K-Means Clustering 41 K-Nearest Neighbors 13

Keras 62 LLMOps 25 LLMs 59

Linear Algebra 42 Linear Regression 48

Logistic Regression 25 MATLAB 59

ML Design Patterns 30 MLOps 37

Model Evaluation 33 NLP 38

Naïve Bayes 20 Neural Networks 140

NumPy 40 Optimisation 33 PCA 17

Pandas 68 Probability 36 PyTorch 36

Python 139 Q-Learning 12 R R 41

RNN 18 Random Forest 41

Recommendation Systems 30

Reinforcement Learning 55 SQL 47

SVM 56 Scala 39 Scikit-Learn 33

Searching 24 Sorting 26 Statistics 59

Supervised Learning 32 TensorFlow 31

Time Series 38 Unsupervised Learning 27

35 Advanced Statistics Interview Questions For Data Scientists and ML Engineers

Statistics 59

> Statistics is one of the most important disciplines to provide tools and methods to find structure in and to give deeper insight into data. And Data Science in return is a scientific discipline is influenced by informatics, computer science, mathematics, operations research, and statistics. Follow along and check the 35 most common and advanced Statistics and Probability Interview Questions and Answers any data scientists, data analysts,s and machine learning engineers must know before the next ML Interview.

Q1: **What is a *Probability Distribution*?**

Entry Probability 36

Answer

A **Probability Distribution** is a statistical function that describes all the possible values and likelihood that a random variable can take within a given range.

There are two main types of probability distribution:

- **Discrete probability distributions**: used for random variables with discrete outcomes, for example, the number of heads in five consecutive coin tosses, the number of rainy days in a given week, the number of goals scored by a player, and so on.
- **Continuous probability distributions**: used for random variables with continuous outcomes, for example, the height of male students,

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: www.statisticshowto.com

Junior Probability 36

Q2: Box challenge: Which box has a higher probability of getting cards of the same color and why?

Problem

A box has `12` red cards and `12` black cards. Another box has `24` red cards and `24` black cards. You want to draw two cards at random from one of the two boxes, one card at a time. Which box has a higher probability of getting cards of the same color and why?

Answer

Let's say the first card we draw from each deck is a red Ace. This means that in the deck of the first box with `12` reds and `12` blacks, there's now `11` reds and `12` blacks. Therefore the odds of drawing another red are equal to:

$$red_{firstBox} = \frac{11}{11+12} = \frac{11}{23} = 0.4782\cdots$$

In the second box that has a deck with `24` reds and `24` blacks, there would then be `23` reds and `24` blacks. Therefore the odds of drawing another red are equal to:

$$red_{secondBox} = \frac{23}{23+24} = \frac{23}{47} = 0.4893\cdots$$

Since `23/47` > `11/23`, the second deck with more cards has a higher probability of getting the same two cards.

Q3: Can there be more than one *Mode*?

Junior Statistics 59

Answer

The **mode** is the value that appears most frequently in a data set. A set of data may have one mode, more than one mode, or no mode at all. A data set can often have no mode, one mode, or more than one mode – it all depends on how many different values repeat most frequently.

> For example, in the following list of numbers, 16 is the mode since it appears more times in the set than any other number:
>
> - 3, 3, 6, 9, **16, 16, 16**, 27, 27, 37, 48

Your data can be:

- without any mode
- **unimodal**, with one mode,
- **bimodal**, with two modes,
- **trimodal**, with three modes, or
- **multimodal**, with four or more modes.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: www.scribbr.com

Q4: Rolling a fair die event challenge

Junior Probability 36

Problem

Let's suppose I roll a fair die. Let `A` be the event that an outcome is an odd number and let `B` be the event that the outcome is less than or equal to `3`. What is the probability `P(A|B)` ?

outcome must be among `{1,2,3}`. For `A` to also happen the outcome must be in `A ∩ B = {1,3}`. Since all die rolls are equally likely, then we calculate the conditional probability as:

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{2}{3}$$

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: www.probabilitycourse.com

Q5: What is *Central Tendency*?

Junior Statistics 59

Answer

A measure of **central tendency** is a single value that attempts to *describe* a set of data by identifying the **central position** within that set of data. In the simplest of terms, it attempts to find a single value that best represents an entire distribution of scores.

