Statistics Interview Questions

Statistics Question

- It is time to test your ability with Statistics!
- Keep in mind there is some cross over here with basic probability.
- A few of these questions require knowledge of distributions, use the resource guidebook to help you out!

Let’s get started!

Statistics Interview Question 1

Statistics Question

You're about to get on a plane to Seattle. You want to know if you should bring an umbrella. You call 3 random friends of yours who live there and ask each independently if it's raining.

Statistics Question

- Each of your friends has a 2/3 chance of telling you the truth and a 1/3 chance of messing with you by lying.
- All 3 friends tell you that "Yes" it is raining.

Statistics Question

- You also know that there is a 25% it’s raining on any given day in Seattle.
- What is the probability that it's actually raining in Seattle?

Statistics Interview Question 2

Statistics Question

A new quantum message system has a probability of 0.8 of success in any attempt to send a message through.

Calculate the probability of having 7 successes in 10 attempts.

Statistics Question

Hint: You may want to look at the Binomial Distribution to help you with this!

Statistics Interview Question 3

Statistics Question

What is the difference between Type I vs Type II error?

Statistics Interview Question 4

Statistics Question

- A new medical test for a virus has been created.
- 1% of the population has the virus.
- 99% of sick people with the virus test positive (indicating they have the virus).

Statistics Question

- 99% of healthy individuals test negative for the virus.
- If a patient tested positive, what is the probability that they have the virus?

Statistics Interview Question 5

Statistics Question

- The average life of a certain type of motor is 10 years, with a standard deviation of 2 years.
- If the manufacturer is willing to replace only 3% of the motors because of failures, how long a guarantee should she offer?

Statistics Question

- Assume that the lives of the motors follow a normal distribution.

Statistics Question

Quick Note: To fully solve this question you will need to look up some values in a z-table. Treat this as part of a take home task so you can use those resources!

SOLUTIONS

SOLUTIONS
ARE UP NEXT!

Solution to Statistics Interview Question 1

PIERIAN DATA

Statistics Question

You're about to get on a plane to Seattle. You want to know if you should bring an umbrella. You call 3 random friends of yours who live there and ask each independently if it's raining.

Statistics Question

Each of your friends has a 2/3 chance of telling you the truth and a 1/3 chance of messing with you by lying. All 3 friends tell you that "Yes" it is raining. (You also know that there is a 25% it's raining on any given day in Seattle). What is the probability that it's actually raining in Seattle?

Statistics Question

There are two situations where the friends all say “Yes.” One is if it is raining and they are all telling the truth.

Each tells the truth 2/3 of the time, so all three telling the truth is a probability of $(2/3)^3 = 8/27$.

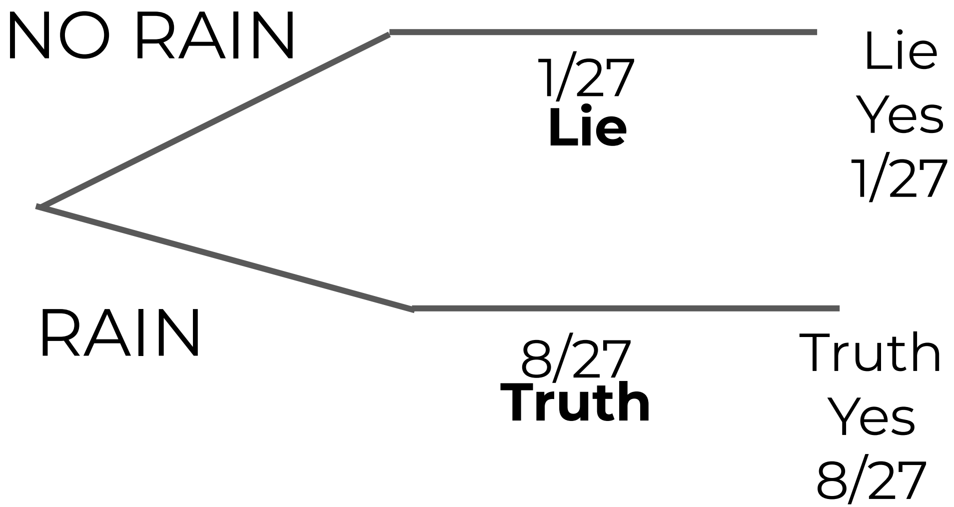
Statistics Question

The other situation is if it is not raining and all three are lying.

