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 1234. With 50% confidence, how many widgets are out there?

   (a) somewhere between 0 and 100000.

   (b) somewhere between 1234 and 1234*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) In our company, about 10% of the employees will leave the company within a year. We notice that of the folks who leave the company, about 80% were recently promoted. Of the folks who stay, only 5% were recently promoted. We decide to promote an employee. Use Bayes rule to determine probability that they will leave the company within a year?

(answer)

15. (5 points) Continuing from above, we notice that of the folks who leave the company, 60% have been at the company less than than 5 years. Of the folks who stay, 90% were with the company for less than 5-years. An employee has been with the company less than 5-years. Use Bayes rule to determine probability that they will leave the company within a year?

(answer)

16. (5 points) Continuing from above, we decide to promote an employee who has been with the company for 3 years. Use Bayes rule to determine probability that they will leave the company within a year?

(answer)

17. (5 points) Continuing from above, we decide to promote an employee who has been with the company for 3 years. Use Naive Bayes rule to determine probability that they will leave the company within a year?

(answer)

18. (5 points) You form a hypothesis that promotions cause employees to leave. To test your hypothesis you:

    (a) Need to count the number of employees who leave, and seeing if that is significant.
    (b) Collect data on all employees, and see what fraction of them were promoted.
    (c) Collect data on all promotions and see what fraction of them leave.
    (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: