

Statistics, Day 1

Probability Distributions

Do all work on a separate piece of paper.

1. A student randomly guesses the answer for each question of a 6-question True/False exam. Let the random variable X be the number of questions answered correctly.
 - a. List the possible numerical outcomes (sample space) of X .
 - b. Find the probability of each outcome (Hint: it's binomial!)
 - c. Construct a histogram showing the probability distribution. Be sure to label your horizontal and vertical axis.
 - d. What is the probability that the student will get 4 or more questions correct?
 - e. What is the expected value of the number of questions the student will answer correctly?
2. You are going to roll a dice 50 times.
 - a. What is the sample space?
 - b. What is the probability of each outcome?
 - c. Construct a histogram showing the probability distribution. Label your horizontal and vertical axes.
 - d. What is the expected value of a single dice roll? Will you ever actually roll that value?
 - e. Does your probability distribution represent theoretical or experimental probability?
3. Now you will construct an experimental probability distribution for rolling a dice 50 times. Use your calculator MATH PRB randInt (1, 6) to simulate rolling a dice one time. Repeat 50 rolls of the dice and record the frequency of each outcome in a probability distribution table.

- a. Use your data to construct a frequency table and a probability distribution.

Outcome	1	2	3	4	5	6
Observed Frequency						
Probability						

- b. Construct a probability histogram for your data. Compare your experimental distribution to the theoretical distribution you drew for #2.
4. The following data show the salaries (in thousands of dollars) of 28 employees of a small company:
30, 33, 35, 41, 42, 45, 45, 50, 52, 53, 55, 57, 60, 60, 61, 62, 63, 65, 67, 70, 71, 72, 73, 98, 105, 125, 150, 175
 - a. Construct a frequency table for the data. What class intervals will you use?
 - b. Draw a frequency histogram for the data. Include a title and label your axes.
 - c. Suppose that an adult is randomly selected from this company. What is the probability that s/he earns more than \$75,000?
 5. In the Pick 3 Lottery you choose a 3 digit number. If your number matches the number selected by the lottery board, you win \$500. Otherwise, you win nothing.
 - a. What are the possible numerical outcomes for your winnings?
 - b. What is the probability of each outcome?
 - c. Construct a histogram showing the probability distribution.
 - d. Find the expected value of your winnings.
 - e. A game is considered "fair" if the cost to play is equal to the expected value of the winnings. What would be a fair price to charge for a lottery ticket?

Statistics, Day 2

Describing Distributions

1. Describe the shape of each of the distributions you constructed for Day 1.
2. The following data show the salaries (in thousands of dollars) of 28 employees of a small company:
(Note: Data is the same as Day 1, # 4)
30, 33, 35, 41, 42, 45, 45, 50, 52, 53, 55, 57, 60, 60, 61, 62, 63, 65, 67, 70, 71, 72, 73, 98, 105, 125, 150, 175
 - a. What are the mean and median salaries? Which one is a better measure of center? Why?
 - b. Find the range and IQR of the data. Which one is a better measure of the spread of the data? Why?
 - c. Find the standard deviation.
 - d. Use the 1.5 IQR rule to determine if the data has any outliers.
 - e. Draw a box plot of the data showing outliers (if present).

3. Use the formula $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$ to calculate the standard deviation of 0, 2, 5, 8, 10. What does this value tell you about the data?

4. The heights of a random sample of 19 men are recorded below:

69.9, 71.8, 72.1, 73.1, 73.8, 70.6, 69.4, 69.6, 76.2, 71.8, 74.6, 66.9, 69.1, 66.7, 70.4, 71.8, 69.3, 72.3, 71.5

- a. Use your calculator to construct a histogram of the data. Use class interval widths of 2 inches.
 - b. Describe the shape of the distribution.
 - c. Find the mean, median and mode (class interval with the highest frequency). What do you notice about the three measures of center?
 - d. Construct a box plot for the data. Are there any outliers?
 - e. What interval of heights contains the middle 50% of the men?
 - f. What height interval contains the shortest 25% of the men? The tallest 25%?
5. A student randomly guesses the answer for each question of a 6-question True/False exam.
(Note: you constructed the probability distribution on Day 1, Q #1)
 - a. Enter the numerical values of the sample space into L1 of your calculator. Enter the probability of each outcome into L2.
 - b. To have your calculator construct the probability histogram, press 2nd STAT PLOT 1: PLOT 1...On. Choose the histogram. Set Xlist: L1 and Freq: L2. (This allows the probabilities in L2 to be used as frequencies). Use the same window as you used for the graph yesterday. Is the calculator's graph the same as yours?
 - c. Follow these instructions carefully: Press STAT CALC 1-Var Stats ENTER L1, L2 ENTER . (The calculator will know to use the values in L2 as frequencies for L1). What is the mean of the distribution? Is it the same as the expected value you found yesterday?
 - d. Does the calculator report a sample standard deviation? For a probability distribution, the calculator provides the population standard deviation σ . What is this value?
 - e. Report Q1, Median, and Q3 for the distribution.
 - f. Use your calculator to construct a box plot.