Mean, **Median** and **Mode** are average values or **central tendency** of a numerical data set.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: statistics.laerd.com

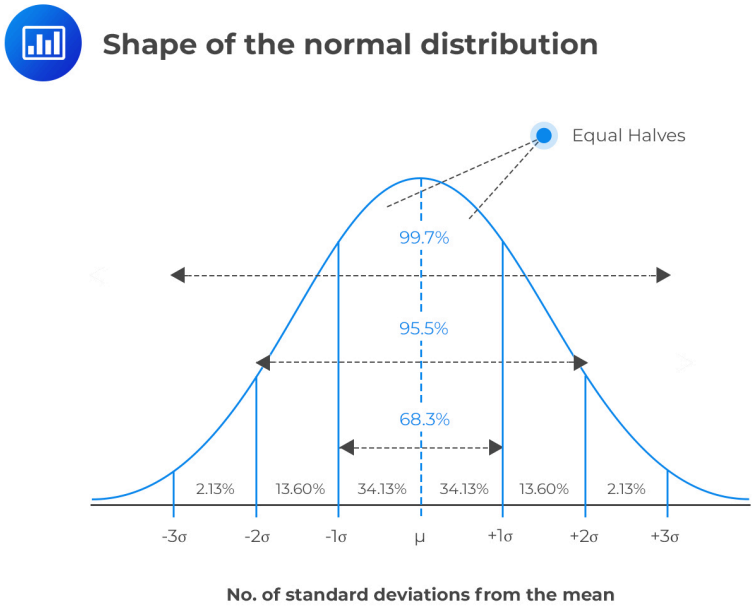
23 Apache Spark Interview Questions (ANSWERED) To Learn Before ML & Big Data Interview

Apache Spark 30

Q6: What is *Normal Distribution*?

Junior Statistics 59

statistics because it fits many natural phenomena. For example, heights, blood pressure, measurement error, and IQ scores follow the normal distribution. It is also known as the Gaussian distribution and the bell curve.



Normal distributions have the following features:

- symmetric bell shape
- *mean* and *median* and *mode* are equal; both located at the center of the distribution
- ≈68% of the data falls within 1 standard deviation of the mean
- ≈95% of the data falls within 2 standard deviations of the mean
- ≈99.7% of the data falls within 3 standard deviations of the mean

---

Having Machine Learning, Data Science or Python Interview? Check 👉 59 Statistics Interview Questions

Source: statisticsbyjim.com

Q7: **What is *Statistical Significance*?**

Junior Statistics 59

Answer

probability value.

Statistical significance is arbitrary – it depends on the threshold, or alpha value, chosen by the researcher. The most common threshold is $p < 0.05$, which means that the data is likely to occur less than 5% of the time under the null hypothesis.

When the *p*-value falls below the chosen alpha value, then we say the result of the test is statistically significant.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

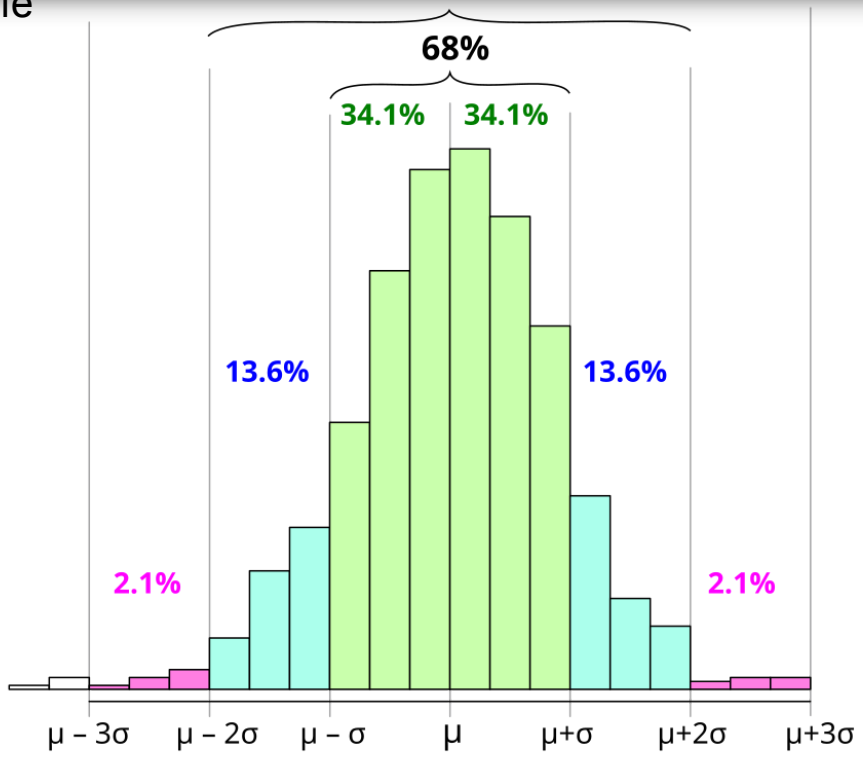
Source: www.scribbr.com

Q8: What is the `68 - 95 - 99.7` rule for *Normal Distribution*?

