# Cumulative AP Practice Test 2

**Section I: Multiple Choice** *Choose the best answer for each question.*

**AP2.1.** The five-number summary for a data set is given by min $= 5$, $Q_1 = 18$, $M = 20$, $Q_3 = 40$, max $= 75$. If you wanted to construct a modified boxplot for the data set (that is, one that would show outliers, if any existed), what would be the maximum possible length of the right-side "whisker"?

(a) 33
(c) 45
(e) 55
(b) 35
(d) 53

**AP2.2.** The probability distribution for the number of heads in four tosses of a coin is given by

| Number of heads: | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Probability: | 0.0625 | 0.2500 | 0.3750 | 0.2500 | 0.0625 |

The probability of getting at least one *tail* in four tosses of a coin is

(a) 0.2500.
(c) 0.6875.
(e) none of these.
(b) 0.3125.
(d) 0.9375.

**AP2.3.** In a certain large population of adults, the distribution of IQ scores is strongly left-skewed with a mean of 122 and a standard deviation of 5. Suppose 200 adults are randomly selected from this population for a market research study. The distribution of the sample mean of IQ scores is

(a) left-skewed with mean 122 and standard deviation 0.35.

(b) exactly Normal with mean 122 and standard deviation 5.

(c) exactly Normal with mean 122 and standard deviation 0.35.

(d) approximately Normal with mean 122 and standard deviation 5.

(e) approximately Normal with mean 122 and standard deviation 0.35.

**AP2.4.** A 10-question multiple-choice exam offers 5 choices for each question. Jason just guesses the answers, so he has probability 1/5 of getting any one answer correct. You want to perform a simulation to determine the number of correct answers that Jason gets. One correct way to use a table of random digits to do this is the following:

(a) One digit from the random digit table simulates one answer, with 5 = right and all other digits = wrong. Ten digits from the table simulate 10 answers.

(b) One digit from the random digit table simulates one answer, with 0 or 1 = right and all other digits = wrong. Ten digits from the table simulate 10 answers.

(c) One digit from the random digit table simulates one answer, with odd = right and even = wrong. Ten digits from the table simulate 10 answers.

(d) Two digits from the random digit table simulate one answer, with 00 to 20 = right and 21 to 99 = wrong. Ten pairs of digits from the table simulate 10 answers.

(e) Two digits from the random digit table simulate one answer, with 00 to 05 = right and 06 to 99 = wrong. Ten pairs of digits from the table simulate 10 answers.

**AP2.5.** Suppose we roll a fair die four times. The probability that a 6 occurs on exactly one of the rolls is

(a) $4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^1$
(c) $4\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^3$
(e) $6\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^3$

(b) $\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^1$
(d) $\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^3$

**AP2.6.** You want to take an SRS of 50 of the 816 students who live in a dormitory on a college campus. You label the students 001 to 816 in alphabetical order. In the table of random digits, you read the entries

95592 94007 69769 33547 72450 16632 81194 14873

The first three students in your sample have labels

(a) 955, 929, 400.
(d) 929, 400, 769.
(b) 400, 769, 769.
(e) 400, 769, 335.
(c) 559, 294, 007.

**AP2.7.** The number of unbroken charcoal briquets in a twenty-pound bag filled at the factory follows a Normal distribution with a mean of 450 briquets and a standard deviation of 20 briquets. The company expects that a certain number of the bags will be underfilled, so the company will replace for free the 5% of bags that have too few briquets. What is the minimum number of unbroken briquets the bag would have to contain for the company to avoid having to replace the bag for free?

(a) 404 (c) 418 (e) 448 (b) 411 (d) 425

**AP2.8.** You work for an advertising agency that is preparing a new television commercial to appeal to women. You have been asked to design an experiment to compare the effectiveness of three versions of the commercial. Each subject will be shown one of the three versions and then asked about her attitude toward the product. You think there may be large differences between women who are employed and those who are not. Because of these differences, you should use

(a) a block design, but not a matched pairs design.

(b) a completely randomized design.

(c) a matched pairs design.

(d) a simple random sample.

