Statistics Assignment

Student Name: Madhusudhan
Student email: maddymaster@gmail.com

Question 1:

The quality assurance checks on the previous batches of drugs found that — it is 4 times more likely that a drug is able to produce a satisfactory result than not.

Given a small sample of 10 drugs, you are required to find the theoretical probability that at most, 3 drugs are not able to do a satisfactory job

a. Propose the type of probability distribution that would accurately portray the above scenario, and list out the three conditions that this distribution follows.

Answer below:
I propose binomial probability distribution as it is either a yes or a no situation.
The conditions that this distribution follows are:
The binomial distribution describes the behavior of a count variable X if the following conditions apply:
1: The number of observations n is fixed.
2: Each observation is independent.
3: Each observation represents one of two outcomes ("success" or "failure").

B. Calculate the required probability: Answer below

Answer

(a) We use Binomial Distribution Here

(b) $P(X) = \frac{n!}{(n-X)!X!} * (P)^X . (q)^{n-X}$

$P = (1/5)$

$q = (4/5)$

$n = 10$

X from 0 to 3

$P(x <= 3)$

$P(X <= 3) = \frac{10}{(10-1)!10} \times \left(\frac{1}{5}\right)^{10} . \left(\frac{4}{5}\right)^{10-1}$

$P(0) = \binom{10}{0} = \frac{10}{10!(10)} \times \left(\frac{1}{5}\right)^{10} \times \left(\frac{4}{5}\right)^{10-0}$

$= 0.1 \times 1.02 \times 1.07$

$= 1.172$

$P(0) = 0.1074$

$P(1) = 0.268$

$P(2) = 0.301$

$P(3) = 0.201$

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

$\therefore \quad P(0) + P(1) + P(2) + P(3)$

Required probability $= 0.107 + 0.268 + 0.301 + 0.201$

$= \underline{\underline{0.877}}$

2. For the effectiveness test, a sample of 100 drugs was taken. The mean time of effect was 207 seconds, with the standard deviation coming to 65 seconds. Using this information, you are required to estimate the range in which the population mean might lie — with a 95% confidence level.

a.)Discuss the main methodology using which you will approach this problem.

State all the properties of the required method. Limit your answer to 150 words.

Answer:

Now there is a sample of 100 drugs, which is a good sample size for the 80,000 new products that exist. The Central Limit Theorem (CLT for short) basically says that for non-normal data, the distribution of the sample means has an approximate normal distribution, no matter what the distribution of the original data looks like, as long as the sample size is large enough (usually at least 30 and in our case its 100 so thats a good thing) and all samples have the same size. And it doesn't just apply to the sample mean; the CLT is also true for other sample statistics, such as the sample proportion

The central limit theorem states that if you have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population with replacement , then the distribution of the sample means will be approximately normally distributed. This will hold true regardless of whether the source population is normal or skewed, provided the sample size is sufficiently large (usually $n \geq 30$). If the population is normal, then the theorem holds true even for samples smaller than 30. In fact, this also holds true even if the population is binomial, provided that $\min(np, n(1-p)) \geq 5$, where n is the sample size and p is the probability of success in the population. This means that we can use the normal probability model to quantify uncertainty when making inferences about a population mean based on the sample mean.

For the random samples we take from the population, we can compute the mean of the sample means:

$$\mu_{\bar{X}} = \mu$$

and the standard deviation of the sample means:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

(a) We will use Central Limit theorem.

(b) Sun pharma creates 80,000 new products. It is not possible to go and test every product [illegible]