Junior · Anomaly Detection 47

Answer

- The **68-95-99.7 rule**, is also known as the **Empirical Rule**.
- It is a shorthand used to remember the percentage of values that lie within an *interval estimate* in a **Normal Distribution**.
- It is shown below:



The rule states that `68%`, `95%`, and `99.7%` of the values lie within one, two, and three *standard deviations* of the *mean*, respectively.

Having Machine Learning, Data Science or Python Interview? Check 👉
47 Anomaly Detection Interview Questions

Source: en.wikipedia.org

Difference between *Descriptive Statistics* and *Inferential Statistics*?

Answer

- **Descriptive statistics**, as its name suggests, focus on *describing* the characteristics or features of a dataset. Here we look for measures of *distribution*, *central tendency* and *variability* in order to draw conclusions based on known data.
- **Inferential statistics** focus on making *generalizations* about a larger population based on a representative sample of that population, It also allows us to make *predictions* so its results are usually in the form of a *probability*. Here, we perform *hypothesis testing*, compute *confidence intervals*, make *regression* and *correlation* analyses, in order to draw conclusions that go beyond the available data.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: careerfoundry.com

23 Azure Machine Learning Interview Questions (ANSWERED)

Azure ML 30

Q10: What is the difference between a *Combination* and a *Permutation*?

Junior Probability 36

- A **Combination** is the choice of *r* elements from a set of *n* elements *without replacement* and where *order does not matter*. Is most used to group data. For example, picking three team members from a group, picking two colors from a color brochure, etc. It is mathematically defined as:

$$C_r^n = \frac{n!}{(n-r)r!}$$

- A **Permutation** is the choice of *r* elements from a set of *n* elements *without replacement* and where the *order matters*. Is used to list data, for example picking first, second and third place winners, picking two favorite colors -in order- from a color brochure, etc. It is mathematically defined as:

$$P_{n,r} = \frac{n!}{(n-r)!}$$

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: byjus.com

Junior Probability 36

Q11: What is the difference between the *Bernoulli* and *Binomial* distribution?

Answer

- The **Bernoulli distribution** is the *discrete probability distribution* of a random variable which takes a **binary** output: `1` with probability `p`, and `0` with probability `(1-p)`. The idea is that, whenever we are running an experiment that might lead either to *success* or to a *failure*, we can associate with our *success* (labeled with `1`) a probability `p`, while our *failure* (labeled with `0`) will have probability `(1-p)`.

- In the **Binomial distribution** we keep the same idea as before: we count *probability* of a *success* or *failure* outcome in an experiment, but this time is is *repeated multiple times*.

Source: www.statisticshowto.com

Q12: What's the difference between *Confidence Interval* and *Confidence Level*?

Junior Statistics 59

Answer

- The **confidence level** is the percentage of times we expect to get close to the same estimate if we run our experiment again or resample the population in the same way.
- The **confidence interval** is the actual upper and lower bounds of the estimate we expect to find at a given level of confidence.

For example, if we are estimating a `95%` *confidence interval* around the mean proportion of female babies born every year based on a random sample of babies, we might find an upper bound of `0.56` and a lower bound of `0.48`. These are the upper and lower bounds of the *confidence interval* for a *confidence level* of `95%`.

This means that `95%` of the time, we can expect our estimate to fall between `0.56` and `0.48`.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: www.scribbr.com

Q13: A coin was flipped `1000` times, and `550` times it showed up heads.

