1. In a study of the effects of acid rain, a random sample of 100 trees from a particular forest is examined. Forty percent of these show some signs of damage. Which of the following statements is correct?
   (a) 40% is a parameter
   (b) 40% is a statistic
   (c) 40% of all trees in the forest show some signs of damage
   (d) More than 40% of the trees in the forest show some signs of damage
   (e) Less than 40% of the trees in the forest show some signs of damage

2. Refer to the previous problem. Which of the following statements is correct?
   (a) The sampling distribution of the proportion of damaged trees is approximately normal
   (b) If we took another random sample of trees, we would find that 40% of these would show some signs of damage
   (c) If a sample of 1000 trees was examined, the variability of the sample proportion would be larger than in a sample of 100 trees
   (d) This is a comparative experiment
   (e) None of the above

3. A machine is designed to fill 16-ounce bottles of shampoo. When the machine is working properly, the mean amount poured into the bottles is 16.05 ounces with a standard deviation of 0.05 ounce. If four bottles are randomly selected each hour and the number of ounces in each bottle is measured, then 95% of the observations should occur in which interval?
   (a) 16.05 and 16.15 ounces
   (b) –.30 and +.30 ounces
   (c) 15.95 and 16.15 ounces
   (d) 15.90 and 16.20 ounces
   (e) None of the above

4. A survey asks a random sample of 1500 adults in Ohio if they support an increase in the state sales tax from 5% to 6%, with the additional revenue going to education. Let \( X \) denote the number in the sample that say they support the increase. Suppose that 40% of all adults in Ohio support the increase. The probability that \( X \) is more than 650 is
   (a) less than 0.0001.
   (b) less than 0.001.
   (c) less than 0.01.
   (d) 0.9960.
   (e) None of these.

5. The probability that a certain machine will produce a defective item is 0.20. If a random sample of 6 items is taken from the output of this machine, what is the probability that there will be 5 or more defectives in the sample?
   (a) 0.0001
   (b) 0.0154
   (c) 0.0015
   (d) 0.2458
   (e) 0.0016

6. A professional basketball player sinks 80% of his foul shots, in the long run. If he gets 100 tries during a season, then the probability that he sinks between 75 and 90 shots (inclusive) is approximately equal to:
   (a) \( \Pr (-1.25 \leq Z \leq 2.5) \)
   (b) \( \Pr (-1.125 \leq Z \leq 2.625) \)
   (c) \( \Pr (-1.125 \leq Z \leq 2.375) \)
   (d) \( \Pr (-1.375 \leq Z \leq 2.375) \)
   (e) \( \Pr (-1.375 \leq Z \leq 2.625) \)

7. Suppose we are planning on taking an SRS from a population. If we double the sample size, then \( \sigma_x \) will be multiplied by:
   (a) \( \sqrt{2} \)
   (b) \( 1/\sqrt{2} \)
   (c) 2
   (d) \( 1/2 \)
   (e) 4
8. If a statistic used to estimate a parameter is such that the mean of its sampling distribution is different from the true value of the parameter being estimated, the statistic is said to be
   (a) Random   (b) Biased
   (c) A proportion   (d) not biased
   (b) None of the above. The answer is ____________________________.

9. Dr. Stats plans to toss a fair coin 10,000 times in the hope that it will lead him to a deeper understanding of the laws of probability. Which of the following statements is true?
   (a) It is unlikely that Dr. Stats will get more than 5000 heads.
   (b) Whenever Dr. Stats gets a string of 15 tails in a row, it becomes more likely that the next toss will be a head.
   (c) The fraction of tosses resulting in heads should be close to 1/2.
   (d) The chance that the 100th toss will be a head depends somewhat on the results of the first 99 tosses.
   (e) All of the above statements are true.

