

Interpreting Categorical and Quantitative Data

S-ID

Summarize, represent, and interpret data on a single count or measurement variable

1. Represent data with plots on the real number line (dot plots, histograms, and box plots). (S-ID.1.)

An election involving four candidates for mayor has been held. Of the following, which is the best way to present the percentage of votes each candidate received?

- A. Circle graph
- B. Line graph
- C. Box plot
- D. Scatterplot
- E. Histogram

2005-12-12-10

Source: National Assessment of Educational Progress, 2005, Grade 12 Mathematics Assessment.

2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets. (S-ID.2.)

A certain company keeps a list of 50 employees and their annual salaries. When the salary of the very highly paid president is added to this list, which of the following statistics is most likely to be approximately the same or nearly the same for the original list and the new list?

- A. The highest salary
- B. The range
- C. The mean
- D. The median
- E. The standard deviation

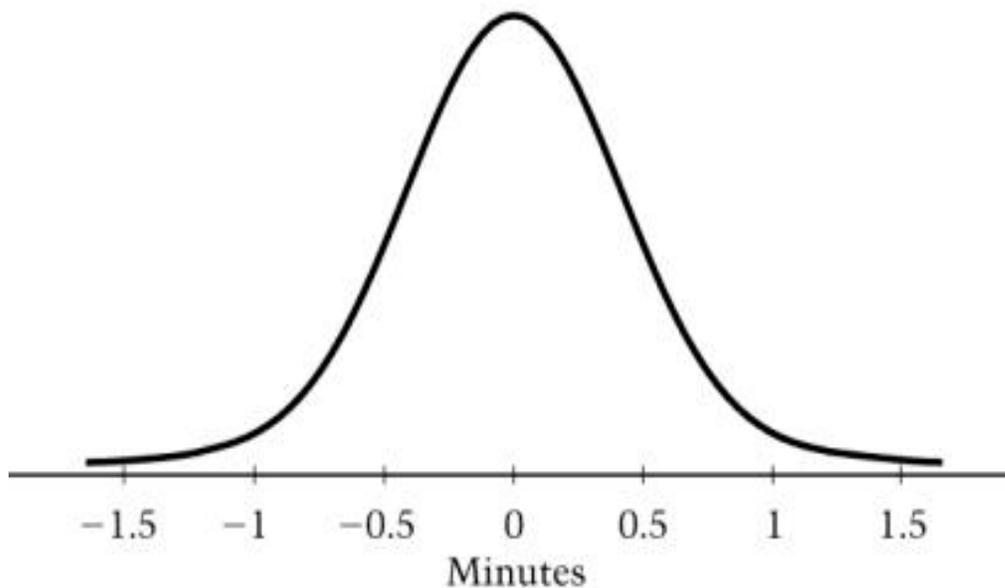
1992-12-15-7

Source: National Assessment of Educational Progress, 1992, Grade 12 Mathematics Assessment.

3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers). (S-ID.3.)

4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve. (S-ID.4.)

A clock manufacturer has found that the amount of time their clocks gain or lose per week is normally distributed with a mean of 0 minutes and a standard deviation of 0.5 minute, as shown below.



In a random sample of 1,500 of their clocks, which of the following is closest to the expected number of clocks that would gain or lose more than 1 minute per week?

- A. 15
- B. 30
- C. 50
- D. 70
- E. 90

2005-12-12-16

Source: National Assessment of Educational Progress, 2005, Grade 12 Mathematics Assessment.

Summarize, represent, and interpret data on two categorical and quantitative variables