Each lies 1/3 of the time, so all three lying is a probability of $(1/3)^3 = 1/27$.

Statistics Question

Here is a probability tree of the situation if No Rain and Rain were equally likely.



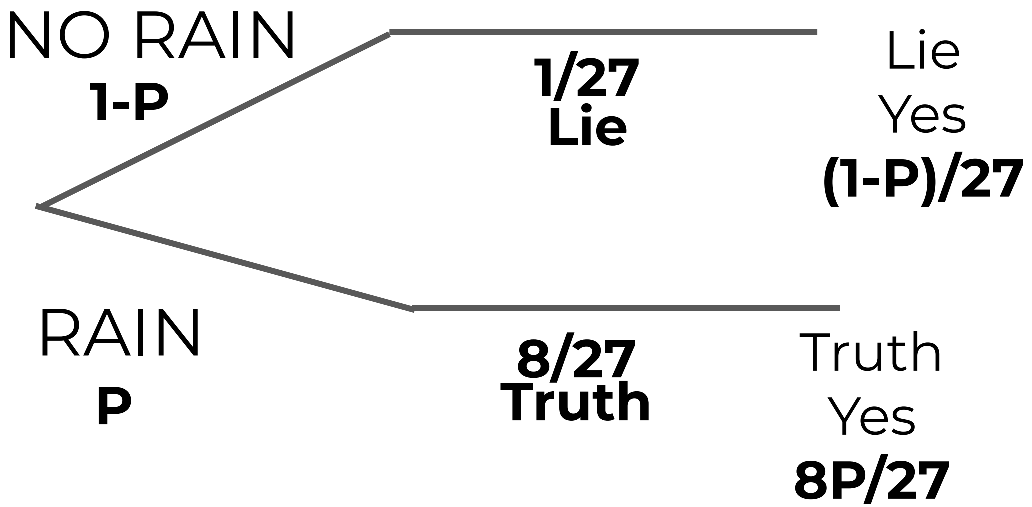
Statistics Question

The event “all three say yes” happens 1/3 = 8/27 + 1/27 of the time.

Out of these times, there is an 88.9% =>
8/9 = (8/27)/(1/3) chance that it is actually raining, and a 1/9 chance it is not raining.

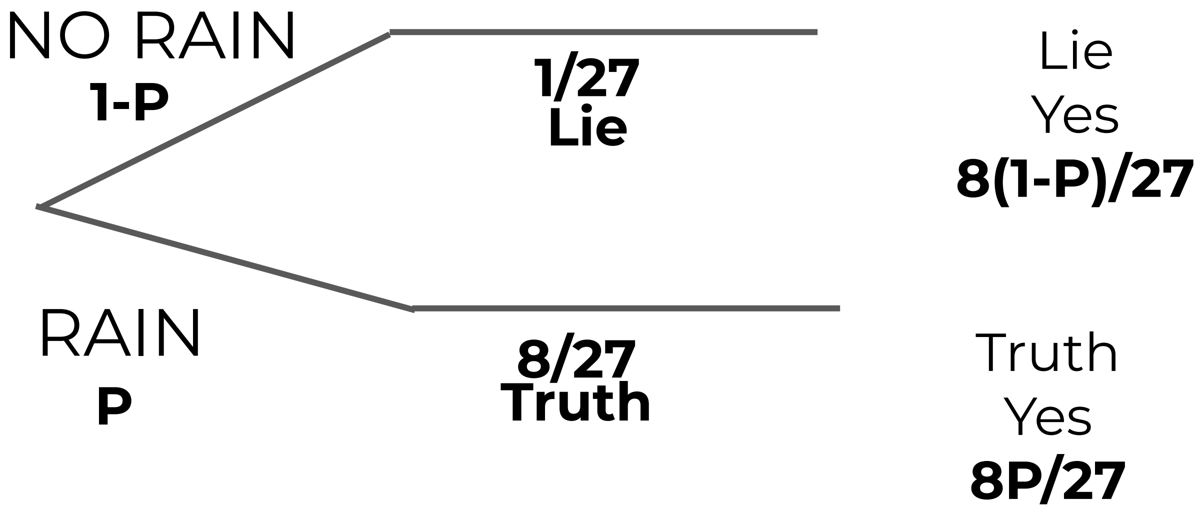
Statistics Question

However, we know that the probability of it raining on any given day was not 50/50, here is general case tree:



