

Math 107

Review for final test

1. You roll a six sided die two times. What is the probability that you do not get a three on either roll? $5/6 * 5/6 = 25/36 \approx .694$
2. Consider a box with 5 blue balls, 7 red balls and 3 black balls. A single ball is drawn from the box
 - a) What is the probability that the ball is NOT red? $8/15$
 - b) What is the probability for a blue ball being drawn? $5/15$

3. For ten days a dieter counts calories for their diet:

1200	1150	980	1450	1200
1000	1350	900	1500	1300

- a) Calculate the mean. 1203
- b) What is the median? 1200
- c) What value is the mode? 1200
- d) Give the five number summary and make a box plot

$$\text{Min} = 900$$

$$\text{Q1} = 1000$$

$$\text{Med} = 1200$$

$$\text{Q3} = 1350$$

$$\text{Max} = 1500$$

4. You roll a pair of six –sided dice and add up the total shown on the top
 - a) What is the probability that you get a sum of 7 or a sum less than 4? $9/36$
 - b) What is the probability you get “doubles” given the sum is greater than 7? $3/15$
 - c) What is the probability you get at least one three? $11/36$
5. A roulette wheel has 38 slots around the rim, the first 36 slots are numbered from 1 to 36. Half of these slots are red and the other half are black. The remaining 2 slots are numbered 0 and 00 and are green. A roulette wheel is spun in one direction, a small ivory ball is rolled along the rim in the opposite direction. The ball has an equally likely chance of falling into any of the 38 slots. Find:
 - a) The probability of a red. $18/38$
 - b) The probability you do not get an odd number. $20/38$
 - c) If you bet on the number 19 and win you get \$35 if you lose you lose you \$1 bet. What is the expected value of this game? Explain the meaning of this value?
$$1/38 * 35 + 37/38 * (-1) = -2/38$$
On average, every time you play this game you will lose 5.26 cents.

6. Suppose a drug test has a sensitivity of 95% and a specificity of 95%. Further suppose that in a certain population 5% of the population has the disease. Suppose the population is numbered 10,000

a) Fill in the accompanying table.

	Has disease	Does not have disease	Totals
Test positive	475	475	950
Test negative	25	9025	9050
Totals	500	9500	10,000

b) What is the PPV and NPV for this disease with this population? **PPV=50%**
NPV=99.7%

7. Consider the information given in the table below, which is from a medical research experiment

	Has emphysema	No emphysema
Smoker	5460	9850
Non-Smoker	1150	12,150

- a) What is the probability a person selected from this chart has emphysema? **23.1%**
 b) What is the probability that a person selected from this chart has emphysema or is a smoker? **57.5%**
 c) What is the probability that a person selected from this chart has emphysema given they are a smoker? **35.7%**
8. Suppose a test for a disease has a sensitivity of 75% and a specificity of 85%. Further suppose that in a certain country with a population of 2,000,000, 20% of the population has the disease

a) Fill in the following chart.

	Has disease	Does not have disease	Totals
Test positive	300,000	240,000	540,000
Test negative	100,000	1,360,000	1,460,000
Totals	400,000	1,600,000	2,000,000

b) What are the PPV and NPV for this test in this country?

PPV=55.6% NPV=93.2%

9. A test for HIV has both a sensitivity and a specificity of 99%. This test applied to a population of 1,000,000 people. Suppose 0.1% of this population is actually infected with the disease

a) How many people with a positive test are actually disease free? 9990

b) What is the PPV and NPV? 990/10,980 989,010/989,020

	Has disease	Does not have disease	Totals
Test positive	990	9990	10,980
Test negative	10	989,010	989,020
Totals	1,000	999,000	1,000,000

10. Consider the information given in the table below, which is from a medical research experiment

	Has Lung Cancer	No Lung Cancer
Smoker	13,222	9,778
Non-Smoker	3153	14,847

a) What is the probability a person selected from this chart is a smoker? 56%

b) What is the probability that a person selected from this chart has lung cancer or is a smoker? 63.8%

c) What is the probability that a person selected from this chart has lung cancer given they are a smoker? 57.5%

11. Suppose you wish to study the amount of time the 3500 students at your college spend exercising each day. You ask 100 students who are residents in the dorms at your college to fill out a survey on daily exercise. 85 students fill out and return the survey. After receiving the results, you find the average amount of daily exercise is 30 minutes.