Mid Probability 36

Answer

To answer this question let's say X is the number of heads and let's assume that the coin is not biased. Since each individual flip is a Bernoulli random variable, we can assume it has a probability of showing up heads as `p = 0.5`, so this will lead to the following expected number of heads:

$$\mu = np = 1000 \times 0.5 = 500$$

And the following standard deviation:

$$\sigma = \sqrt{np(1-p)} = \sqrt{1000 \times 0.5 \times 0.5} = \sqrt{250} \approx \sqrt{16}$$

Given that we got a `1000` sample size, we can apply the Central Limit Theorem to approximate the total number of heads as normal distribution and calculate the corresponding z-score to test the hypothesis that the coin is fair.

$$z = \frac{x - \mu}{\sigma} = \frac{550 - 500}{16} = 3.125 > 3$$

This means that, if the coin were fair, the event of seeing 550 heads should occur with a < `1%` chance under normality assumptions. Therefore, the coin is likely biased.

Having Machine Learning, Data Science or Python Interview? Check 👉 36 Probability Interview Questions

Source: www.nicksingh.com

19 Feature Engineering Interview Questions (EXPLAINED) for Data Scientists

Data Mining 13 Feature Engineering 28

Q14: Explain how to use *Standard*

Mid Anomaly Detection 47

Detection?

Answer

In statistics, if a **data distribution** is approximately **normal** then about 68% of the data values lie within **one standard deviation** of the mean and about 95% are within two standard deviations, and **about 99.7%** lie within three standard deviations:

Therefore, if you have any data point that is more than 3 times the standard deviation, then those points are very likely to be **anomalous** or **outliers**.

Having Machine Learning, Data Science or Python Interview? Check 👉
47 Anomaly Detection Interview Questions

Source: towardsdatascience.com

Q15: Given two fair dices, what is the probability that two dices sum to `8` ?

Mid Naïve Bayes 20

Problem

What is the probability that two dices sum to `8` when the first dice is `6` ?

Answer

Given the *probability* formula:

$$P = \frac{n}{n_{total}}$$

where n is the number of desired outcomes and n_{total} is the number of possible outcomes. For one single fair dice, we have `6` possible outcomes, so for two fair dices, we have `6*6 = 36` possible outcomes.

We want that the two dices sums `8` so the possible combinations are:

- `(2,6),(3,5),(4,4)`

$$P = \frac{5}{36} = 0.13888\cdots$$

For the second part we use the *Bayes Theorem* and *conditional probability* formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Let's say:

- `A` is the event of getting an `8`
- `B` is the event of getting a `3` .

Now, from the previous combination, we got that `(3,5)` is the only event that satisfies both `A` and `B` so then `P(A ∩ B) = 1/36` . On other hand, the probability of getting a `3` in a fair dice is just `1/6` , we now put these values on the previous formula and compute:

$$P(A|B) = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} = 0.1666\cdots$$

Having Machine Learning, Data Science or Python Interview? Check 👉
20 Naïve Bayes Interview Questions

Source: www.quora.com

Linear Regression 48 · Hadoop 32 · Optimisation 33 · Reinforcement Learning 55 · Azure ML 30 · NLP 38 · MLOps 37 · Q-Learning 12 · Big-O Notation 22 · Bias & Variance 19

Q16: How many *Sampling*

Mid · Statistics 59

Answer

- **Simple Random Sampling**: Every element has an *equal chance* of getting selected to be the part sample. It is used when we don't have any kind of prior information about the target population.
- **Stratified Sampling**: we divide the elements of the population into small *subgroups* (or *strata*) based on the *similarity* in such a way that the elements within the group are *homogeneous* and *heterogeneous* among the other subgroups formed. Then, the *elements* are randomly selected from each of these *strata*.
- **Cluster Sampling**: we divide the entire population into *subgroups*, wherein, each of those subgroups has *similar* characteristics to that of the population when considered in totality. Also, instead of sampling individuals, we randomly select the *entire subgroups*.
- **Systematic Clustering**: is about sampling items from the population at regular *predefined intervals*. For example, every `5th` element, `21st` element and so on. All the elements are put together in a sequence first where each element has an *equal chance* of being selected.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: towardsdatascience.com

Q17: How many types of *Descriptive Statistics* do you know?

Mid Statistics 59

Answer

Descriptive statistics, in short, help describe and understand the features of a specific data set by giving short summaries about the sample and measures of the data.

