

STATISTICS & PROBABILITY

DESCRIPTIVE STATISTICS

1. What are mean, median, and mode? Explain the differences.
2. When can the mean be misleading?
3. When is the median better than the mean?
4. What is weighted mean? Give an example.
5. What is geometric mean? When is it used?
6. What is the use case of harmonic mean?
7. What is trimmed mean?
8. How is mode used in discrete data?
9. What is bimodal and multimodal distribution?
10. What is the relationship between mean, median, and mode in a skewed distribution?
11. What is variance? Explain the formula and intuition.
12. What is standard deviation? What is its relation to variance?
13. What is range? What is its limitation?
14. What is Interquartile Range (IQR)?
15. What is MAD (Mean Absolute Deviation)?
16. What is the Coefficient of Variation? When is it used?
17. What is the difference between percentile and quartile?
18. What do the 25th, 50th, and 75th percentiles mean?
19. What is the five number summary?
20. How do you detect outliers? Explain the IQR method.

PROBABILITY BASICS

21. What is probability? State the classical definition.
22. What are sample space and event?
23. What are mutually exclusive events? Give an example.
24. What is the difference between independent events and dependent events?
25. What are complementary events?
26. What is the addition rule of probability?

27. Explain the multiplication rule of probability.
28. What is conditional probability? State the formula.
29. What is the difference between $P(A|B)$ and $P(B|A)$?
30. What is the Law of Total Probability?
31. What is Bayes' theorem? Derive the formula.
32. What are prior probability and posterior probability?
33. Give a real-world example of Bayes' theorem (medical testing).
34. How is Bayes' theorem used in the Naive Bayes algorithm?
35. What are false positive and false negative?
36. What is a permutation? What is the formula?
37. What is a combination? How is it different from permutation?
38. What are the formulas for nCr and nPr ?
39. Explain the birthday paradox problem.
40. What is the difference between probability with replacement and without replacement?

PROBABILITY DISTRIBUTIONS

41. What is a random variable? What is the difference between discrete and continuous?
42. What is a Probability Mass Function (PMF)?
43. What is a Cumulative Distribution Function (CDF)?
44. What is expected value ($E[X]$)? How do you calculate it?
45. How do you calculate the variance of a random variable?
46. What is Bernoulli distribution? What are its parameters?
47. What is binomial distribution? Give a use case.
48. What are the mean and variance of binomial distribution?
49. What is the difference between binomial and Bernoulli?
50. Give an example of binomial distribution using coin toss.
51. What is Poisson distribution? When is it used?
52. What does the parameter lambda represent in Poisson distribution?
53. What are the mean and variance of Poisson distribution?
54. Give a real-world example - calls per hour in a call center.
55. What is the relationship between binomial and Poisson distribution?

56. What is a Probability Density Function (PDF)?
57. What is the difference between PDF and PMF?
58. What is uniform distribution? Both continuous and discrete types.
59. What is normal distribution (Gaussian distribution)?
60. What are the properties of normal distribution?
61. How do mean and standard deviation affect the normal curve?
62. What is the 68-95-99.7 rule (Empirical rule)?
63. What is standard normal distribution? What is Z-score?
64. How do you calculate Z-score? What is the formula?
65. How do you use the Z-table?

SAMPLING & ESTIMATION

66. What is the difference between population and sample?
67. What is random sampling? What are its types?
68. What is simple random sampling?
69. What is stratified sampling? When is it used?
70. What is cluster sampling?
71. Explain systematic sampling.
72. What is sampling bias? How do you avoid it?
73. Why is sample size selection important?
74. What is the Central Limit Theorem (CLT)? Very important!
75. What is the practical importance of CLT?
76. What is the distribution of the sample mean (large sample)?
77. What is standard error? How is it different from standard deviation?
78. What is the Law of Large Numbers?

HYPOTHESIS TESTING

79. What is hypothesis testing? When is it used?
80. What are null hypothesis (H_0) and alternative hypothesis (H_1)?
81. What are Type I error (α) and Type II error (β)?
82. What is significance level (alpha)? What are common values?

83. What is p-value? How do you interpret it?
84. What does $p\text{-value} < 0.05$ mean?
85. What is the relationship between confidence level and significance level?
86. What is the difference between one-tailed test and two-tailed test?
87. What is critical value?
88. What is test statistic?
89. What is Z-test? When is it used?
90. What is T-test? When is it different from Z-test?
91. Explain one-sample t-test, two-sample t-test, and paired t-test.
92. What is Chi-square test? What are its use cases?
93. What is ANOVA (Analysis of Variance)?
94. What is F-test?
95. When do you use which statistical test?

CORRELATION & REGRESSION

96. What is correlation? What are positive and negative correlation?
97. What is Pearson correlation coefficient? What is its range (-1 to +1)?
98. What is the difference between correlation and causation? (Very important!)
99. What is Spearman rank correlation? When is it different from Pearson?
100. How do you interpret the correlation coefficient?