Sample $(X_1, X_2, X_3, \ldots, X_n)$

Note: Sample size $= n$

Sample Mean $\bar{X} = \sum_{i=1}^{i=n} X_i$

Sample Variance $\Rightarrow S^2$

$$= \sum_{i=1}^{i=n} \frac{(X_i - \bar{X})^2}{n-1}$$

Population = 80,000

mean = 2

variance = 1.6

Std deviation = 1.2649

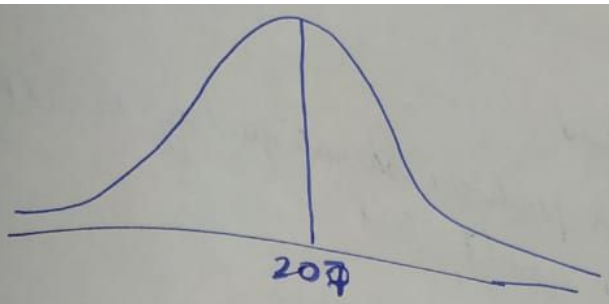
$n = 100 \rightarrow$ (sample size)

mean time = 207 [for sample of 100]

Std time deviation = 65 [—"—]

$$\sum_{i=1}^{i=n} (X_i - \bar{X})^2 = \qquad \mu_X = 100$$

$S^2 = 100$



Y% Confidence level = 95%

$$\bar{X} - \frac{2 * s}{\sqrt{n}} \; , \; \bar{X} + \frac{2 * s}{\sqrt{n}}$$

= 95% range $\rightarrow \pm 1.96 \approx 2.31$

Question 3:

Question 3:

a) The painkiller drug needs to have a time of effect of at most 200 seconds to be considered as having done a satisfactory job. Given the same sample data (size, mean, and standard deviation) of the previous question, test the claim that the newer batch produces a satisfactory result and passes the quality assurance test. Utilize 2 hypothesis testing methods to make your decision. Take the significance level at 5 %. Clearly specify the hypotheses, the calculated test statistics, and the final decision that should be made for each method.

Answer:

1. Making a decision - p-value method:

- Calculate the value of Z-score for the sample mean point on the distribution
- Calculate the p-value from the cumulative probability for the given z-score using the z-table
- Make the decision on the basis of the p-value with respect to the given value of α (significance level)

Question 3 (a)

(a) P-value method

H_0 → new batch produces ~~sb~~ satisfactory result and pass quality test

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \qquad 1.64$$

$$\Rightarrow \mu_{\bar{x}} = \mu = 100$$

2 decision can be made

H_0

- Sample mean from 200 far → Reject H_0 & accept H_1
- Sample mean close to 200 → H_0 cannot be rejected

Sample of $n = 100$

$$\frac{\sigma}{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$= 1.64$ and 1.07

