Probability (with Solutions)

Please read Chapter 6 of DSS and answer the following questions (each counts for 2.5 points):

1. According to the frequentist interpretation of probability, the probability of an event is the proportion of its occurrence among infinitely many identical trials. For example, if we flip a coin 1 million times and 300,000 of those times we get heads, what's the probability of heads?

   A. it is impossible to know with the information given
   B. 30%
   C. 0.3%
   D. 300%

   Answer: B (Note: 300,000/1,000,000= 0.3; 0.3*100=30%)

2. According to the Bayesian interpretation, probabilities represent one's subjective beliefs about the relative likelihood of events. For example, if we state that the probability of rain today is of 0%, we are describing that:

   A. the proportion of rain events over multiple days
   B. we are certain that it will rain today
   C. we are certain that it won't rain today
   D. none of the above

   Answer: C

3. In this class, we distinguish between two types of numeric random variables based on the number of values the variables can take. These are:

   A. treatment and control
   B. predictors and outcomes
   C. binary and non-binary
   D. treatment and outcome

   Answer: C (Note: Binary variables can only take two values; non-binary variables can take more than two values. Treatment variables, outcome variables, and predictors are distinctions based on the role the variables take in the research question, not on the number of values the variables can take.)

4. Each random variable has a probability distribution, which characterizes the likelihood of each value the variable can take. The Bernoulli distribution is the probability distribution of:

A. a non-binary variable
B. a binary variable
C. a variable that can take more than two values
D. none of the above

Answer: B

5. If X={1,1,1,0,0}, X is:

A. a non-binary variable
B. a binary variable
C. an outcome variable
D. a predictor

Answer: B

6. If X={1,1,1,0,0}, what is $P(X=1)$, also known as p?

A. $P(X=1) = p = 0.6$
B. $P(X=1) = p = 3$
C. $P(X=1) = p = 0.4$
D. none of the above

Answer: A (Note: 3/5=0.6, which means that the probability that X equals 1 is of 60%.)

7. If in a Bernoulli distribution $p=0.7$, what is $P(X=0)$?

A. $P(X=0) = 0.2$
B. $P(X=0) = 0.7$
C. $P(X=0) = 1$
D. $P(X=0) = 0.3$

Answer: D (Note: Since all probabilities in a distribution must add up to 1, $P(X=1)+P(X=0)=1$. Therefore, if $P(X=1)=p$, then $P(X=0)=1-p$. In this case, $P(X=0)=1-p=1-0.7=0.3$, which means that the probability that X equals 0 is of 30%.)

8. The Bernoulli distribution is characterized by:

A. one parameter
B. two parameters
C. three parameters
D. four parameters

Answer: A (Note: The Bernoulli distribution is characterized by one parameter: p. Once we know p, we know everything there is to know about the distribution. The probability that X equals 1 is p and the probability that X equals 0 is 1-p. The mean of the distribution is p and the variance of the distribution is $p(1-p)$.)

9. The normal distribution is the probability distribution we commonly use as a good approximation for many non-binary variables. The normal distribution is characterized by:

   A. one parameter
   B. two parameters
   C. three parameters
   D. four parameters

   Answer: B (Note: The normal distribution is characterized by two parameters: μ (pronounced mu) and σ^2 (pronounced sigma-squared). The mean of the distribution is μ and the variance of the distribution is σ^2.)
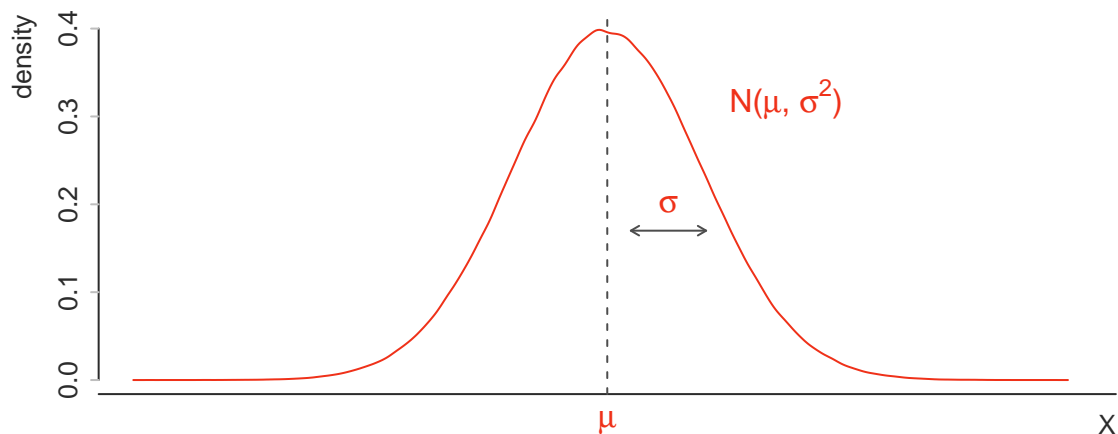
10. In mathematical notation, we write a normal random variable X as $X \sim N(\mu, \sigma^2)$, where:

    A. μ stands for the mean of the distribution and σ^2 stands for the variance
    B. μ stands for the mean of the distribution and σ^2 stands for the standard deviation
    C. μ stands for the variance of the distribution and σ^2 stands for the mean
    D. none of the above

    Answer: A

11. If $X \sim N(20, 25)$, then X is:

    A. distributed like a normal distribution with mean 25 and variance 20.
    B. distributed like a normal distribution with mean 20 and standard deviation 5.
    C. distributed like a normal distribution with mean 20 and standard deviation 25.
    D. distributed like a Bernoulli distribution with mean 20 and variance 5.