> For instance, consider a simple number used to summarize how well a batter is performing in baseball, the batting average. This single number is simply the number of hits divided by the number of times at bat (reported to three significant digits). A batter who is hitting

- **Frequency Distribution**: Used for both *quantitative* and *qualitative* data, frequency distribution depicts the count of the different outcomes in a data set and summarizes it in tables or graphs. It allows for a more structured and organized way to present raw data.
- **Central Tendency**: refers to a dataset's descriptive summary using a single value reflecting the *center* of the data distribution. The three most common measures of central tendency are the *mode*, *median*, and *mean*.
- **Variability**: it concerns how far apart the data points appear to fall from the center. It is most commonly measured with *range*, *interquartile range*, *standard deviation*, and *variance*.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: corporatefinanceinstitute.com

23 Classification Interview Questions (ANSWERED) For ML Engineers And Data Scientists

Classification 43

Q18: How many types of measures of *Variability* do you know?

Mid Statistics 59

Answer

- **Range**: is the difference between the *largest* and *smallest* values in that dataset. This measure is easy to calculate but it is very much affected by *extreme values*.
- **The Interquartile Range (IQR)**: Is the difference between *upper quartile (Q3)* and *lower quartile (Q1)*, we can see it as the middle half of the data.

is in relation to the *mean*.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: online.stat.psu.edu

Q19: State *null* and *alternate* hypothesis for a relationship between gender and height

Mid Statistics 59

Problem

You want to test whether there is a relationship between gender and height. Based on your knowledge of human physiology, you formulate a hypothesis that men are, on average, taller than women. State your *null* and *alternate* hypothesis.

Answer

The **alternate hypothesis** is usually your initial hypothesis that predicts a relationship between variables. The **null hypothesis** is a prediction of no relationship between the variables you are interested in.

For this example:

- Ho (null hypothesis): Men are, on average, not taller than women.
- Ha (alternate hypothesis): Men are, on average, taller than women.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: www.scribbr.com

Q20: Three Ants Challenge

Mid Probability 36

Problem

Answer

Each ant has two possible ways to go: the edge on its left ***L*** and the edge on its right ***R***. Now the only way no ant will collide is if they all walk in the same direction along the triangle (assuming they all move at the same speed). Overall the ways how the ants can move are:

1. LLL
2. LLR
3. LRL
4. RLL
5. LRR
6. RLR
7. RRL
8. RRR

We have a total of 8 ways how the 3 Ants can move, out of these, only RRR and LLL are the ways by which the Ants will never meet. So the probability of it is `2/8 = 0.25`

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: www.quora.com

Q21: True Positive and False Positive challenge

Mid Probability 36

Problem

A test has a true positive rate of `100%` and a false positive rate of `5%`. There is a population with a `1/1000` rate of having the condition the test identifies. Considering a positive test, what is the probability of having that condition?

Answer

Let's denoted as A a person who has the disease, and let B a positive test. To calculate P(A|B) and we will use the Bayes' theorem:

where `P(B)` is given by the law of total probability:

$$P(B) = P(B|A)P(A) + P(B|notA)P(notA)$$

Now, we know that:

- Out of `1000` people, `1` person who has the disease will get true positive result, so `P(A) = 0.001` and `P(notA) = 1 - P(A) = 0.999`.
- The test has a true positive rate of `100%`, so `P(B|A) = 1`
- The test has a false positive rate of `5%`, so `P(B|notA) = 0.05`.

Therefore,

$$P(A|B) = \frac{1 \cdot 0.001}{1 \cdot 0.001 + 0.05 \cdot 0.999} = 0.0198 \approx 0.02$$

Thus, there is only a `2%` probability of one person having the disease even if the reports say that it has the disease.

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: books.google.co.ve

27 Reinforcement Learning Interview Questions (ANSWERED) for Machine Learning Engineers

Reinforcement Learning 55

Q22: What *Bayes' Theorem (Bayes Rule)* is all about?

Mid Naïve Bayes 20

Answer

Often, we know how frequently some particular evidence is observed, *given a known outcome*. We have to use this known fact to compute the reverse, to compute the chance of that *outcome happening*, given the

for the calculation of certain **conditional probabilities**. Conditional probabilities are just those probabilities that reflect the influence of one event on the probability of another.

