VEPP BIOSTATISTICS MIDTERM NOV 6, 2012
Choose one correct answer and record it on the last sheet of this exam.
Commonly used z values: 95% =1.96, 90% =1.645, 99% = 2.576
*indicates correct option

1. A survey is conducted on students taking a statistics class. Several variables are measured in the survey. Which of these variables listed below is not quantitative?
    a. The number of credit hours taken during the quarter
*   b. The gender
    c. The parents' annual income
    d. The high school GPA

2. Stream quality is based on the levels of many variables, including the following. Which of these variables is quantitative?
    a. The amount of phosphorus
    b. The number of distinct species present
    c. The amount of dissolved oxygen
*   d. All of the above

3. A survey records many variables of interest to the researchers conducting the survey. The following are some of the variables from a survey conducted by the USDA. Which of these variables is categorical?
*   a. The type of grain grown
    b. The number of acres owned
    c. The total farm income, before taxes, in 1993
    d. The number of children under the age of 12 on the farm

4. Which of the following is a discrete variable?
*   a. Number of toxins present in a fish
    b. Length of a fish
    c. Weight of a fish
    d. None of the above

Use the following to answer Questions 5-9.
The Insurance Institute for Highway Safety publishes data on the total damage caused by compact automobiles in a series of controlled, low-speed collisions. The following costs are for a sample of 9 cars, in hundreds of dollars.

    10.0   6.0   8.0   10.0   4.0   3.5   7.5   8.0   9.0

5. What is the mean of the total damage suffered for this sample of cars?
    a. $239
*   b. $733
    c. $800
    d. $950

6. Using the correct units, what is the value of the variance?
    a. 224.85 dollars
    b. 238.48 dollars^2
    c. 50,555.56 dollars^2
*   d. 56,875 dollars^2

7. What is the first quartile for these data?
    a. $350
*   b. $500
    c. $600
    d. $800

8. What is interquartile range for these data?
    a. $300
    b. $350
    c. $400
*   d. $450

9. To be an outlier, an observation must fall outside the range
    a. 0.50 to 14.00.
    b. 5.00 to 9.50.
    c. 3.50 to 10.00.
*   d. −1.75 to 16.25.

10. Sickle-cell disease is a painful disorder of the red blood cells that in the United States affects mostly African-Americans. To investigate whether the drug hydroxyurea can reduce the pain associated with sickle-cell disease, a study by NIH gave the drug to 150 sickle-cell sufferers and the placebo to another 150. Neither doctors nor patients were told who received the drug. The number of episodes of pain reported by each subject was recorded. This is an example of
    a. an observational study.
*   b. an experiment.
    c. convenience sampling.
    d. multistage sampling.

11. A health club is interested in finding out which of two brands of aerobic exercise equipment provides a more vigorous workout. They purchase 10 machines of each type and, for five days between 9 a.m. and 12 p.m., they measure the average pulse rate of each person who is working out on one of these 20 machines. This is an example of
    a. an experiment.
*   b. an observational study.
    c. a paired data experiment.
    d. a stratified experiment.

12. At a local health club, a researcher samples 75 people whose primary exercise is cardiovascular and 75 people whose primary exercise is strength training. The researcher's objective is to assess the effect of type of exercise on cholesterol. Each subject reported to a clinic to have his or her cholesterol measured. The subjects were unaware of the purpose of the study, and the technician measuring the cholesterol was not aware of the subject's type of exercise. This is
*   a. an observational study.
    b. an experiment, but not a double blind experiment.
    c. a double blind experiment.
    d. a matched pairs experiment.

Use the following to answer Questions 13-15.
A researcher is interested in investigating the relationship between sugar consumption and weight gain for high school students. Fifteen volunteers were randomly assigned to one of two groups. The first group contained six volunteers who were put on a low-sugar diet. The second group consisting of the remaining nine volunteers was put on a diet with sugar constituting approximately 15% of their diet. After eight weeks, the change in weight was recorded for each of the volunteers.

13. The response is
    a. the percentage of sugar in the diet.
    b. the eight-week time period.
    c. the assignment to groups.
*   d. the change in weight.

14. This is an example of
    a. a randomized observational study.
    b. a matched pairs experiment.
    c. an experiment, but not a double blind experiment.
*   d. a randomized comparative experiment.

15. This study uses the principles of
*   a. randomization.
    b. confounding.
    c. blocking.
    d. All of the above

Use the following to answer Questions 16-17.
The Physicians Health Study followed 22,000 male physicians for a period of several years. About 11,000 took an aspirin every second day, while the rest took a placebo. At the completion of the study, it was noted whether a subject had

experienced a heart attack during the period of the study. It was found that the aspirin group had significantly fewer heart attacks than the placebo group.

16. The factor in the experiment is
    a. the severity of the heart attack.
    b. use of a placebo.
*   c. medication used (aspirin or placebo).
    d. the length of the study.