10. It has been estimated that about 30% of frozen chickens contain enough salmonella bacteria to cause illness if improperly cooked. A consumer purchases 12 frozen chickens. What is the probability that the consumer will have more than 6 contaminated chickens?
   (a) 0.961   (b) 0.118   (c) 0.882   (d) 0.039   (e) 0.079

11. Refer to the previous question. Suppose that a supermarket buys 1000 frozen chickens from a supplier. The number of frozen chickens that may be contaminated that are within two standard deviations of the mean is between A and B. The numbers A and B are
   (a) (90, 510)   (b) (290.8, 309.2)   (c) (0, 730)   (d) (271, 329)   (e) (255, 345)

12. A fair coin (one for which both the probability of heads and the probability of tails are 0.5) is tossed six times. Use the binomial formula to evaluate the probability that less than 1/3 of the tosses are heads is
   (a) 0.344.   (b) 0.33.   (c) 0.109.   (d) 0.09.   (e) 0.0043.

13. Which of the following statements is (are) true?
   I. The sampling distribution of \( \bar{x} \) has standard deviation \( \sigma/\sqrt{n} \) even if the population is not normally distributed.
   II. The sampling distribution of \( \bar{x} \) is normal if the population has a normal distribution.
   III. When n is large, the sampling distribution of \( \bar{x} \) is approximately normal even if the population is not normally distributed.
   (a) I and II   (b) I and III   (c) II and III   (d) I, II, and III
   (b) None of the above gives the correct set of responses.

14. Suppose we select an SRS of size \( n = 100 \) from a large population having proportion \( p \) of successes. Let \( X \) be the number of successes in the sample. For which value of \( p \) would it be safe to assume the sampling distribution of \( X \) is approximately normal?
   (a) 0.01   (b) \( \frac{1}{9} \)   (c) 0.975   (d) 0.9999   (e) All of these.
For questions 15-17: In a test for ESP (extrasensory perception), the experimenter looks at cards that are hidden from the subject. Each card contains either a star, a circle, a wave, or a square. As the experimenter looks at each of 20 cards in turn, the subject names the shape on the card. When the ESP study described discovers a subject whose performance appears to be better than guessing, the study continues at greater length. The experimenter looks at many cards bearing one of five shapes (star, square, circle, wave, and cross) in an order determined by random numbers. The subject cannot see the experimenter as he looks at each card in turn, in order to avoid any possible nonverbal clues. The answers of a subject who does not have ESP should be independent observations, each with probability 1/5 of success. We record 1000 attempts.

15. What are the mean and standard deviation of the proportion of successes among the 1000 attempts?

\[
\text{mean: } p = 0.2 \quad \text{Since } 10(1000) < N \quad \text{All attempts possible}
\]
\[
\text{standard deviation: } \sqrt{\frac{0.2(0.8)}{1000}} = 0.0126
\]

16. What is the probability that a subject without ESP will be successful in at least 24% of the 1000 attempts?

\[
P(\hat{p} \geq 0.24) = P \left( z \geq \frac{0.24 - 0.2}{0.0126} \right) = P(z \geq 3.17) = 0.00076
\]

Or \( \text{normalcdf}(0.24, 100000000, 0.2, 0.0126) = 0.00075 \)

17. The researcher considers evidence of ESP to be a proportion of successes so large that there is only probability 0.01 that a subject could do this well or better by guessing. What proportion of successes must a subject have to meet this standard?

\[
1 - 0.01 = 0.99 \quad \text{(looking for the top right of normal curve)}
\]
\[
\text{invNorm}(0.99, 0.2, 0.0126) = 0.2293
\]

or: find the z-score of 0.99 (chart or invnorm), then use z-score formula to find \( \hat{p} \).

\[
2.33 = \frac{\hat{p} - 0.2}{0.0126} \rightarrow \hat{p} = 0.229
\]

For questions 18-19: Harlan comes to class one day, totally unprepared for a pop quiz consisting of ten multiple-choice questions. Each question has five answer choices, and Harlan answers each question randomly.

