

Join Data Science Interview MasterClass (in 4 weeks) 🚀 led by FAANG Data Scientists | Just 8 seats remaining...



Prep Now



Statistics Interview Questions

Dan Lee

Updated Nov 12, 2024 — 9 min read

 **Table of Contents**

Looking for REAL statistics interview questions asked by top companies? Here's a comprehensive guide with REAL questions!

These questions are frequently used in interviews for data scientist and data analyst roles. Companies like Google, Amazon, Meta, Microsoft, and many others use these to gauge candidates' statistical knowledge and analytical skills.

In this guide, we'll explore critical statistics question areas, prep tips, and a list of questions to help you ace your interviews!

Let's dive in ↴



What is a Statistics Interview?

Here's a breakdown of common statistics topics assessed in data-related interviews.

Area 1 - Probability Fundamentals

A solid understanding of probability is essential. Expect questions that test your grasp of probability concepts like conditional probability, Bayes' theorem, distributions, and expected values. This foundational knowledge is crucial for problem-solving in data science.

Sample Questions

- What's the difference between independent and mutually exclusive events?
- Explain Bayes' theorem and provide an example of its use.
- How do you calculate the probability of at least one event occurring?
- Describe the concept of conditional probability with an example.
- What's the expected value of a random variable?



Tip: Memorize basic probability formulas and do drills to hone in solving probability teaser problems.

Area 2 - Distributions and Hypothesis Testing

Proficiency in statistical distributions and hypothesis testing is vital for data analysis. You'll likely face questions on distributions (normal, binomial, Poisson) and on framing and interpreting hypothesis tests.

Sample Questions

- What is the difference between a t-test and a z-test?
- Explain the central limit theorem and its significance.
- How do you interpret a p-value?
- Describe the properties of a normal distribution.
- What is a Type I and Type II error?



Tip: Interviewers often ask candidates how to explain the p-value and confidence interval to non-technical stakeholders. When you study these concepts, practice explaining these concepts plainly.

Area 3 - Regression and Correlation

Data science interviews often include questions on regression analysis and correlation to assess your ability to interpret relationships between variables.

Sample Questions

- How do you interpret the slope in a linear regression?
- What's the difference between correlation and causation?
- Explain multicollinearity and how to detect it.
- What's the difference between R-squared and adjusted R-squared?
- How do you handle outliers in a regression analysis?



Tip: Be prepared to solve case studies involving regression models. One area to cover in such cases includes assumptions check in regression models.

Area 4 - Sampling and Experimental Design

Proficiency in sampling techniques and experimental design is crucial for data-driven decision-making. Expect questions that test your knowledge of random sampling, sampling bias, and experiment setup.

Sample Questions

- What's the difference between stratified and cluster sampling?
- How do you minimize sampling bias?
- Explain the concept of statistical power.
- What is an A/B test, and how do you interpret the results?
- Describe the difference between observational and experimental studies.



Tip: Be prepared to discuss sampling and design questions in the A/B testing interviews.

Area 5 - Advanced Statistical Modeling

Data science and machine learning roles require knowledge of advanced statistical models. Be ready for questions on topics like logistic regression, decision trees, and model evaluation.

Sample Questions

- How do you interpret coefficients in logistic regression?
- Explain the bias-variance tradeoff.

- What is cross-validation, and why is it important?
- How do you handle imbalanced datasets in classification problems?
- Explain the use of ROC curves and AUC.

Area 6 - Time Series Analysis

In roles focused on forecasting, you may be tested on time series analysis, including methods like ARIMA, seasonality, and trend analysis.

Sample Questions

- How do you decompose a time series?
- Explain the difference between seasonality and trend.
- What is an ARIMA model, and when would you use it?
- How do you handle missing data in time series?
- What is autocorrelation, and why is it important?

★ Statistics Interviews Across Data Roles

Here's how statistics questions vary slightly between the data analyst and data scientist interviews. Be prepared to discuss your experience in statistical analysis based on projects you have achieved in academics or past work experiences.

Data Analyst Interview

The questions tend to be fairly light compared to the data scientist interview. Expect to cover basics in statistics, such as the definition of the p-value.

- **Format:** Primarily focused on 30-60 minute technical screens with discussions on data analysis projects.

- **Number of Questions:** Typically, 4-6 questions per round related to statistical analysis and interpretation.
- **Focus Areas:** Expect questions on basic statistics, descriptive analysis, and hypothesis testing. Common tools include Excel, SQL, and sometimes Python. The interviewer may ask you to analyze datasets, interpret statistical results, or describe how you'd approach real-world data problems.

Data Scientist Interview

Compared to the data analyst role, there's more rigor in statistical questions posed across interviews from technical screens to on-site.