The formula is:

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

Which tells us:

- how often A happens *given that B happens*, written $P(A|B)$,

When we know:

- how often B happens *given that A happens*, written $P(B|A)$
- and how likely A is on its own, written $P(A)$
- and how likely B is on its own, written $P(B)$

As an example let us say **P(Fire)** means how often there is fire, and **P(Smoke)** means how often we see smoke, then

- **P(Fire | Smoke)** means how often there is fire when we can see smoke
- **P(Smoke | Fire)** means how often we can see smoke when there is fire

So the formula kind of tells us "forwards" **P(Fire | Smoke)** when we know "backwards" **P(Smoke | Fire)**

Having Machine Learning, Data Science or Python Interview? Check 👉
20 Naïve Bayes Interview Questions

Source: www.mathsisfun.com

Q23: What is the *Bayesian approach* to probability?

Mid | Probability 36

Answer

The ***Bayesian approach*** defines probability as the *measure of believability* one has about how a particular event occurs. It uses the ***Bayes theorem*** to help us update beliefs about random events once we've seen new evidence about the event, so we use Bayesian statistics to create different beliefs after new evidence is uncovered. It differs from *frequentist statistics* that rely only on data from repeated trials.

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: www.indeed.com

Q24: What is the *Central Limit Theorem (CLT)*?

Mid | Statistics 59

Answer

In probability theory, the **Central Limit Theorem (CLT)** states that the distribution of a sample variable *approximates a normal distribution* (i.e., a “bell curve”) as the sample size becomes larger, assuming that all samples are identical in size, and regardless of the population's actual distribution shape.

Put another way, CLT is a statistical premise that, given a sufficiently large sample size from a population with a finite level of variance, the *mean* of all sampled variables from the same population *will be approximately equal* to the mean of the whole population.

Mathematically, the **Central Limit Theorem (CLT)** is a statement about the cumulative distribution function of the random variable

mean μ and standard deviation σ. The CLT asserts that for each a, $-\infty < a < \infty$,

$$F_{Z_n}(a) = P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right\} \to \Phi(a) = \int_{-\infty}^{a} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \mathrm{d}x$$

as $n \to \infty$.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: www.investopedia.com

Q25: What's the difference between *Homoskedasticity* and *Heteroskedasticity*?

Mid Statistics 59

Problem

How would you identify each one and what is their importance?

Answer

- **Homoskedasticity** occurs when the *variance* of the error term in a regression model is *constant*.
- **Heteroskedasticity** happens when the *standard errors* of a variable, monitored over a specific amount of time, are **non-constant**.

We can identify each case by plotting the *residual values* vs the *fitted values* of a linear regression model.

The importance of each one case relies on the context of the **ordinary least squares (OLS)** estimator, which is a common way to perform *linear regression*. In *OLS*, we must satisfy the assumption of *homoskedasticity* in order to get an *efficient* estimation.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

23 Logistic Regression Interview Questions (SOLVED) To Nail On ML Interview

Logistic Regression 25

Q26: What's the difference between *Kurtosis* and *Skewness*?

Mid Statistics 59

Answer

- **Kurtosis** is a measure of the *tailedness* of the probability distribution of a real-valued random variable, it can also be seen as the *heaviness* of the distribution tails. In normal distributions, the *kurtosis* value is `3`, so for any other distribution we measure **excess kurtosis**, defined as `kurtosis - 3`. According to that value, we can define `3` types of excess kurtosis:

- **Skewness** is a distortion or *asymmetry* that deviates from the *symmetrical bell curve*, or normal distribution, in a set of data. It has two types:

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: www.analyticsvidhya.com

Q27: What's the difference between *Binomial Distribution* and

Mid Probability 36

Answer

- The **Binomial distribution** describes the probability of obtaining *k* successes in *n Bernoulli experiments*, i.e an experiment which has only two possible outcomes, often call them *success* and *failure*. Its probability function describes the probability of getting exactly `k` successes in `n` independent *Bernoulli* trials:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- The **Geometric distribution** describes the probability of experiencing a *certain amount* of *failures* before experiencing the *first success* in a series of *Bernoulli experiments*. This probability is given by:

$$P(X = k) = p^k (1-p)^{n-k}$$

  So as we can see, the key difference is that in a *binomial distribution*, there is a *fixed number of trials* meanwhile in a *geometric distribution*, we're interested in the *number of trials* required until we obtain a *success*.

Having Machine Learning, Data Science or Python Interview? Check 👉
36 Probability Interview Questions

Source: www.statology.org

Q28: in which use-case we should use *Mean* and when to use *Median*?

Mid Statistics 59

Answer

The **arithmetic mean** is denoted as $\bar{x}$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

where each x_i represents a unique observation. The arithmetic mean measures the average value for a given set of numbers.

compared to the rest) which would *skew the mean*.