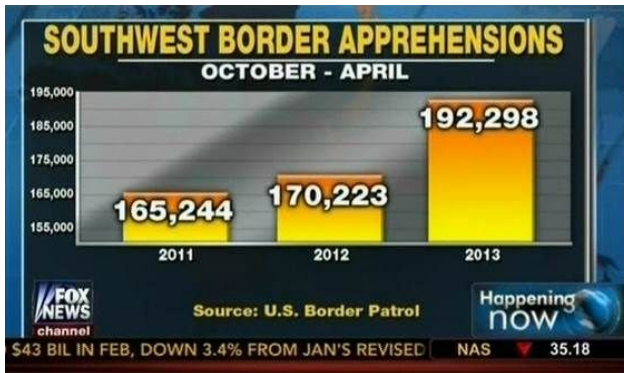
From this you conclude that, on average, the 3500 students at your college exercise about 30 minutes per day

- a) What is the population for this study? What is the sample for this study?
3500 students. 85 students.
- b) What is the raw data for this study? Information from the 85 students.
- c) What is the sample statistic for the study? 30 minutes
- d) Criticize the sampling technique of the experimenter - do you feel the results will be accurate? Why or why not? No, students all lived in a dorm which isn't representative of student population as a whole. Only 85 students returned survey out of 100. The data is self-reported and is probably inflated.

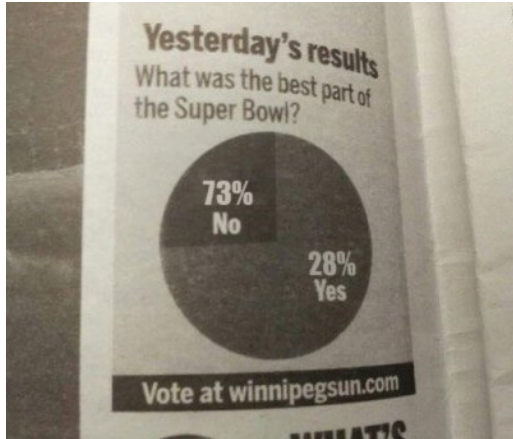
12. Consider the following graphs – explain in detail why they are misleading



The dates along the bottom aren't evenly spaced.
The increase in job loss looks over inflated due to the scaling of the vertical axis.

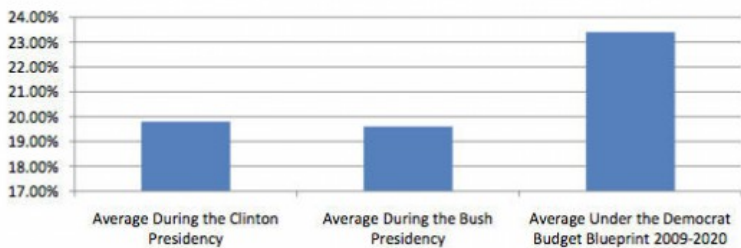


The increase is over emphasized due to scaling of the vertical axis.



Doesn't answer the question asked.
Proportions are off.

Federal Spending as a Share of the Economy

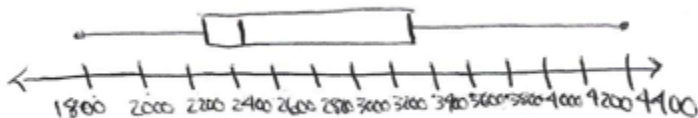


Vertical axis scaling over emphasizes the increase.

13. Here are monthly incomes of 20 employees for a small company

2600	1800	2500	4400	3200	2400
3400	2400	1900	3700	2300	2300

- Calculate the mean. 2742
- What is the median? 2450
- What is the mode? 2300, 2400
- Give the five number summary. Min = 1800, Q1=2300, Median=2450, Q3=3300, max=4400
- Draw (and correctly label) a boxplot for this data



For the information in problems 14-16, determine if the given study is an observational study or an experiment. Determine the population(s) being studied, the sample(s), and the statistic given

14. A Gallup survey done for CNN asked 2005 randomly selected Americans and found that 80% of Americans think we're less civil than we were 10 years ago and 67% think we are more likely to use vulgar language than we were 10 years ago.

Observational study. Population: Americans, Sample is 2005 Americans. Stats are 80% and 60%.

15. A survey of 235,812 first year college students revealed that 32.4% of these students had an A average in high school.

Observational Study. Population = first year college students. Sample is 235,812 first year students. Stat is 32.4%

16. To determine the validity of new flu drug, you take a sample of 2000 individuals with the disease. You divide the individuals into two groups randomly and give one group the new drug and another group a placebo. You discover that 85% of group which took the drug recovered, while only 43% of those with the placebo recovered.

Experiment. Population= people with the flu. Sample 2000 flu patients. Stats 85% and 43%

In problems 17-19, identify the sampling method used - is it a valid method?

17. You are doing a survey to investigate racial prejudice. You classify the United States population into five different racial groups: White, Hispanic, African-American, American Indian and Asian. You take a random sample from each group and uses these people as the sample for your study.

Stratified. Valid.

18. You manufacturer light bulbs and are interested in how many bulbs in each batch of 1000 bulbs are defective. You test each 10th bulb off the assembly line to determine if it is defective

Systematic. Valid.

19. To determine the opinions on the validity of the IRAN nuclear agreement you ask everyone in the food court at the Riverpoint mall on a given Friday afternoon.

Convenience. Not valid.

20. You have a plate of 40 cookies. Ten have chocolate chips and 15 have pecans. Five cookies have both chocolate chips and pecans. You select a cookie at random, what is the probability that

a. you get a cookie with both chocolate chips and pecans? **5/40**

b. you get cookie without either chocolate chips or pecans? **20/40**

c. you get a cookie with chocolate chips OR a cookie with pecans? **20/40**

21. You play the following game. You roll a single dice and get the dollar amount of the outcome (for example if you roll a 4 you get \$4). What is the expected value of this game? **\$3.50**

If you pay \$5 to play the game in problem 20 is this a "fair" game – why or why not? **No, you would lose on average \$1.50 per play.**

22. The following table deals with patients who have hepatitis B. They were grouped into the following categories

Source of Disease					
	Job related	Drug use	Sexual contact	Other	Totals
Female	150	155	150	19	474
Male	200	165	250	31	646
Totals	350	320	400	50	1120

Suppose one of the persons with hepatitis is randomly selected. Compute the following:

- $P(\text{person is female}) = 474/1120$
- $P(\text{person has source of disease job related}) = 350/1120$
- $P(\text{person is female OR has source of disease drug use}) = 639/1120$
- $P(\text{person is female AND has source sexual contact}) = 150/1120$
- $P(\text{person is male AND has source of disease other}) = 31/1120$
- $P(\text{male GIVEN person got the disease from sexual contact}) = 250/400$

23. Jeff and Jerry play the following game. They toss two coins. If both coins come up heads Bill pays Larry \$4. Otherwise Larry pays Bill \$1

- What is the expected value of this game for Bill? **-\$3**
- What is the expected value of this game for Larry? **\$3**
- If they play the game 100 times who should come out ahead and by how much?
Larry by \$300.

24. Suppose a company charges \$450 for a premium for a fire insurance policy. In the case of a fire claim the company will pay out \$100000. Suppose the probability of fire claim in a given year is 0.004. What is the expected annual profit (or loss) for the insurance company on this policy? What annual profit should the company expect if they 1000 such policies?

$$450 \cdot 1 + (-100,000) \cdot 0.004 = \$50 \text{ per policy or } 50,000 \text{ per } 1000 \text{ policies.}$$

25. You roll two dice and look at the sum you receive. If the sum is 2, 3, 4, 10, 11, or 12 you receive \$8. If the sum is 5,6,7,8,9 you lose \$2, Determine if this is fair game or not.

$$12/36 \cdot 8 + 24/36 \cdot (-2) = \$1.33$$