(e) a stratified random sample.

**AP2.9.** Suppose that you have torn a tendon and are facing surgery to repair it. The orthopedic surgeon explains the risks to you. Infection occurs in 3% of such operations, the repair fails in 14%, and both infection and failure occur together 1% of the time. What is the probability that the operation is successful for someone who has an operation that is free from infection?

(a) 0.0767
(c) 0.8400
(e) 0.9900
(b) 0.8342
(d) 0.8660

**AP2.10.** Social scientists are interested in the association between high school graduation rate (HSGR) and the percent of U.S. families living in poverty (POV). Data were collected from all 50 states and the District of Columbia, and a regression analysis was conducted. The resulting least-squares regression line is given by $\widehat{POV} = 59.2 - 0.620(HSGR)$ with $r^2 = 0.802$. Based on the information, which of the following is the best interpretation for the slope of the least-squares regression line?

(a) For each 1% increase in the graduation rate, the percent of families living in poverty is predicted to decrease by approximately 0.896.

(b) For each 1% increase in the graduation rate, the percent of families living in poverty is predicted to decrease by approximately 0.802.

(c) For each 1% increase in the graduation rate, the percent of families living in poverty is predicted to decrease by approximately 0.620.

(d) For each 1% increase in the percent of families living in poverty, the graduation rate is predicted to increase by approximately 0.802.

(e) For each 1% increase in the percent of families living in poverty, the graduation rate is predicted to decrease by approximately 0.620.

Here is a dotplot of the adult literacy rates in 177 countries in 2008, according to the United Nations. For example, the lowest literacy rate was 23.6%, in the African country of Burkina Faso. Use the dotplot below to answer Questions AP2.11 to AP2.13.



**Literacy rate**
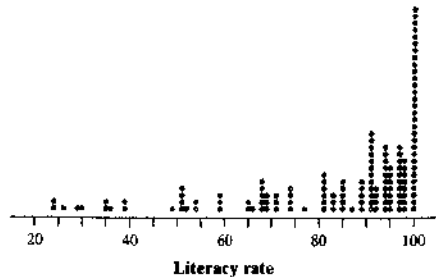
**AP2.11.** The overall shape of this distribution is

(a) clearly skewed to the right.

(b) clearly skewed to the left.

(c) roughly symmetric.

(d) uniform.

(e) There is no clear shape.

**AP2.12.** The mean of this distribution (*don't* try to find it) will be

(a) very close to the mode.

(b) greater than the median.

(c) less than the median.

(d) You can't say, because the median is random.

(e) You can't say, because the mean is random.

**AP2.13.** Based on the shape of this distribution, what numerical measures would best describe it?

(a) The five-number summary

(b) The mean and standard deviation

(c) The mean and the quartiles

(d) The median and the standard deviation

(e) It is not possible to determine which numerical values to use.

**AP2.14.** The correlation between the age and height of children under the age of 12 is found to be $r = 0.60$. Suppose we use the age $x$ of a child to predict the height $y$ of the child. What can we conclude?

(a) The height is generally 60% of a child's weight.

(b) About 60% of the time, age will accurately predict height.

(c) The fraction of the variation in heights explained by the least-squares regression line of $y$ on $x$ is 0.36.

(d) The least-squares regression line of $y$ on $x$ has a slope of 0.6.

(e) Thirty-six percent of the time, the least-squares regression line accurately predicts height.

**AP2.15.** An agronomist wants to test three different types of fertilizer (A, B, and C) on the yield of a new variety of wheat. The yield will be measured in bushels per acre. Six one-acre plots of land were randomly assigned to each of the three fertilizers. The treatment, experimental unit, and response variable are, respectively,

(a) a specific fertilizer, bushels per acre, a plot of land.

(b) a plot of land, bushels per acre, a specific fertilizer.

(c) random assignment, a plot of land, bushels per acre.

(d) a specific fertilizer, a plot of land, bushels per acre.

(e) a specific fertilizer, the agronomist, bushels per acre.