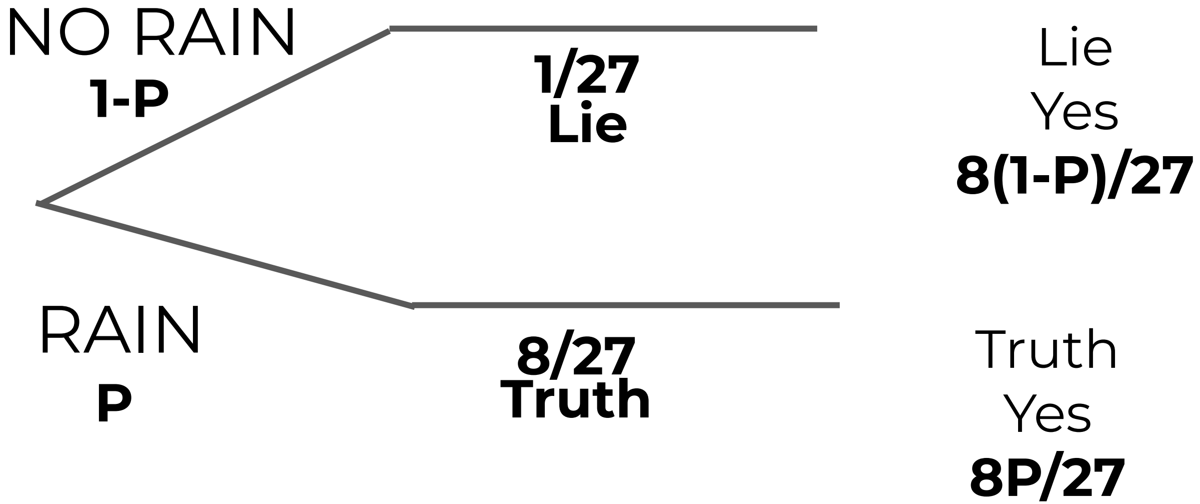
Statistics Question

This means the probability that they all say yes (for any situation) is (8P/27) + (8(1-P)/27) = (7p+1)/27.



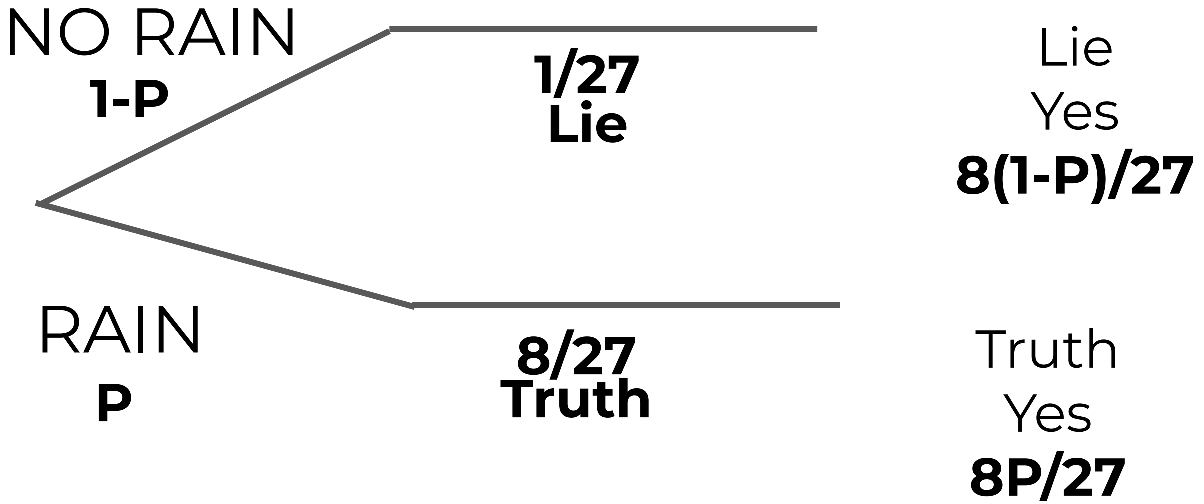
Statistics Question

Which means that out of those times it was actually raining (8P/27) / ((7p+1)/27) = 8P /(7P+1).



Statistics Question

Plugging in P=0.25 gives us 8P /(7P+1) → 0.72 chance that it is actually raining .



Solution to Statistics Interview Question 2

Statistics Question

A new quantum message system there was a probability of 0.8 of success in any attempt to send a message through.

Calculate the probability of having 7 successes in 10 attempts.

Statistics Question

Hint: You may want to look at the Binomial Distribution to help you with this!

