Great Learning
POWER AHEAD

Statistical Analysis Interview Questions

1. A bag contains 3 white and 2 black balls. Another bag contains 5 white and 3 black balls. If a bag is chosen at random, and a ball is drawn from it, what is the chance that it is white?
2. A carnival game involves throwing 3 dice and summing up the numbers on the tops of the dice. The item associated with this sum is then given to the thrower as a prize. The most expensive prizes are to be kept with which sum to minimize the chance of collecting it?
3. You have 25 horses and you can race only 5 of them simultaneously. Assuming you do not have access to a stopwatch, how many times would you need to race the horses to find the 3 fastest horses.
4. What is the role of probability in inferential statistics?
5. How would you estimate the number of crocins used in Karnataka in a year?
6. How would you statistically proceed to estimate the average weight of all fishes in a lake?
7. What is the Central Limit Theorem and why is it extremely important for making inferences?
8. What is a P-value? What is its significance in an experimental setup?
9. What is the role of confidence intervals in the estimation process?
10. Researchers have developed a chemical additive to EV batteries and claim that the average lifespan of this new battery is greater than 1100 charge cycles. A sample of 25 EV batteries is chosen to test this claim.
    a. 1. State the hypotheses you will use?
    b. 2. Is it one-tailed or two-tailed?
    c. 3. What kind of test will you use and why?
11. Explain the difference between Type 1 and Type 2 Errors in a Hypothesis Testing Scenario using real-life examples.
12. What are degrees of freedom in general? For an ANOVA Test, what are the degrees of freedom?
13. When do we use a t-test vs a z-test?
14. What is precision and what is recall and what do they signify? Explain using real-life scenarios?
15. What is the Chi-Square Test of Independence? Explain with real-life examples?
16. An analog clock breaks into 4 pieces. The sum of numbers in each piece is 15. Find the pattern in which the clock is broken.
17. Statistically, how would you ensure that the sample you choose to train a machine learning model is representative of the population?
18. Statistically, how can you ensure that the features chosen for a machine learning model are significant?
19. What is the Statistical power of a hypothesis test? On what factors does statistical power depend on.
20. We want to increase the confidence level of our estimate but we don't want to compromise precision. How can we achieve this?