17. The response variable in this experiment is
*   a. whether a heart attack occurred.
    b. the placebo effect.
    c. whether the symptoms lessened.
    d. the length of the interview.

Use the following to answer Questions 18-19.
Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers give the new medication to 50 adult volunteers who suffer from allergies. Forty-four of these volunteers report a significant reduction in their allergy symptoms without any drowsiness.

18. This study could be improved by
    a. including people who do not suffer from allergies in the study in order to represent a more diverse population.
    b. repeating the study using only the volunteers who reported a significant reduction in their allergy symptoms without
       any drowsiness and giving them a higher dosage this time.
 * c. using a control group.
    d. All of the above

19. The experimental units are the
    a. medication taken by the 50 adult volunteers.
*   b. 50 adult volunteers.
    c. 44 volunteers who reported a significant reduction in their symptoms without any drowsiness.
    d. 6 volunteers who didn't report a significant reduction in their symptoms without any drowsiness.

20. A researcher is trying to determine the proportion of a certain species of fish in a local lake. After sampling 40 fish, she found 32 of them were the species of interest. She estimates the probability that the next fish is of the species of interest to be
    a. 0.32.
    b. 0.50.
*   c. 0.80.
    d. 1.25.

21. A researcher in art education found that 17% of all adult Americans had taken at least one dance course in their life. If we select an adult American at random, we would estimate the probability that the person selected has taken at least one dance class in their life to be
*   a. 0.17.
    b. 0.83.
    c. 17.
    d. $17 \times 100 = 1700$.

Use the following information for Questions 22-24.
Event A has probability 0.4. Event B has probability 0.5.
22. If A and B are disjoint (mutually exclusive), then the probability that both events occur is
*   a. 0.0.
    b. 0.1.
    c. 0.2.
    d. 0.9.

23. If A and B are disjoint, then the probability of A or B is
    a. 0.0.
    b. 0.1.
    c. 0.2.
*   d. 0.9.

24. An assignment of probability must obey which of the following?
    a. The probability of any event must be a number between 0 and 1, inclusive.
    b. The sum of all the probabilities of all outcomes in the sample space must be exactly 1.
    c. It must obey the addition rule for disjoint events.
*   d. All of the above

25. I flip a coin twice and count the number of heads. Which of the following is a valid assignment of probabilities for the number of heads observed in two flips?
    a.

Number of heads	0	1	2
Probability	1/3	1/3	1/3

    b.

Number of heads	0	1	2
Probability	1/4	1/4	1/4

*   c.

Number of heads	0	1	2
Probability	1/4	1/2	1/4

    d. All of the above

Use the following to answer questions 26-28.
If you draw an M&M candy (a kind of chocolate in different colors) at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made. The following table gives the probability that a randomly chosen M&M has each color. (These probabilities came from the M&M Web site.)

Color	Brown	Red	Yellow	Green	Blue	Orange
Probability	0.13	0.13	?	0.16	0.24	0.20

26. The probability of drawing a yellow candy is
*   a. 0.14.
    b. 0.16.
    c. 1/6.
    d. impossible to determine from the information given.

27. The probability that I do not draw a red candy is
    a. 0.13.
    b. 0.27.
*   c. 0.87.
    d. 5/6.

28. The probability that I draw either a brown or a green candy is
    a. 0.13.
    b. 0.16.
*   c. 0.29.
    d. 0.56.

Use the following to answer questions 29-30.
All human blood can be typed as one of O, A, B, or AB. The distribution of the types varies a bit with race. Choose an African-American at random. Here are the approximate probabilities that the person you choose will have blood type O, B, or AB.

Blood Type	O	A	B	AB
Probability	0.50	?	0.20	0.05

29. The probability that the person chosen has blood type A is
    a. 0.04.
*   b. 0.25.
    c. 0.27.
    d. impossible to determine from the information given.

30. The probability that the person chosen has a blood type other than O is
    a. 0.25.
*   b. 0.50.
    c. 0.75.
    d. 1.00.

Use the following to answer Questions 31–32.
According to the Center for Disease Control and Prevention (CDC), up to 20% of Americans contract the influenza virus each year, and approximately 3% of all births in the United States result in birth defects each year. Consider two babies being born independently of one another.

31. The probability that both babies have birth defects is
*   a. 0.0009.
    b. 0.0400.
    c. 0.0606.
    d. 0.2000.

32. The probability that neither baby catches the flu in a given year is
    a. 0.024.
    b. 0.040.
    c. 0.230.
*   d. 0.640.

33. Event A occurs with probability 0.1. Event B occurs with probability 0.6. If A and B are independent, then
*   a. $P(A \text{ and } B) = 0.06$.
    b. $P(A \text{ or } B) = 0.70$.
    c. $P(A \text{ and } B) = 0.70$.
    d. $P(A \text{ or } B) = 0.06$.