- **Format:** May include technical assessments, take-home assignments, and live coding/statistical analysis. The first technical round is generally 30 to 60 minutes, followed by a take-home, which requires 3-5 days to submit, and then an on-site with multiple rounds. 1 or 2 of those onsite rounds tend to focus on statistics. Here's an example of how statistical interviews surface in the [Google Data Scientist interviews](#).
- **Number of Questions:** 4-6 per technical or on-site round with follow-up discussions.
- **Focus Areas:** Emphasis on inferential statistics, regression analysis, and A/B testing. The interviewer might probe your knowledge of statistical modeling, hypothesis testing, and predictive analytics. Expect questions that require critical thinking and application of statistical methods to complex datasets.

👉 Looking for targeted interview prep? Consider the [Data Scientist Interview MasterClass](#)—a hands-on workshop from top industry experts to boost your skills!



More Statistics Interview Questions

Here is a comprehensive list of 120 statistical interview questions, categorized across the six key areas:

1. Probability Fundamentals

1. What is the probability of getting at least one head when flipping two coins?
2. Define and differentiate between discrete and continuous probability distributions.
3. Explain conditional probability and provide an example.
4. What is Bayes' theorem, and how is it applied?
5. How would you calculate the probability of independent events occurring together?
6. What is a joint probability, and how is it different from conditional probability?
7. Explain the law of total probability.
8. What is an expected value, and how do you calculate it?
9. Describe the concept of variance in probability.
10. Explain the difference between mutually exclusive and independent events.
11. What is the probability of drawing two aces from a standard deck of cards?
12. Define cumulative distribution function (CDF).
13. Explain what a probability density function (PDF) is.
14. What is the difference between a permutation and a combination?
15. How would you explain the concept of a random variable?
16. What is a Markov chain, and where is it used?
17. How does probability differ from likelihood?
18. What is conditional independence?
19. Describe the difference between prior, likelihood, and posterior in Bayesian analysis.
20. Explain the Monty Hall problem and its solution.

2. Distributions and Hypothesis Testing

1. What is a normal distribution, and why is it important?
 2. Explain the difference between a t-test and a z-test.
 3. What is the central limit theorem, and why is it significant?
 4. Define p-value and explain its role in hypothesis testing.
 5. What are Type I and Type II errors?
 6. Describe a binomial distribution and give an example.
 7. Explain the Poisson distribution and a scenario where it's applicable.
 8. What is a chi-square test, and when would you use it?
 9. Define the null and alternative hypotheses.
 10. How do you interpret a confidence interval?
 11. Explain the concept of statistical power.
 12. How is an ANOVA test conducted, and what does it test?
 13. What is the F-test, and when would you use it?
 14. Explain homoscedasticity and heteroscedasticity.
 15. How do you test if a dataset follows a normal distribution?
 16. Describe what a one-sample t-test is.
 17. What is an effect size, and why is it important?
 18. How do you control for multiple comparisons in hypothesis testing?
 19. Explain the difference between a one-tailed and two-tailed test.
 20. What are some limitations of p-values in hypothesis testing?
-

3. Regression and Correlation

1. How do you interpret the coefficients in a linear regression?

2. What is the difference between correlation and causation?
 3. Define multicollinearity and describe how to detect it.
 4. How would you handle multicollinearity in a regression model?
 5. Explain R-squared and adjusted R-squared.
 6. What is logistic regression, and when would you use it?
 7. Describe the assumptions of linear regression.
 8. How do you interpret the intercept term in a regression model?
 9. Explain the concept of regularization in regression.
 10. What is the difference between Lasso and Ridge regression?
 11. How do you handle outliers in a regression analysis?
 12. What is the purpose of polynomial regression?
 13. Explain the difference between simple linear regression and multiple regression.
 14. How do you interpret the odds ratio in logistic regression?
 15. What is a residual, and why is it important in regression analysis?
 16. Describe the difference between bias and variance in modeling.
 17. What is cross-validation, and how is it used in regression?
 18. How would you implement stepwise regression?
 19. What is the difference between overfitting and underfitting?
 20. How do you choose the best model for a regression analysis?
-

4. Sampling and Experimental Design

1. What is random sampling, and why is it important?
2. Explain the difference between stratified sampling and cluster sampling.

3. How do you minimize sampling bias?
 4. Define and explain the concept of statistical power.
 5. What is an A/B test, and how is it conducted?
 6. How would you calculate the sample size needed for an experiment?
 7. Describe the concept of control and treatment groups in experiments.
 8. What is a placebo effect, and why is it important in experimental design?
 9. Explain the difference between observational and experimental studies.
 10. How would you interpret the results of an A/B test?
 11. What is a confounding variable, and how do you control for it?
 12. Explain the concept of random assignment.
 13. What is the difference between within-subjects and between-subjects design?
 14. How do you account for selection bias in an experiment?
 15. What is a crossover design, and when is it used?
 16. Describe the purpose of a power analysis.
 17. How do you calculate the margin of error in a sample?
 18. Explain what a matched pairs design is.
 19. What is external validity, and why is it important?
 20. Describe a factorial design and give an example.
-

5. Advanced Statistical Modeling

1. How does logistic regression differ from linear regression?
2. Explain the bias-variance tradeoff.
3. What is the purpose of cross-validation in model evaluation?