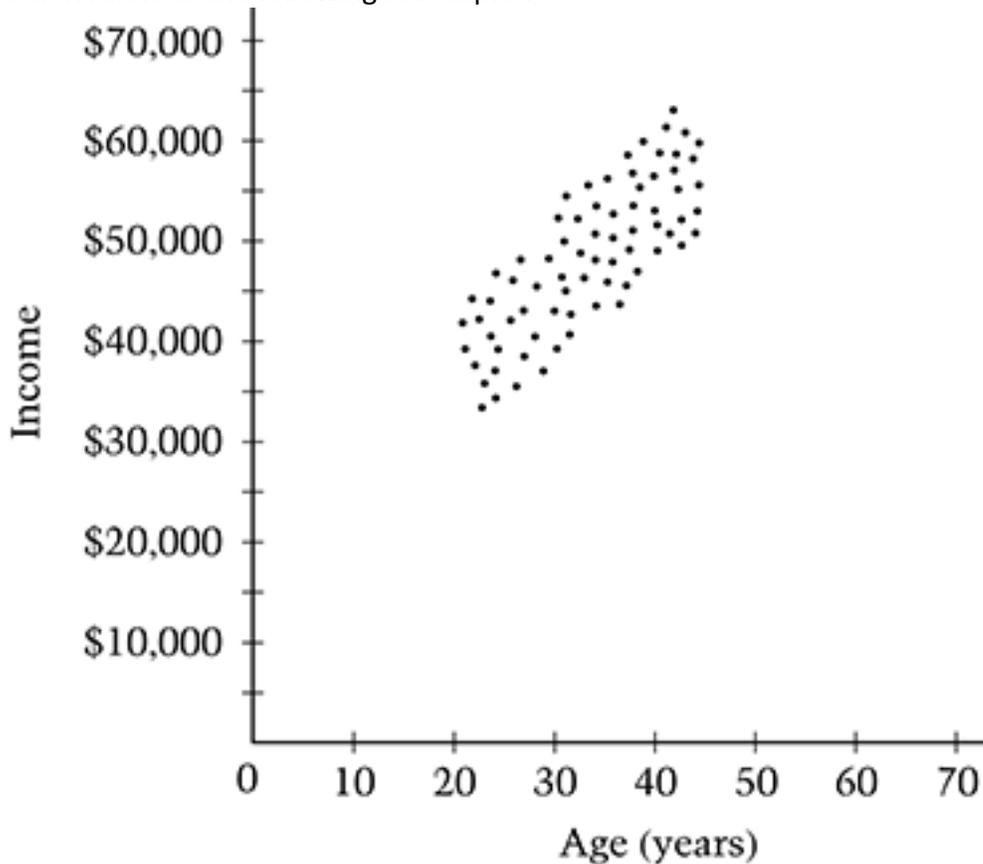
5. Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data. (S-ID.5.)

6. Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.
 - a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.
 - b. Informally assess the fit of a function by plotting and analyzing residuals.
 - c. Fit a linear function for a scatter plot that suggests a linear association. (S-ID.6.)

Interpret linear models

7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. (S-ID.7.)

This question refers to the following scatterplot.