Statistics Question

A binomial experiment is one that possesses the following properties:

- The experiment consists of n repeated trials;
- Each trial results in an outcome that may be classified as a success or a failure
- The probability of a success, denoted by p, remains constant from trial to trial and repeated trials are independent.

Statistics Question

The number of successes X in n trials of a binomial experiment is called a binomial random variable.

The probability distribution of the random variable X is called a binomial distribution, and is given by the formula:

$$Pr(k; n, p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for k = 0, 1, 2, ..., n, where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Statistics Question

- Probability of success **p=0.8** so **q=0.2**

$$\text{Probability} = P(X = 7)$$

$$= C_7^{10}(0.8)^7(0.2)^{10-7}$$

$$= 0.20133$$

Solution to Statistics Interview Question 3

Statistics Question

What is the difference between Type I vs Type II error?

Statistics Question

Table of error types		Null hypothesis (H_0) is	
		True	**False**
Decision About Null Hypothesis (H_0)	**Reject**	Type I error (False Positive)	Correct inference (True Positive)
	Fail to reject	Correct inference (True Negative)	Type II error (False Negative)

Example One

A type I error occurs when the null hypothesis (H0) is true, but is rejected. It is asserting something that is absent, a false hit.

A type I error may be likened to a so-called false positive (a result that indicates that a given condition is present when it actually is not present).

Example One

A type II error occurs when the null hypothesis is false, but erroneously fails to be rejected. It is failing to assert what is present, a miss.

A type II error may be compared with a so-called false negative (where an actual 'hit' was disregarded by the test and seen as a 'miss') in a test checking for a single condition with a definitive result of true or false.

Example One

Hypothesis: "Adding water to toothpaste protects against cavities."

Null hypothesis (H0): "Adding water to toothpaste has no effect on cavities."

This null hypothesis is tested against experimental data with a view to nullifying it with evidence to the contrary.

Example One

- A type I error occurs when detecting an effect (adding water to toothpaste protects against cavities) that is not present.
- The null hypothesis is true (i.e., it is true that adding water to toothpaste has no effect on cavities), but this null hypothesis is rejected based on bad experimental data.

Example Two

Hypothesis: "Adding fluoride to toothpaste protects against cavities."

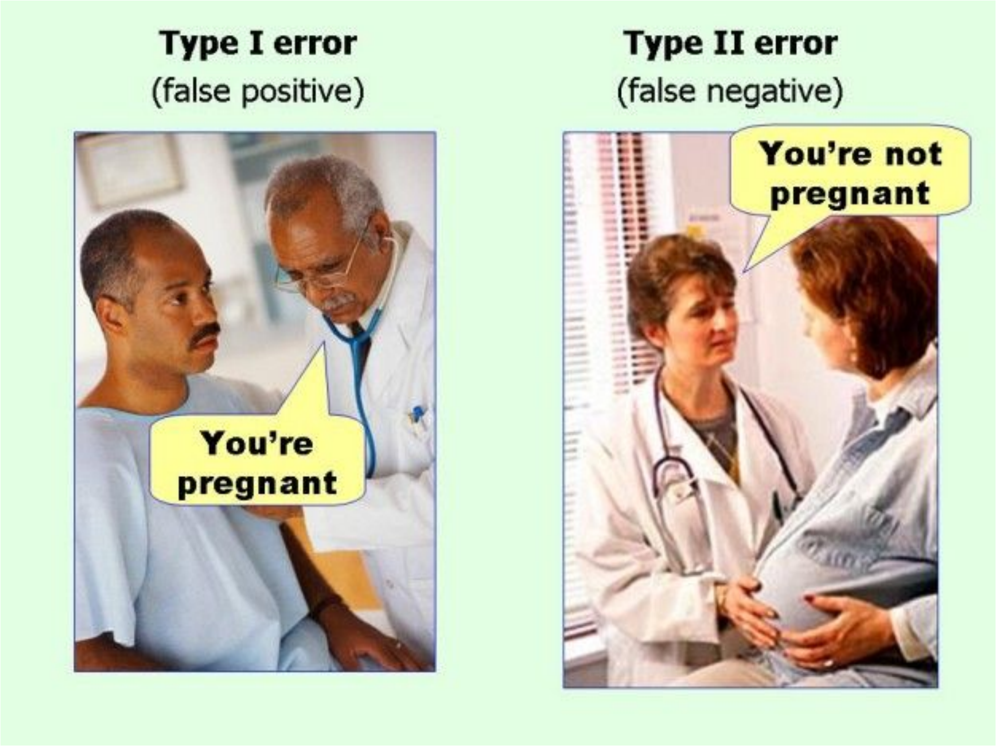
Null hypothesis (H0): "Adding fluoride to toothpaste has no effect on cavities."