**AP2.16.** According to the U.S. census, the proportion of adults in a certain county who owned their own home was 0.71. An SRS of 100 adults in a certain section of the county found that 65 owned their home. Which one of the following represents the approximate probability of obtaining a sample of 100 adults in which fewer than 65 own their home, assuming that this section of the county has the same overall proportion of adults who own their home as does the entire county?

(a) $\binom{100}{65}(0.71)^{65}(0.29)^{35}$ (b) $\binom{100}{65}(0.29)^{65}(0.71)^{35}$

(c) $P\left(z < \dfrac{0.65 - 0.71}{\sqrt{\dfrac{(0.71)(0.29)}{100}}}\right)$ (d) $P\left(z < \dfrac{0.65 - 0.71}{\sqrt{\dfrac{(0.65)(0.35)}{100}}}\right)$

(e) $P\left(z < \dfrac{0.65 - 0.71}{\dfrac{(0.71)(0.29)}{\sqrt{100}}}\right)$

**AP2.17.** Which one of the following would be a correct interpretation if you have a $z$-score of $+2.0$ on an exam?

(a) It means that you missed two questions on the exam.

(b) It means that you got twice as many questions correct as the average student.

(c) It means that your grade was two points higher than the mean grade on this exam.

(d) It means that your grade was in the upper 2% of all grades on this exam.

(e) It means that your grade is two standard deviations above the mean for this exam.

**AP2.18.** Records from a random sample of dairy farms yielded the information below on the number of male and female calves born at various times of the day.

|         | Day | Evening | Night | Total |
|---------|-----|---------|-------|-------|
| Males   | 129 | 15      | 117   | 261   |
| Females | 118 | 18      | 116   | 252   |
| **Total** | 247 | 33    | 233   | 513   |

What is the probability that a randomly selected calf was born in the night or was a female?

(a) $\dfrac{369}{513}$ (b) $\dfrac{485}{513}$ (c) $\dfrac{116}{513}$ (d) $\dfrac{116}{252}$ (e) $\dfrac{116}{233}$

**AP2.19.** When people order books from a popular online source, they are shipped in standard-sized boxes. Suppose that the mean weight of the boxes is 1.5 pounds with a standard deviation of 0.3 pounds, the mean weight of the packing material is 0.5 pounds with a standard deviation of 0.1 pounds, and the mean weight of the books shipped is 12 pounds with a standard deviation of 3 pounds. Assuming that the weights are independent, what is the standard deviation of the total weight of the boxes that are shipped from this source?

(a) 1.84        (c) 3.02        (e) 9.10

(b) 2.60        (d) 3.40

**AP2.20.** A grocery chain runs a prize game by giving each customer a ticket that may win a prize when the box is scratched off. Printed on the ticket is a dollar value ($500, $100, $10) or the statement, "This ticket is not a winner." Monetary prizes can be redeemed for groceries at the store. Here are the distribution of the prize values and the associated probabilities for each prize:

| Amount won: | $500 | $100 | $25 | $0 |
|---|---|---|---|---|
| Probability: | 0.01 | 0.05 | 0.20 | 0.74 |

Which of the following are the mean and standard deviation, respectively, of the winnings?

(a) $15.00, $2900.00     (d) $156.25, $53.85

(b) $15.00, $53.85       (e) $156.25, $26.93

(c) $15.00, $26.93

**AP2.21.** A large company is interested in improving the efficiency of its customer service and decides to examine the length of the business phone calls made to clients by its sales staff. A cumulative relative frequency graph is shown below from data collected over the past year. According to the graph, the shortest 80% of calls will take how long to complete?



Length of Phone Calls
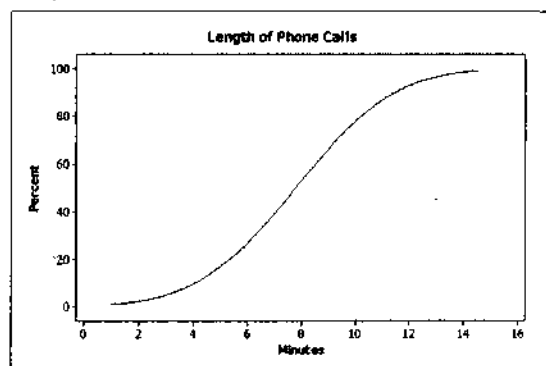
(a) Less than 10 minutes.

(b) At least 10 minutes.

(c) Exactly 10 minutes.

(d) At least 5.5 minutes.

(e) Less than 5.5 minutes.

## Section II: Free Response

**AP2.22.** A health worker is interested in determining if omega-3 fish oil can help reduce cholesterol in adults. She obtains permission to examine the health records of 200 people in a large medical clinic and classifies them according to whether or not they take omega-3 fish oil. She also obtains their latest cholesterol readings and finds that the mean cholesterol reading for those who are taking omega-3 fish oil is 18 points lower than the mean for the group not taking omega-3 fish oil.
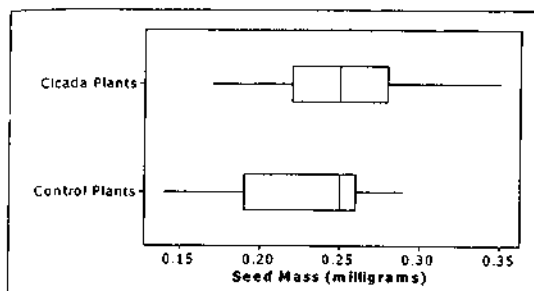
(a) Is this an observational study or an experiment? Explain.

(b) Do these results provide convincing evidence that taking omega-3 fish oil lowers cholesterol?

(c) Explain the concept of confounding in the context of this study and give one example of a possible confounding variable.

**AP2.23.** There are four major blood types in humans: O, A, B, and AB. In a study conducted using blood specimens from the Blood Bank of Hawaii, individuals were classified according to blood type and ethnic group. The ethnic groups were Hawaiian, Hawaiian-White, Hawaiian-Chinese, and White. Suppose that a blood bank specimen is selected at random.

| Blood type | Ethnic Group | | | |
|---|---|---|---|---|
| | Hawaiians | Hawaiian-White | Hawaiian-Chinese | White |
| O | 1903 | 4469 | 2206 | 53,759 |
| A | 2490 | 4671 | 2368 | 50,008 |
| B | 178 | 606 | 568 | 16,252 |
| AB | 99 | 236 | 243 | 5001 |

(a) Find the probability that the specimen contains type O blood or comes from the Hawaiian-Chinese ethnic group. Show your work.

(b) What is the probability that the specimen contains type AB blood, given that it comes from the Hawaiian ethnic group? Show your work.

(c) Are the events "type B blood" and "Hawaiian ethnic group" independent? Give appropriate statistical evidence to support your answer.

Now suppose that two blood bank specimens are selected at random.

(d) Find the probability that at least one of the specimens contains type A blood from the White ethnic group.

**AP2.24.** Every 17 years, swarms of cicadas emerge from the ground in the eastern United States, live for about six weeks, and then die. (There are several different "broods," so we experience cicada eruptions more often than every 17 years.) There are so many cicadas that their dead bodies can serve as fertilizer and increase plant growth. In a study, a researcher added 10 cicadas under 39 randomly selected plants in a natural plot of American bellflowers on the forest floor, leaving other plants undisturbed. One of the response variables measured was the size of seeds produced by the plants. Here are the boxplots and summary statistics of seed mass (in milligrams) for 39 cicada plants and 33 undisturbed (control) plants:



Seed Mass (milligrams)

| Variable: | n | Minimum | $Q_1$ | Median | $Q_3$ | Maximum |
|---|---|---|---|---|---|---|
| Cicada plants: | 39 | 0.17 | 0.22 | 0.25 | 0.28 | 0.35 |
| Control plants: | 33 | 0.14 | 0.19 | 0.25 | 0.26 | 0.29 |

(a) Is this an observational study or an experiment? Explain.

(b) Based on the graphical displays, which distribution has the larger mean? Justify your answer.

(c) Do the data support the idea that dead cicadas can serve as fertilizer? Give graphical and numerical evidence to support your conclusion.