34. A Normal density curve has which of the following properties?
    a. It is symmetric.
    b. It has a peak centered above its mean.
    c. The spread of the curve is proportional to the standard deviation.
*   d. All of the above

35. Scores on a university exam are Normally distributed with a mean of 68 and a standard deviation of 9. Using the 68-95-99.7 rule, what percent of students score above 77?
    a. 2.5%
    b. 5%
*   c. 16%
    d. 32%

36. The time to complete a standardized exam is approximately Normal with a mean of 70 minutes and a standard deviation of 10 minutes. Using the 68-95-99.7 rule, if students are given 90 minutes to complete the exam, what percent of students will not finish?
    a. 32%
    b. 5%
*   c. 2.5%
    d. 0.0015%

Use the following information for Questions 37-39.
The proportion of American births that result in a birth defect is approximately 1/33 according to the Center for Disease Control and Prevention (CDC). A local hospital randomly selects five births and lets the random variable X count the number not resulting in a defect. Assume the births are independent.

37. X follows a
    a. normal distribution with mean 5 and variance 1/33.
    b. binomial distribution with $n = 5$ and $p = 1/33$.
*   c. binomial distribution with $n = 5$ and $p = 32/33$.
    d. None of the above

38. What is the probability that two of the five births do not result in defects?
*   a. 0.0003
    b. 0.0008
    c. 0.0084
    d. 0.3125

39. A small school club has 16 students with 12 males and 4 females. Two representatives are needed to meet with the principal. The names of the 16 students are put in a hat, and 2 are selected at random to represent the club. Let X be the number of males selected. Then X, has
    a. a binomial distribution with mean 0.75.
    b. a binomial distribution with mean 1.5.
    c. a binomial distribution with 12 trials.
    *d. None of the above
40. If X has a binomial distribution with 20 trials and a mean of 5, then the success probability p is
    *a. 0.25.
    b. 0.50.
    c. 0.75.
    d. not known without taking a sample.

41. Sale of eggs that are contaminated with salmonella can cause food poisoning among consumers. A large egg producer takes a SRS of 200 eggs from all the eggs shipped in one day. The laboratory reports that 11 of these eggs had salmonella contamination. Unknown to the producer, 0.2% (two-tenths of one percent) of all eggs shipped had salmonella. In this situation
*   a. 0.2% is a parameter and 11 is a statistic.
    b. 11 is a parameter and 0.2% is a statistic.
    c. both 0.2% and 11 are statistics.
    d. both 0.2 % and 11 are parameters.

42. The law of large numbers states that as the number of observations drawn at random from a population—with finite mean μ and variance σ^2—increases, the mean, $\bar{x}$, of the observed values
a. tends to get larger.
b. tends to get smaller.
*c. tends to get closer to the population mean μ.
d. All of the above

43. The fact that the sample mean doesn't tend to over or under estimate the population mean makes the sample mean
    a. consistent.
    *b. unbiased.
    c. efficient.
    d. a statistic.

44. I flip a coin ten times and record the proportion of heads I obtain. I then repeat this process of flipping the coin ten times and recording the proportion of heads obtained many, many times. When done, I make a histogram of my results. This histogram represents
a. the bias, if any, which is present.
b. the true population parameter.
c. simple random sampling.
*d. the sampling distribution of the proportion of heads in ten flips of the coin.

45. The variability of a statistic is described by
*a. the spread of its sampling distribution.
b. the amount of bias present.
c. the vagueness in the wording of the question used to collect the sample data.
d. the stability of the population it describes.

ANSWER SHEET. Circle your correct answer here. Return this sheet only. Keep the exam. Good luck.

NAME: INSTITUTION:

1	a	b	c	d
2	a	b	c	d
3	a	b	c	d
4	a	b	c	d
5	a	b	c	d
6	a	b	c	d
7	a	b	c	d
8	a	b	c	d
9	a	b	c	d
10	a	b	c	d
11	a	b	c	d
12	a	b	c	d
13	a	b	c	d
14	a	b	c	d
15	a	b	c	d
16	a	b	c	d
17	a	b	c	d
18	a	b	c	d
19	a	b	c	d
20	a	b	c	d
21	a	b	c	d
22	a	b	c	d
23	a	b	c	d
24	a	b	c	d
25	a	b	c	d
26	a	b	c	d
27	a	b	c	d
28	a	b	c	d
29	a	b	c	d
30	a	b	c	d
31	a	b	c	d
32	a	b	c	d
33	a	b	c	d
34	a	b	c	d
35	a	b	c	d
36	a	b	c	d
37	a	b	c	d
38	a	b	c	d
39	a	b	c	d
40	a	b	c	d
41	a	b	c	d
42	a	b	c	d
43	a	b	c	d
44	a	b	c	d

45   a b c d