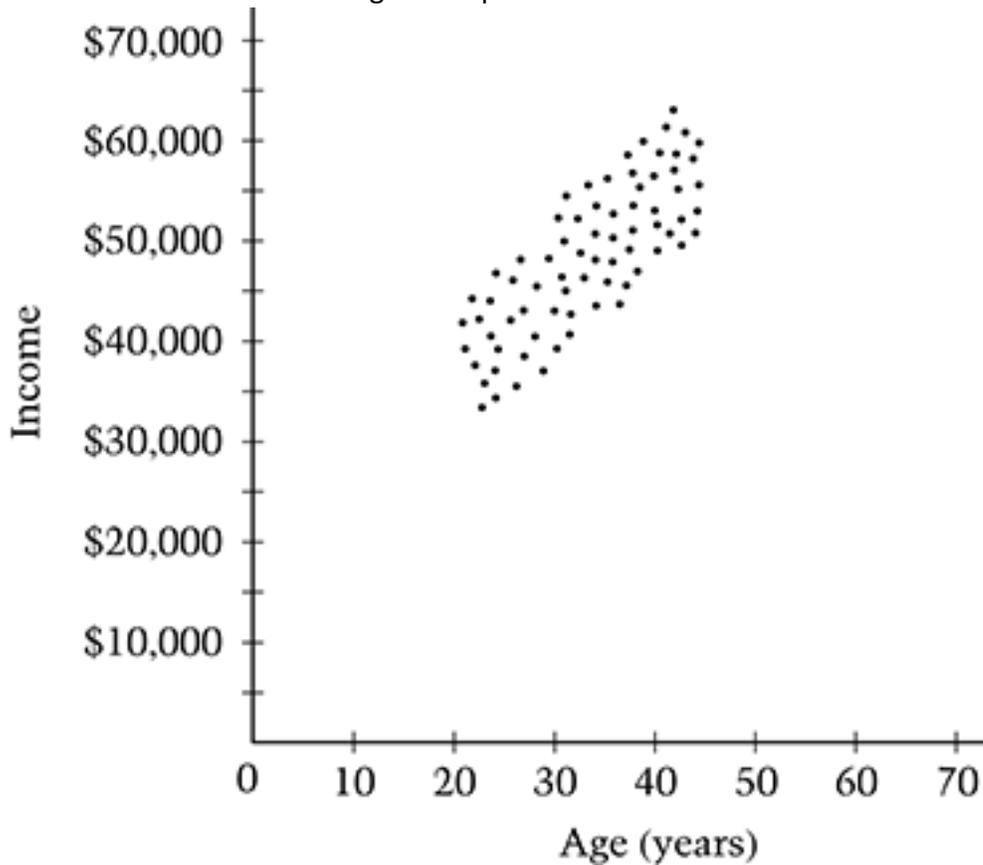


A random sample of graduates from a particular college program reported their ages and incomes in response to a survey. Each point on the scatterplot above represents the age and income of a different graduate.

Of the following equations, which best fits the data above?

- A. $y = -1,000x + 15,000$
- B. $y = 1,000x$
- C. $y = 1,000x + 15,000$
- D. $y = 10,000x$
- E. $y = 10,000x + 15,000$

This question refers to the following scatterplot.



A random sample of graduates from a particular college program reported their ages and incomes in response to a survey. Each point on the scatterplot above represents the age and income of a different graduate.

Based on the data in the scatterplot, predictions can be made about the income of a 35 year old and the income of a 55 year old. For which age is the prediction more likely to be accurate?

35 year old 55 year old

Justify your answer.

2009-12-2-4

Source: National Assessment of Educational Progress, 2009, Grade 12 Mathematics Assessment.

8. Compute (using technology) and interpret the correlation coefficient of a linear fit. (S-ID.8.)

9. Distinguish between correlation and causation. (S-ID.9.)

Making Inferences and Justifying Conclusions

S-IC

Understand and evaluate random processes underlying statistical experiments

1. Understand statistics as a process for making inferences about population parameters based on a random sample from that population. (S-IC.1.)

The principal of a high school would like to determine why there has been a large decline during the year in the number of students who buy food in the school's cafeteria. To do this, 25 students from the school will be surveyed. Which method would be the most appropriate for selecting the 25 students to participate in the survey?

- A. Randomly select 25 students from the senior class.
- B. Randomly select 25 students from those taking physics.
- C. Randomly select 25 students from a list of all students at the school.
- D. Randomly select 25 students from a list of students who eat in the cafeteria.
- E. Give the survey to the first 25 students to arrive at school in the morning.

2009-12-2-8

Source: National Assessment of Educational Progress, 2009, Grade 12 Mathematics Assessment.

2. Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. For example, a model says a spinning coin falls heads up with probability 0.5. Would a result of 5 tails in a row cause you to question the model? (S-IC.2.)

Make inferences and justify conclusions from sample surveys, experiments, and observational studies

3. Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each. (S-IC.3.)

4. Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling. (S-IC.4.)

5. Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant. (S-IC.5.)

6. Evaluate reports based on data. (S-IC.6.)

Conditional Probability and the Rules of Probability

S-CP

Understand independence and conditional probability and use them to interpret data

1. Describe events as subsets of a sample space (the set of outcomes) using characteristics (or categories) of the outcomes, or as unions, intersections, or complements of other events ("or," "and," "not") (S-CP.1.)

2. Understand that two events A and B are independent if the probability of A and B occurring together is the product of their probabilities, and use this characterization to determine if they are independent. (S-CP.2.)

3. Understand the conditional probability of A given B as $P(A \text{ and } B)/P(B)$, and interpret independence of A and B as saying that the conditional probability of A given B is the same as the probability of A, and the conditional probability of B given A is the same as the probability of B. (S-CP.3.)

4. Construct and interpret two-way frequency tables of data when two categories are associated with each object being classified. Use the two-way table as a sample space to decide if events are independent and to approximate conditional probabilities. For example, collect data from a random sample of students in your school on their favorite subject among math, science, and English. Estimate the probability that a randomly selected student from your school will favor science given that the student is in tenth grade. Do the same for other subjects and compare the results. (S-CP.4.)