For example, salaries are usually discussed using medians. This is due to the large disparity between the majority of people and a very few people with a lot of money (with the few people with a lot of money being the outliers). Thus, looking at the 50% percentile individual will give a more representative value than the mean in this circumstance.

Alternatively, grades are usually described using the mean (average) because most students should be near the average and few will be far below or far above.

Having Machine Learning, Data Science or Python Interview? Check 👉
59 Statistics Interview Questions

Source: datascience.stackexchange.com

Q29: How could I (statistically) find *features* that are more important than others?

Senior | Statistics 59

Answer

Classification 43 | Azure ML 30 | NLP 38 | Gradient Descent 28 | Reinforcement Learning 55 | Curse of Dimensionality 14

23 Classification Interview Questions (ANSWERED) For ML Engineers And Data Scientists

Classification 43

Q30: How would you calculate a *Confidence Interval* for *non normally distributed* data?

Senior Statistics 59

Answer

Q31: How would you increase the *Statistical Power*?

Senior Statistics 59

Answer

missing data acceptable practice? Why or why not?

Answer

Q33: What's the difference between *Covariance* and *Correlation*?

Senior Statistics 59

Answer

23 Amazon & AWS Machine Learning Interview Questions (ANSWERED)

AWS Machine Learning 30

Q34: Which measures of

Senior Statistics 59

data?

Answer

Join **MLStack.Cafe** to open this Answer. It's Free!

Q35: How does an *ANOVA* test work?

Expert Statistics 59

Answer

23 Amazon & AWS Machine Learning Interview Questions (ANSWERED)

AWS Machine Learning 30

Amazone runs the internet as we know it. Amazon Web Services (AWS) offers a comprehensive suite of machine learning (ML) services that cater to various needs and expertise levels. Follow along and learn the 23 most common AWS machine-learning intervi...

23 Azure Machine Learning Interview Questions (ANSWERED)

Azure ML 30

Follow along and learn the 23 most common Azure Mach...

27 Hadoop Interview Questions (ANSWERED) For Your Next Big Data and ML Interview

Hadoop 32

Hadoop is an open-source big data processing framework. It leverages distributed computing to store and process large datasets in a fault-tolerant manner. According to recent reports, Apache Hadoop is one of the most sought-after big data skills with...

23 Apache Spark Interview Questions (ANSWERED) To Learn Before ML & Big Data Interview

Apache Spark is a unified analytics engine for large-scale data processing. It is built to handle various use cases in big data analytics, including data processing, machine learning, and graph processing. Follow along and learn the 23 most common an...

Apache Spark 30

Scala is a powerful language with functional programming capabilities that can be a good choice for data science, especially in big data and distributed computing scenarios. As an example, Apache Spark, a popular distributed data processing framework...

Scala 39

30 PyTorch Interview Questions (ANSWERED) To Beat Your Next Machine Learning Interview

PyTorch popularity as a Deep Learning framework of choice is on the rise. As of December 2022, 62% of the academic papers were implemented in PyTorch whereas only 4% were for TensorFlow. Follow along and prepare effectively with these key 30 PyTorch ...

PyTorch 36

23 Common Artificial Intelligence Interview Questions (SOLVED) For Your Next ML & DS Interview

The use of Artificial Intelligence (AI) in machine learning and data science enabled advancements in areas such as natural language processing, computer vision, recommendation systems, fraud detection, predictive analytics, and personalized medicine....

ChatGPT 42 Machine Learning 18

Neural Networks 140

21 Data Science Optimisation Algorithms Interview Questions (ANSWERED)

Optimization algorithms are extensively used in training machine learning models. Data engineers employ algorithms like gradient

Optimisation 33

27 ChatGPT Interview Questions (ANSWERED) For ML and Prompt Engineers

ChatGPT, an implementation of the GPT (Generative Pre-trained Transformer) model excels in understanding and generating human-like text, making it a powerful tool for NLP tasks. ML engineers and software developers can leverage ChatGPT's capabilities...

ChatGPT 42

27 Large Language Models (LLMs) Interview Questions (ANSWERED) For your Next ML and GPT Interview

Large Language Models (LLMs), such as GPT-3.5, have revolutionized natural language processing by demonstrating the ability to generate human-like text and comprehend context. Follow along to understand the top 27 LLMs-related interview questions and...

ChatGPT 42 LLMs 59

Load More...