4. How do you handle imbalanced datasets in classification problems?
 5. Describe the ROC curve and AUC metric.
 6. What is a decision tree, and how does it work?
 7. Explain the concept of ensemble learning.
 8. What is the purpose of the K-nearest neighbors (KNN) algorithm?
 9. How do you interpret feature importance in a model?
 10. What is a confusion matrix, and how is it used?
 11. Describe the difference between bagging and boosting.
 12. How do you evaluate the accuracy of a classification model?
 13. Explain principal component analysis (PCA).
 14. What is a support vector machine (SVM)?
 15. How would you perform dimensionality reduction on a dataset?
 16. Describe overfitting in machine learning models.
 17. How do you handle missing data in a dataset?
 18. What is the purpose of grid search in model selection?
 19. Describe the K-means clustering algorithm.
 20. What are precision and recall, and why are they important?
-

6. Time Series Analysis

1. What is a time series, and how is it different from other data types?
2. Describe the components of a time series.
3. How do you decompose a time series?
4. Explain the concept of seasonality in time series.
5. What is a trend, and how do you identify it in data?

6. What is an ARIMA model, and when is it used?
7. Define autocorrelation and partial autocorrelation.
8. Explain the purpose of differencing in time series analysis.
9. How do you handle missing values in a time series?
10. What is the Box-Jenkins methodology?
11. Describe exponential smoothing and its applications.
12. How do you perform a stationarity test in time series?
13. What is a moving average, and why is it useful?
14. Explain the concept of a lag in time series.
15. Describe the purpose of a seasonal decomposition of time series (STL).
16. How do you select the best model for time series forecasting?
17. Explain the difference between additive and multiplicative models.
18. What is the purpose of a rolling window in time series analysis?
19. How do you evaluate a time series forecasting model?
20. Describe the Holt-Winters model and its applications.



How to Prepare for Statistics Interviews

Tip 1 - Understand the Statistics Interview Format

Statistics interviews can vary significantly depending on the role—ranging from theoretical questions to practical applications. Be ready to discuss statistical concepts, interpret data, and solve problems related to probability, hypothesis testing, and data analysis. Familiarize yourself with common interview formats, which often include problem-solving exercises, statistical interpretation, and real-world scenario-based questions. For a comprehensive guide, consider joining the [DataInterview Bootcamp](https://www.datainterview.com/blog/statistics-interview-questions) to get in-depth insights on what to expect.

Tip 2 - Join Prep Communities

Joining a community is one of the best ways to prepare for statistics interviews. Connect with others in the [DataInterview Premium Community](#) to network with coaches and peers actively preparing for top tech interviews. Engage in weekly group sessions to review sample questions, discuss challenging concepts, and reinforce your understanding.

Tip 3 - Get Personalized Coaching

Consider scheduling mock interviews with experienced coaches to simulate real interview conditions. Personalized coaching helps you receive valuable feedback, strengthen your communication skills, and refine your statistical problem-solving techniques. Platforms like [DataInterview Coaching](#) offer tailored sessions to prepare you for technical and behavioral questions, ensuring you're interview-ready.

With these strategies, you'll be well-prepared to approach statistics interviews confidently!

Unlock the content that helped 500+ candidates ACE interviews.

Prep Now

- ✓ 300+ Lessons and Practice Problems
- ✓ 40+ Hours of Video Content
- ✓ FAANG Instructors
- ✓ Lifetime Slack Access

✉ Get free interview tips and practice question every week.

Get Free Question

Dan Lee

DataInterview Founder (Ex-Google)

Dan Lee is a former Data Scientist at Google with 8+ years of experience in data science, data engineering, and ML engineering. He has helped 100+ clients land top data, ML, AI jobs at reputable companies and startups such as Google, Meta, Instacart, Stripe and such.



Related Articles

Square (Block) Data Engineer Interview

[Read more](#)

Shopify Data Analyst Interview

[Read more](#)

Morgan Stanley Data Analyst Interview

[Read more](#)

Subscribe to our newsletter

The latest news, articles, and resources, sent to your inbox weekly.

Subscribe

Products

[Courses](#)

[Coaching](#)

[SQL Pad](#)

Support

[Pricing](#)

[Blog](#)

Legal

[Terms](#)

[Privacy](#)



© 2025 DataInterview, LLC. All rights reserved.