GENDER AND COLOR OF PUPPIES

| | Male | Female |
|-------|------|--------|
| Black | 1 | 2 |
| Brown | 1 | 3 |

The table above shows the gender and color of 7 puppies. If a puppy selected at random from the group is brown, what is the probability it is a male?

- A. $\frac{1}{4}$
- B. $\frac{2}{7}$
- C. $\frac{1}{3}$
- D. $\frac{1}{2}$
- E. $\frac{2}{3}$

5. Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations. For example, compare the chance of having lung cancer if you are a smoker with the chance of being a smoker if you have lung cancer. (S-CP.5.)

Use the rules of probability to compute probabilities of compound events in a uniform probability model

6. Find the conditional probability of A given B as the fraction of B's outcomes that also belong to A, and interpret the answer in terms of the model. (S-CP.6.)

7. Apply the Addition Rule, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$, and interpret the answer in terms of the model. (S-CP.7.)

8. (+) Apply the general Multiplication Rule in a uniform probability model, $P(A \text{ and } B) = P(A)P(B|A) = P(B)P(A|B)$, and interpret the answer in terms of the model. (S-CP.8.)

Bob is going on a trip. He will be taking a taxi, a flight, and then a train. Bob chose the following three companies based on their claims.

- Tom's Taxi Service claims that it is on time 95 percent of the time.
- Friendly Flyer Airlines claims that it is on time 93 percent of the time.
- Rapid Railways claims that it is on time 98 percent of the time.

Based on the three companies' claims, what is the approximate probability that all three parts of Bob's trip will be on time, assuming that all three probabilities are independent?

2009-12-7-12

Source: National Assessment of Educational Progress, 2009, Grade 12 Mathematics Assessment.

A particular flu vaccine is effective for 90 percent of the patients who receive it. What is the probability that it will be effective for both of the next two randomly selected patients?

- A. 0.54
- B. 0.63
- C. 0.72
- D. 0.81
- E. 0.90

2005-12-12-9

Source: National Assessment of Educational Progress, 2005, Grade 12 Mathematics Assessment.

9. (+) Use permutations and combinations to compute probabilities of compound events and solve problems. (S-CP.9.)

Using Probability to Make Decisions

S-MD

Calculate expected values and use them to solve problems

1. (+) Define a random variable for a quantity of interest by assigning a numerical value to each event in a sample space; graph the corresponding probability distribution using the same graphical displays as for data distributions. (S-MD.1.)

2. (+) Calculate the expected value of a random variable; interpret it as the mean of the probability distribution. (S-MD.2.)

A quarter is flipped 50 times. Which of the following is most likely to be the number of times heads comes up?

- A. 2
- B. 3
- C. 11
- D. 26
- E. 50

2008-13-21-19
2008-17-21-19

Source: National Assessment of Educational Progress, 2008, Age 13 and Age 17 Mathematics Assessments.

3. (+) Develop a probability distribution for a random variable defined for a sample space in which theoretical probabilities can be calculated; find the expected value. For example, find the theoretical probability distribution for the number of correct answers obtained by guessing on all five questions of a multiple-choice test where each question has four choices, and find the expected grade under various grading schemes. (S-MD.3.)

4. (+) Develop a probability distribution for a random variable defined for a sample space in which probabilities are assigned empirically; find the expected value. For example, find a current data distribution on the number of TV sets per household in the United States, and calculate the expected number of sets per household. How many TV sets would you expect to find in 100 randomly selected households? (S-MD.4.)

Use probability to evaluate outcomes of decisions

5. (+) Weigh the possible outcomes of a decision by assigning probabilities to payoff values and finding expected values.

- a. Find the expected payoff for a game of chance. For example, find the expected winnings from a state lottery ticket or a game at a fast-food restaurant.
- b. Evaluate and compare strategies on the basis of expected values. For example, compare a high-deductible versus a low-deductible automobile insurance policy using various, but reasonable, chances of having a minor or a major accident. (S-MD.5.)

6. (+) Use probabilities to make fair decisions (e.g., drawing by lots, using a random number generator). (S-MD.6.)

7. Analyze decisions and strategies using probability concepts (e.g., product testing, medical testing, pulling a hockey goalie at the end of a game). (S-MD.7.)