This null hypothesis is tested against experimental data with a view to nullifying it with evidence to the contrary.

Example Two

- A type II error occurs when failing to detect an effect (adding fluoride to toothpaste protects against cavities) that is present.
- The null hypothesis is false (i.e., adding fluoride is actually effective against cavities), but the experimental data is such that the null hypothesis cannot be rejected.

Example Two


Type I error
(false positive)
You're pregnant
Type II error
(false negative)
You're not pregnant

Solution to Statistics Interview Question 4

Statistics Question

A new medical test for a virus has been created. 1% of the population has the virus. 99% of sick people with the virus test positive (indicating they have the virus). 99% of healthy individuals test negative for the virus. If a patient tested positive, what is the probability that they have the virus?

Statistics Question

This is yet another question where we can use Bayes' Rule (it is very popular in interview questions).

Let's walk through the steps.

Statistics Question

We know P(sick)=0.01, we also know that 99% of sick patients test positive, meaning P(positive|sick)=0.99. Also 99% of healthy patients test negative, so P(neg|not sick).

We want to solve for P(sick|positive).

Statistics Question

Using Bayes' Rule we have:

$$P(sick|pos) = P(sick)\frac{P(pos|sick)}{P(pos)}$$

Statistics Question

We also can calculate the probability

P(pos) = P(pos|sick)P(sick) + P(pos|not sick)P(not sick)

P(pos) = 0.99*0.01 + 0.01*(1-0.01) = 0.0198

Plugging this back into our previous Bayes' rule we get **P(sick|pos) = 0.5**

Solution to Statistics Interview Question 5

Statistics Question

The average life of a certain type of motor is 10 years, with a standard deviation of 2 years. If the manufacturer is willing to replace only 3% of the motors because of failures, how long a guarantee should she offer? Assume that the lives of the motors follow a normal distribution.

Statistics Question

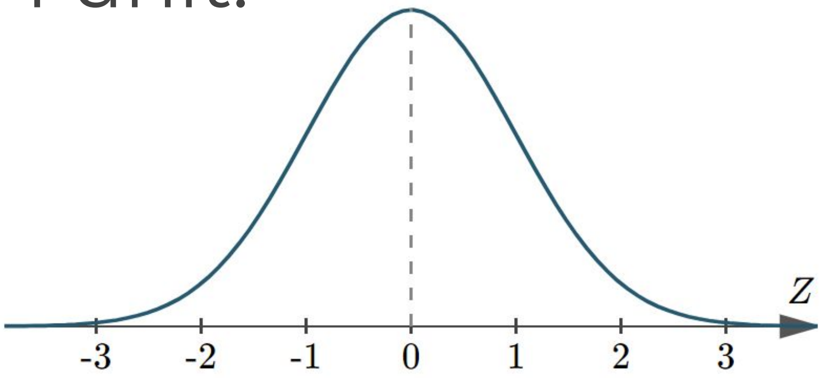
Quick Note: To fully solve this question you will need to look up some values in a z-table. Treat this as part of a take home task so you can use those resources!

Statistics Question

For this question, you will want to make sure you review normal distributions and how to do a look-up in a z-table. We will be using this in order to solve this problem.

Statistics Question

We can use a Standard Normal Curve of with a mean of zero and a standard deviation of 1 unit:



Standard Normal Curve $\mu = 0, \sigma = 1$

Statistics Question

We can transform all the observations of any normal random variable X with mean μ and variance σ to a new set of observations of another normal random variable Z with mean 0 and variance 1 using the following transformation:

$$Z = \frac{X - \mu}{\sigma}$$

Statistics Question

Here we can say **X** is the life of the motor and we are searching for the guarantee period ***x*.**

So we are searching for P(**X**<***x***) = 0.03. Using a z-table we find that the corresponding z-score of **z = -1.88**

Statistics Question

Which means we can then use:

$$Z = \frac{X - \mu}{\sigma}$$

Leaving us with:

$$\frac{x - 10}{2} = -1.88$$

Statistics Question

Solving this leads to x=6.24 , so the guarantee period should be 6.24 years.