Statistics, Day 3
Normal Distributions

Complete the questions using the 68%-95%-99.7% Empirical Rule or the table.

Table entries show the percent (P) of observations in a Normal distribution that are less than the z-score.

z	-3.0	-2.5	-2.0	-1.5	-1.0	-.50	0	.50	1.0	1.5	2.0	2.5	3.0
P	0.13	0.6	2.3	6.7	15.9	30.9	50%	69.1	84.1	93.3	97.7	99.4	99.87

- The lengths of babies born at City Hospital last year were approximately Normally distributed with a mean of 20 inches and a standard deviation of 1 inch.
 - Draw a Normal curve showing the mean and ± 1 , ± 2 , ± 3 std. dev.
 - What percent of the babies born were longer than 20 inches?
 - What percent of the babies born were between 19 and 22 inches?
 - Would it be unusual for a baby to be more than 23 inches? Explain your answer.
 - What interval centered about the mean would contain 95% of all babies born?
- The amount of cola that a machine puts into soda cans is approximately Normally distributed with a mean of 355 mL and a standard deviation of 2 mL. Assume that the machine fills 2000 can a day.
 - Draw a Normal curve showing the mean and ± 1 , ± 2 , ± 3 std. dev.
 - About how many can will contain more than 359 mL of cola?
 - About how many cans will contain between 353 and 357 mL of cola?
 - About how many cans will contain less than 352 mL of cola?
 - If the size of the can is 360 mL, about how many cans will overflow?
- Scores on the Weschler Adult Intelligence Scale are approximately Normally distributed with a mean of 110 and a standard deviation of 25.
 - Approximately what percent of adults will score below 98?
 - Approximately what percent of adults will score above 135?
 - MENSA is an elite organization that will only admit the top 2% of applicants. Approximately what score on the Weschler exam would an adult need to qualify for MENSA?
- A math teacher will be grading a test using a Normal curve. The mean score on the test was 80 and the standard deviation was 8.
 - What percent of the students will get A's if a score of 88 or above is an A?
 - What percent of students will get B's if a score of 80 to 88 is a B?
 - What percent of students will get C's if a score of 68 to 80 is a C?
 - The teacher only wants about 2% of his students to fail. What score should be the lowest passing (D) grade?
- When administered to 4th graders, a test of reading ability has a mean of 75 and standard deviation of 12. Sixth graders have a mean score of 85 with a standard deviation of 8 on a similar test. A young student scored 71 on the reading test as a 4th grader and 79 as a 6th grader.
 - Compute the z-score of the student as a 4th grader and a 6th grader.
 - Relative to his classmates, is his reading performance as measured by the test improving?
 - Do you need to know if the scores are Normally distributed to answer part a & b?

Statistics, Day 4

Review

- A basket ball player who consistently makes 80% of his free throws attempts 5 free-throws during a game. Let the random variable X be the number of free throws made.
 - List the possible numerical outcomes for X .
 - Find the probability of each outcome.
 - Construct a histogram to display the probability distribution.
 - Describe the shape of the distribution.
 - What is the expected value of X . Explain what this number tells you.

- The weight in pounds of children in a 4th grade class are given below

64	71	57	67	74	65	59	62	67	72
84	60	68	72	91	77	69	76	88	99

- Use your calculator to construct a histogram of the data.
 - Describe the shape, center and spread of the data.
 - Use the 1.5 IQR rule to determine if there are any outliers.
 - Construct a box plot of the data.
 - What weight interval contains the middle 50% of the students?
- A survey of 100 randomly selected families in a small community asked how many pets lived in their household. The responses are given below.

Number of Pets	0	1	2	3	4	5	6	7
Frequency	14	22	20	18	10	8	6	2

- Identify the random variable and its possible numerical outcomes.
 - What is the probability that a family selected at random has 5 or more pets?
 - What is the expected value for the number of pets in a household in this community?
- Use the formula $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$ to calculate the standard deviation of 7, 8, 9, 10, 11, 12, 13. Explain what this value tells you about the data.
 - The commute times for employees at a large company follow an approximately Normal distribution with a mean of 33 minutes and a standard deviation of 7 minutes.
 - Draw a Normal curve showing the mean and ± 1 , ± 2 , ± 3 std. dev.
 - What percent of the company's employees commute less than 26 minutes?
 - What is the commute time of the middle 95% of the company's employees?
 - What percent of employees commute between 26 and 47 minutes?
 - Would you be surprised to hear that many employees have commute times over 50 minutes?
 - The midterm scores in Professor Normal's statistic course had a mean of 84 and a standard deviation of 8. The final exam scores had a mean of 77 and a standard deviation of 11. Student A scored 72 on the midterm and 66 on the final. Has her test score relative to the other students in the class improved? Explain your answer.