Answer: B (Note: $X \sim N(\mu, \sigma^2)$ means that X follows a normal distribution with mean μ and variance σ^2. In addition, recall that the standard deviation is the square root of the variance. Thus, X has a mean of 20, a variance of 25, and a standard deviation of 5.)

12. The probability density function of the normal distribution represents the likelihood of each possible value the normal random variable can take. We can use it to compute the probability that X takes a value within a given range:

$$P(x_1 \leq X \leq x_2) = \text{area under the curve between } x_1 \text{ and } x_2$$

If *X* follows the probability density function below, which of the following statements is true?
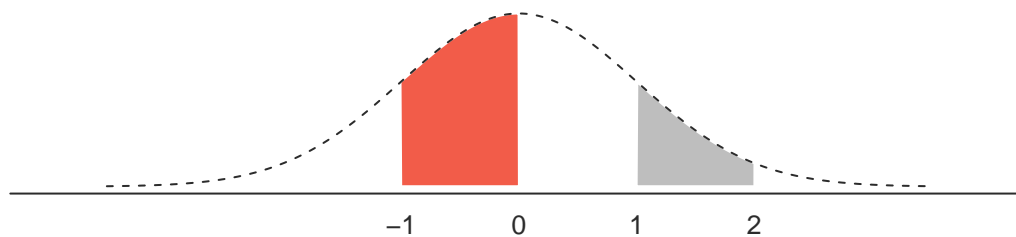
A. $P(-1 \leq X \leq 0) < P(1 \leq X \leq 2)$
B. $P(-1 \leq X \leq 0) > P(1 \leq X \leq 2)$
C. $P(-1 \leq X \leq 0) = P(1 \leq X \leq 2)$
D. none of the above

Answer: B (Note: The area under the curve between –1 and 0 is larger than the area under the curve between 1 and 0.)

13. The standard normal distribution, is a special case of the normal distribution. In mathematical notation, we refer to the standard normal random variable as *Z*. Which of the following statements is true?

A. *Z* is a normal random variable with mean 0 and variance 1
B. *Z* is a normal random variable with mean 0 and standard deviation 1
C. $Z \sim N(0, 1)$
D. all of the above

Answer: D

14. In the standard normal distribution:

A. about 95% of the observations are between –1.96 and 1.96
B. $P(Z \leq -1.96) = P(Z \geq 1.96)$
C. both A. and B.
D. neither A. nor B.

Answer: C

15. Given the properties of normal distributions, we can always transform a normally distributed random variable into a random variable that is distributed like the standard normal distribution:

$$\text{if } X \sim N(\mu, \sigma^2), \text{ then } \frac{X - \mu}{\sigma} \sim N(0,1)$$

So, if $X \sim N(10, 25)$, then:

A. $(X - 10)/25 \sim N(0,1)$
B. $(X - 10)/5 \sim N(0,1)$
C. $(X - 25)/10 \sim N(0,1)$
D. $(X - 5)/10 \sim N(0,1)$

Answer: B (Note that what we use in the denominator is the standard deviation σ, not the variance σ^2. In this case, $\sigma^2 = 25$, therefore $\sigma = \sqrt{25} = 5$.)

16. When we analyze data, we are usually interested in the value of a parameter at the population level, such as the proportion of candidate A supporters among all voters in a country. However, we typically only have access to statistics from a small sample of observations drawn from the target population, such as the proportion of supporters among the voters who responded to a survey. Which of the following statements is true:

A. This is not a problem since the sample statistics always equal the population parameters
B. This is a problem since the sample statistics typically differ from the population parameters because the sample contains noise
C. This is not a problem when we use random sampling to draw the sample. Under these circumstances, the sample statistics equal the population parameters
D. All of the above

Answer: B

17. Sampling variability refers to the fact that the value of a statistic varies from one sample to another because each sample contains a different set of observations drawn from the target population. Which of the following statements is true?

A. Smaller sample size generally leads to greater sampling variability
B. Sampling variability does not depend on the size of the sample
C. Larger sample size generally leads to greater sampling variability
D. None of the above

Answer: A

18. The Law of Large Numbers states that as the sample size increases, the sample mean of X approximates the population mean of X. Now, suppose we drew three samples of data from the same target population: the first sample with ten observations (n=10), the second sample with a thousand observations (n=1,000), and the third sample with a million observations (n=1,000,000). Which sample is most likely to provide us with a sample mean closest to the population mean of X?

A. The first sample
B. The second sample
C. The third sample
D. It is impossible to know with the information given

Answer: C

19. The Central Limit Theorem states that as the sample size increases, the standardized sample mean of X can be approximated by the standard normal distribution. Now, suppose we drew 10,000 samples of 1,000 observations each from the same binary random variable (which by definition follows a Bernoulli distribution). Which of the following statements is true?

A. The standardized sample means will approximately follow a Bernoulli distribution
B. The standardized sample means will approximately follow a normal distribution with mean and variance unknown
C. The standardized sample means will approximately follow a standard normal distribution
D. None of the above

Answer: C

20. Thanks to the Central Limit Theorem, we know that if we were to draw multiple large samples from a random variable centered at ten at the population level, the distribution of the sample means over these multiple samples will be centered at:

A. ten
B. zero
C. the number of observations in each sample
D. it is impossible to know with the information given

Answer: A