CISC 7700X Final Exam

Pick the best answer that fits the question. Not all of the answers may be correct. If none of the answers fit, write your own answer.

1. (5 points) A *model* is:

   (a) A data point.

   (b) A description.

   (c) A fact.

   (d) All of the above.

2. (5 points) For last 3 years, your investment returned: {+25%, +25%, -50%}. What's the arithmetic mean of your returns:

(answer)

3. (5 points) For last 3 years, your investment returned: {+25%, +25%, -50%}. What's the geometric mean of your returns:

(answer)

4. (5 points) This technique allows assigning measures of accuracy to sample estimates of almost any statistic using random sampling methods.

   (a) Normal distribution curve with 95% accuracy

   (b) Bootstrapping

   (c) Standard deviation

   (d) 90% confidence interval

5. (5 points) A permutation test can be used for

   (a) Determining the sorting order of a randomized list.

   (b) Determining the $N!$ (N-factorial) of all permutations.

   (c) Determining the significance

   (d) Same places as Student s-Test.

6. (5 points) If 1-year standard deviation is 7, then 2-year standard deviation is:

   (a) 7

   (b) 10

   (c) 50

   (d) None of the above, the answer is:

7. (5 points) You find a random widget with serial number 4242. With 50% confidence, how many widgets are out there?

   (a) somewhere between 0 and 100000.

   (b) somewhere between 4242 and 4242*4.

   (c) at least 1000000 widgets.

(d) Not enough data to make a guess.

8. (5 points) If $P(x, y) \neq P(x)P(y)$ then

   (a) x is more likely than y.
   (b) x implies y.
   (c) x and y are independent.
   (d) x and y are not independent.
   (e) None of the above, answer is:

9. (5 points) If $P(y|x)P(x) \neq P(x|y)P(y)$ then

   (a) x is more likely after y.
   (b) y causes x.
   (c) x and y are independent.
   (d) x and y are not independent.
   (e) None of the above, answer is:

10. (5 points) The process of computing $P(x)$ from $P(x, y)$ is called

    (a) Bootstrapping
    (b) Generalizing
    (c) Marginalizing
    (d) Specifizing

11. (5 points) In Bayes rule: $P(x|y) = P(y|x)P(x)/P(y)$, the $P(x)$ is:

    (a) The likelihood.
    (b) The prior probability.
    (c) The posterior probability.
    (d) The posterior likelihood.

12. (5 points) In Bayes rule: $P(x|y) = P(y|x)P(x)/P(y)$, the $P(y|x)$ is:

    (a) The likelihood.
    (b) The prior probability.
    (c) The posterior probability.
    (d) The conditional probability of y given x.

13. (5 points) Conditional probability $P(y|x)$ differs from likelihood $P(y|x)$:

    (a) They're both the same.
    (b) They both sum to 1.
    (c) Probability $P(y|x)$ is a function of y, while likelihood $P(y|x)$ is a function of x.
    (d) Likelihood tells us the probability of y given x.

14. (5 points) We notice that 1 in 5 cars driving on Kings Highway are speeding (going faster than 25mph). Of the cars that are speeding, a third of them are BMWs. Of the cars that are not speeding, only a fifth of them are BMWs. You notice a BMW at an intersection, use Bayes rule to determine if it will be speeding once the light turns green:

(answer)

15. (5 points) Continuing from above, Of the cars that are speeding, a third of them are two-door coupes. Of the cars that are not speeding, only a tenth of them are two-door coupes. You notice a two-door coupe at an intersection, use Bayes rule to determine if it will be speeding once the light turns green:

(answer)

16. (5 points) Continuing from above, You notice a two-door BMW at an intersection, use Bayes rule to determine if it will be speeding once the light turns green:

(answer)

17. (5 points) Continuing from above, You notice a two-door BMW at an intersection, use Naive Bayes rule to determine if it will be speeding once the light turns green:

(answer)

18. (5 points) You form a hypothesis that BMWs cause speeding. To test your hypothesis you:

    (a) Need to count the number of speeders who are driving BMWs, and seeing if that is significant.

    (b) Collect data on all speeders, and see what fraction of them drive BMWs.

    (c) Collect data on all MBWs and see what fraction of them are speeders.

    (d) Conduct a controlled experiment.

19. (5 points) To determine if your outcome is significant, you can:

    (a) Collect a *a lot* of data.

    (b) Perform a controlled experiment.

    (c) Perform a permutation test.

    (d) Repeat the data collection and see if anything changes.

20. (5 points) Fair coin flipping game: We start with \$1. Heads we win 50%, tails we lose 50%. After 3 rounds, with a fair coin, the *geometric mean* value we will have:

    (e) Answer is: