STATISTICS WORKSHEET 6

Q1. Which of the following can be considered as random variable?

a) The outcome from the roll of a die

b) The outcome of flip of a coin

c) The outcome of exam

d) All of the mentioned

Ans. D) All of the mentioned

Q2. Which of the following random variable that take on only a countable number of possibilities?

a) Discrete

b) Non Discrete

c) Continuous

d) All of the mentioned

Ans. A)Discrete

Q3. Which of the following function is associated with a continuous random variable?

a) pdf

b) pmv

c) pmf

d) all of the mentioned

Ans.a)pdf

Q4. The expected value or \_\_\_\_\_\_\_ of a random variable is the center of its distribution.

a) mode

b) median

c) mean

d) bayesian inference

Ans.c)mean

5. Which of the following of a random variable is not a measure of spread?

a) variance

b) standard deviation

c) empirical mean

d) all of the mentioned

Ans.a) variance

6. The \_\_\_\_\_\_\_\_\_ of the Chi-squared distribution is twice the degrees of freedom.

a) variance

b) standard deviation

c) mode

d) none of the mentioned

Ans.a)variance

7. The beta distribution is the default prior for parameters between \_\_\_\_\_\_\_\_\_\_\_\_

a) 0 and 10

b) 1 and 2

c) 0 and 1

d) None of the mentioned

Ans.c)0 and 1

8. Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?

a) baggyer

b) bootstrap

c) jacknife

d) none of the mentioned

Ans.b)bootstrap

Q9. Data that summarize all observations in a category are called \_\_\_\_\_\_\_\_\_\_ data.

a) frequency

b) summarized

c) raw

d) none of the mentioned

Ans.b)summarized

**Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.**

Q10. What is the difference between a boxplot and histogram?

Ans. Histograms and box plots are very similar in that they both help to visualize and describe numeric data. Although histograms are better in determining the underlying distribution of the data, box plots allow you to compare multiple data sets better than histograms as they are less detailed and take up less space. It is recommended that you plot your data graphically before proceeding with further statistical analysis.

Q11. How to select metrics?

Ans. 1. Define your primary objective

Before you even begin to sift through the various metrics and statistics available to you, it is essential that your company’s governing objectives have been clearly established. As a B2B finance company, a primary objective could be to increase market share by 3% before the end of the year.

2. Choose your metric(s) – determine cause and effect

Once a clear, overarching objective has been established, most marketing companies look to major metrics to determine their success—factors such as the generation of sales and leads.

But these metrics aren’t the only indicator of a company’s success. Less easily quantifiable factors such as customer satisfaction and brand loyalty also play a significant role in the ability to achieve overall marketing objectives, especially in the long term.

Examining the relationship between these metrics can allow marketers and others to develop a cause-and-effect theory to determine what drives the end results. You may be tempted to aim your campaign at a huge but diverse audience, so as to widen the top of your sales funnel. But as you start to develop a clearer view on the effect of you actions on the achievement of objectives, it may turn out that spending your budget on potential customers further down the funnel would yield more sales

3. Create relevant activities

Digital technology has made it easier than ever to track the engagement of various types of marketing materials, be they a video, article, or even a podcast. Let’s look at how a marketing agency could help its clients to improve performance.

Once a marketing agency has determined that engaging content is what drives sales and leads for their clients, the agency must determine which types of content reliably generates that engagement.

4. Evaluate periodically

Of course, the ever-changing nature of marketing (and the business world as a whole), ensures that the measures you use to link activities with your primary goals must be constantly re-evaluated. The metrics and statistics that drive value for your clients can change over time, especially as new technologies emerge and target demographics shift.

Regularly evaluating your methods and adapting when necessary may cause you to throw away some of your work. But this is by no means a waste. Adjusting course on a regular basis, whether it is for objectives, metrics or activities, will ensure you remain competitive in the years to come.

Q12. How do you assess the statistical significance of an insight?

Ans. To assess statistical significance, you would use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Second, you’d calculate the p-value, which is the likelihood of getting the test’s observed findings if the null hypothesis is true. Finally, you would select the threshold of significance (alpha) and reject the null hypothesis if the p-value is smaller than the alpha — in other words, the result is statistically significant.

Q13. Give examples of data that doesnot have a Gaussian distribution, nor log-normal.

Ans. Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well. Example: Duration of a phone car, time until the next earthquake, etc.

Q14. Give an example where the median is a better measure than the mean.

Ans. Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed. The median indicates that half of all incomes fall below 27581, and half are above it. For these data, the mean overestimates where most household incomes fall.

Q15. What is the Likelihood?

Ans. Likelihood is central to parametric statistical inference. The likelihood is a basis for the likelihood ratio test: a uniformly most powerful test for comparing two point hypotheses. It is also the basis for the maximum likelihood estimate.

In practice one often calculates the natural logarithm of the likelihood function (log-likelihood) as being more convenient (easier to differentiate). The fact that a logarithm is strictly increasing is useful when calculating maximum likelihood: log-likelihood reaches the maximum at the same point as the likelihood.