MATH 338: Statistics Applied to the Natural Sciences
Midterm 1 Review Questions

Department of Mathematics, California State University, Fullerton

Fall 2016

Notes

The personal department keeps records on all employees in a company. Here is the information that they keep in one of their data files: employee identification number, last name, first name, middle initial, department, number of years with the company, salary, education (coded as highschool, some college, or college degree), and age.

- What are the cases for this data?
- Which variables are categorical? Which are quantitative?
- Identify labels if any.

Notes

Notes

Consider the following heights of a sample of students:
$59, 64, 65, 72, 80$

- Make a stemplot of the data. Is the distribution right skewed, left skewed, or about symmetric?
- Find the mean and sd of the data set.
- How many standard deviations above the mean is 80? Would this observation be considered unusual?

Notes

- A statistics instructor wants to know which route will get her to school the fastest. Each day from October 2 to November 15, when she gets to the turn point she checks the odometer on her car. If it shows an even number she takes the freeway; if it shows an odd number, she takes the in-town route. She records the total time each day. This study is a(n):
  a observational study.
  b sample.
  c experiment.

- The administration at a large state university is interested in getting the opinions of students on a proposed instructional fee for use of computer labs on campus. They select a simple random sample of 50 freshman, a simple random sample of 50 sophomores, 50 juniors, and 50 seniors. This is an example of which sampling method: (systematic sample, stratified random sample, or simple random sample)

Notes

- The potential for lurking variables to affect the results is:
  a. less in a experimental study than in an observational study.
  b. maximized in an experimental study.
  c. the same in both an experimental study and an observational study.
  d. minimized in an observational study.
  e. more in an experimental study than in an observational study.

- In a controversial election district, 73% of registered voters are Democrat. A random survey of 500 voters had 68% Democrats. Are the percentage values parameters or statistics?

Notes

Which of the following define a complete and valid probability model?

- $S = \{A, B, C\}$; $P(A) = 1/3, P(B) = 2/3, P(C) = 1$
- $S = \{H, T\}$; $P(H) = 1/4, P(T) = 3/4$
- $S = \{W, X, Y, Z\}$; $P(W) = 1/4, P(X) = 1/4, P(Z) = 1/4$
- $P(RED) = 1/2, P(BLUE) = 1/2$

Notes

Suppose you have a litter of mice which consists of five males and three females. You randomly grab two of them (without replacement).

- What is the probability that they are both male?
- What is the probability that you get at least one female?
- What is the probability that one is male and one is female?

Notes

Consider a discrete random variable X with probability distribution:

$$P(X = x) = c(x^2 - 1) \quad \text{for } x \in \{1, 2, 3, 4\}$$

- Find the value of c.
- Find the expected value and standard deviation of this probability distribution.
- Find $P(X \geq 3)$.

Notes

One of the things that most bothers me about Windows operating system is when is crashes when I'm trying to shut it down. It seems to be more likely to crash when I've put it on standby (sleep) earlier in the day. Suppose the probability that it crashes when it has been on standby is earlier is 0.1, and the probability it crashes when it has not been on standby is 0.01. The probability that I put it on standby during a day's usage is 0.4. If you just walked into the room in time to see me trying to shutdown and have the computer crash, what is the probability that I had put it on standby earlier that day?

Notes

Suppose mouse weights are normally distributed with a mean of 22 grams and a standard deviation of 3 grams. A breeder is shipping out boxes of ten mice and wants no more than 8% of their boxes to have mice below a specified average weight. What weight should they use so that no more than 8% of their boxes will have an average mouse weight below that weight?

- If we took an SRS of size 10 of all cereals and counted how many were hot cereals, would that have a binomial distribution?
- If we took an SRS of size 10 of all cereals and had the same person taste each and say whether they liked that cereal or not, would the number of cereals the person likes have a binomial distribution?
- Suppose that 20% of all cereals have sugar content under 20grams per serving. If we took an SRS of size 10 of all cereals, what is the probability that more 3 of the sample cereals has a sugar content under 20 grams?

Notes

- Suppose that 20% of all cereals have sugar content under 20grams per serving. If we took an SRS of size 1000 of all cereals, what is the probability that more than 30% of the sample cereals has a sugar content under 20 grams?

Notes

Important terms from Chapter 1

Notes

- cases, labels, variables, values
- categorical and quantitative variables
- distribution of a variable
- bar graph, pie chart, histogram, boxplot, stem-and-leaf plot, time plot
- outliers, skewness, symmetric
- mean, median, mode, standard deviation, variance, IQR, quantile/percentile, minimum, maximum
- density curve, standard normal distribution, normal distribution, 68-95-99.7 rule, standardizing (z-scores), normal quantile plot

Important terms from Chapter 3

Notes

- anectdotal data and available data
- sample and population
- sample survey, census, observational study, experiment, randomized comparative experiment
- lurking variables, counfounding, placebo effect, placebo treatment, control group, treatment group
- experimental units, subjects, treatments, outcome, explanatory variables, response variables, bias, double-blind, block, matched pairs,
- simple random sampling, voluntary response sample, stratified random sample, multistage random sample,
- undercoverage/nonresponse, poorly worded questions, response bias
- parameter, statistic, statistical inference
- sampling distribution, sampling variability, margin of error

Important terms from Chapter 4

Notes

- (discrete/continuous) random variable, probability distribution, probability histogram
- events, outcomes, sample space, probability, law of large numbers
- independent rvs, dependent rvs, mutually exclusive/disjoint
- Rules of Probability
- Addition Rule, Multiplication Rule, Conditional Rule, Complement, Bayes Theorem
- Expected Value, Variance, and SD of random variables

Important terms from Chapter 5

Notes

- Central Limit Theorem
- sampling distribution of the sample mean, sampling distribution of the sample sum
- Binomial distribution, Binomial probability
- Normal approximation to the Binomial, sampling distribution of binomial counts, sampling distribution of the sample proportion