18. Find the probability that Harlan guesses more answers correctly than would be expected by chance.

\[
\text{Expect to get correct} = E(x) = np = 10(0.2) = 2
\]
\[
P(x > 2) = \binom{10}{3}(0.2)^3(0.8)^7 + \binom{10}{4}(0.2)^4(0.8)^6 + \cdots + \binom{10}{10}(0.2)^3(0.8)^7 = 0.3222
\]
\[
P(x > 2) = 1 - P(x \leq 2) = 1 - \text{binomcdf}(10,0.2,2) = 0.3222
\]

19. Find the probability that Harlan’s first correct answer occurs on or after the fourth question.

\[
P(x \geq 4) = P(x > 3) = (1 - p)^n = (0.8)^3 = 0.512
\]
20. The Internal Revenue Service estimates that 8% of all taxpayers filling out long forms make mistakes. Suppose that a random sample of 10,000 forms is selected. What is the approximate probability that more than 800 forms have mistakes?

\[
p = 0.08 \quad P(x > 800) = 1 - P(x \leq 800) = 1 - \text{binomcdf}(10000, 0.08, 800) = 0.4906
\]

For questions 21-25: Amarillo Slim, a professional dart player, has an 80% chance of hitting the bullseye on a dartboard with any throw. Suppose that he throws 10 darts, one at a time, at the dartboard.

21. Find the probability that Slim hits the bullseye exactly six times.

\[
P(X = 6) = \text{binompdf}(10, 0.8, 6) = 0.0881
\]

22. Find the probability that he hits the bullseye at least four times.

\[
P(x \geq 4) = 1 - P(x \leq 3) = 1 - \text{binomcdf}(10, 0.8, 3) = 0.9991
\]

23. Compute the mean and variance of the number of bullseyes in 10 throws.

\[
\mu_X = 10(0.8) = 8 \\
\sigma_X^2 = 10(0.8)(1 - 0.8) = 1.6
\]

24. Find the probability that Slim’s first bullseye occurs on the fourth throw.

\[
P(X = 4) = \text{geometpdf}(0.8, 4) = 0.0064
\]

25. Find the probability that it takes Amarillo more than 2 throws to hit the bullseye.

\[
P(X > 2) = (1 - 0.8)^2 = 0.04
\]

For questions 26-29: Sheila’s doctor is concerned that she may suffer from gestational diabetes (high blood glucose levels during pregnancy). There is variation both in the actual glucose level and in the blood test that measures the level. A patient is classified as having gestational diabetes if the glucose level is above 140 milligrams per deciliter (mg/dl) one hour after a sugary drink is ingested. Sheila’s measured glucose level one hour after ingesting the sugary drink varies according to the normal distribution with \( \mu = 125 \) mg/dl and \( \sigma = 10 \) mg/dl.

26. If a single glucose measurement is made, what is the probability that Sheila is diagnosed as having gestational diabetes?

\[
P(x > 140) = P(z > \frac{140 - 125}{10}) = P(z > 1.5) = 0.0668
\]

27. If measurements are made instead on four separate days and the mean result is computed, describe the center, shape and spread of the resulting sampling distribution.

center: \( \mu_X = 125 \) 
shape: normal (since parent population is normal)
spread: \( \sigma_X = \frac{10}{\sqrt{4}} = 5. 
\]

28. If the sample mean of Sheila’s four readings is compared with the criterion 140 mg/dl, what is the probability that Sheila is diagnosed as having gestational diabetes?

\[
P(\bar{x} \geq 140) = P\left(z \geq \frac{140 - 125}{5}\right) = P(z \geq 3) = 0.0013
\]

Or \( \text{normalcdf}(140, 100000, 125, 5) = 0.0013 \)
29. What is the level $L$ such that there is probability only 0.05 that the mean glucose level of four test results falls above $L$ for Sheila’s glucose level distribution?

(Falls above) so the area under the normal curve is $1 - .05 = .95$

$\text{Invnorm (.95, 0, 1)} = 1.645$ is z-score

$\text{or } \text{InvNorm(.95, 125, 5)} = 133.22$

$1.645 = \frac{\bar{x} - 125}{5} \rightarrow \bar{x} = 133.23$

30. A study of the health of teenagers plans to measure the blood cholesterol level of an SRS of youth of ages 13 to 16 years. The researchers will report the mean $\bar{x}$ from their sample as an estimate of the mean cholesterol level $\mu$ in this population. Explain to someone who knows no statistics what it means to say that $\bar{x}$ is an “unbiased” estimator of $\mu$.

In repeated samples, the values of $\bar{x}$ will center around the true mean $\mu$. 