER 6

23. In the game of Scrabble, each player begins by drawing 7 tiles from a bag containing 100 tiles. There are 42 vowels, 56 consonants, and 2 blank tiles in the bag. Cait chooses her 7 tiles and is surprised to discover that all of them are vowels. Can we use a binomial distribution to approximate this probability? Justify your answer.

Geometric Setting

- B
- I
- T
- S

Geometric Random Variable –

Geometric Distribution -

Geometric Probability

Mean (Expected Value) of a Geometric Random Variable

CHAPTER 6

25. Alan decides to use a different strategy for the 1-in-6 wins game. He keeps buying one 20-ounce bottle of the soda at a time until he gets a winner.
- a. Find the probability that he buys exactly 5 bottles.

- b. Find the probability that he buys no more than 8 bottles.

26. According to Benford's law, the probability that the first digit of the amount of a randomly chosen invoice is an 8 or a 9 is 0.097. Suppose you examine randomly selected invoices from a vendor until you find one whose amount begins with an 8 or a 9.

a. How many invoices do you expect to examine until you get one that begins with an 8 or 9? Justify your answer.

b. In fact, you don't get an amount starting with an 8 or 9 until the 40th invoice. Do you suspect that the invoice amounts are not genuine? Compute an appropriate probability to support your answer.