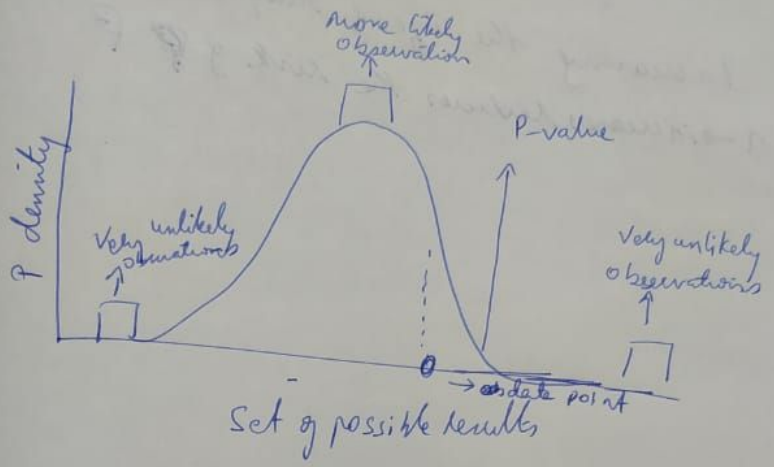
(b) P-

Pvalu

(8) P-value → probability if the null value is correct

Pvalue higher → Null hypothesis is true

lower → then its not true



z table → .9251

$= 1 - 0.9251$

$= 0.0749$

P-value (2 tailed) $= (0749) \times 2$

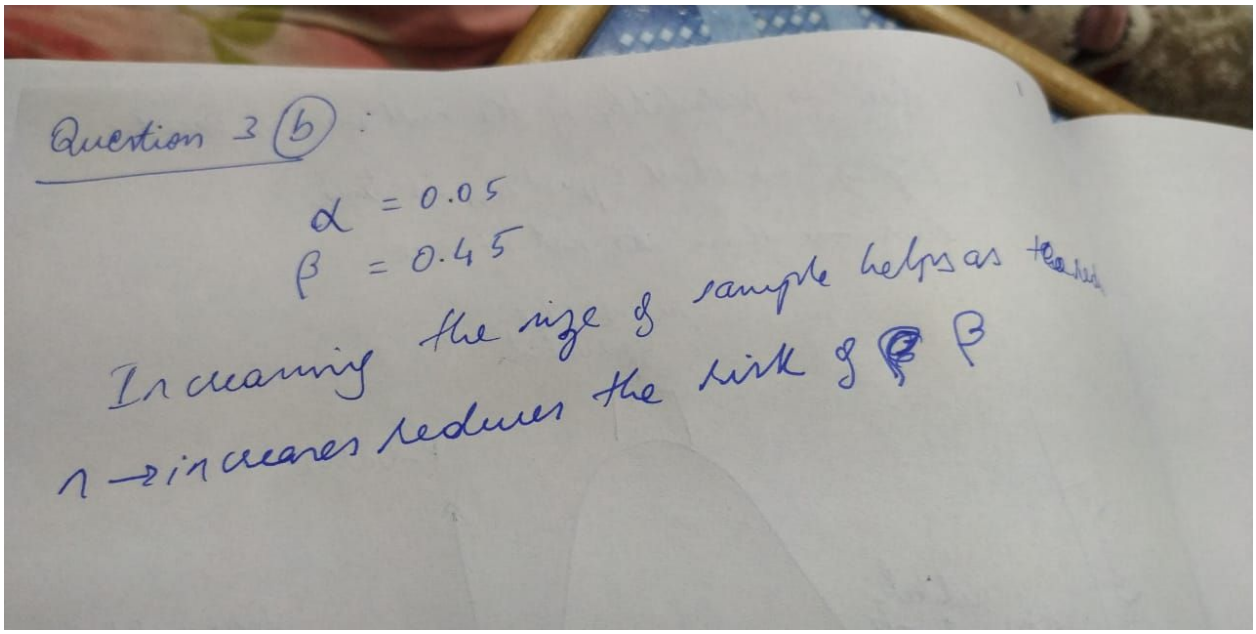
$= 0.1498$ (14.9%)

$\alpha = 0.05$ (5%)

14.9% > 5%

∴ We cannot reject the null hypothesis

B



1. Types of errors:

- Type-I error - Occurs when you reject a null hypothesis even when it is true
  - Its probability is represented by α
- Type-II error - Occurs when you fail to reject the null hypotheses even though it is false
  - Its probability is represented by β

Our two hypotheses have special names: the **null hypothesis** represented by H_0 and the **alternative hypothesis** by H_a. Historically, the null (invalid, void, amounting to nothing) hypothesis was what the researcher hoped to reject. These days it is common practice not to associate any special meaning to which hypothesis is which. (But this common practice may not yet have extended into behavioral science. The **research hypothesis**becomes the alternate hypothesis and the null hypothesis or "straw man" to be knocked down is so determined.) Although simple hypotheses would be easiest to test, it is much more common to have one of each type or perhaps for both to be composite. If the values specified by H_a are all on one side of the value specified by H_0, then we have a **one-sided test** (one-tailed), whereas if the H_a values lie on both sides of H_0, then we have a**two-sided test** (two-tailed). A one-tailed test is sometimes called a **directional test** and a two-tailed test is sometimes called a **nondirectional test**.

The outcome of our test regarding the population parameter will be that we either **reject** the null hypothesis or **fail to reject** the null hypothesis. It is considered poor form to "accept" the null hypothesis, although if we fail to reject it, that is in fact essentially what we are doing. When we reject the null hypothesis we have only shown that it is highly unlikely to be true---we have not proven it in the mathematical sense. The research hypothesis is **supported** by rejecting the null hypothesis. The null hypothesis locates the sampling distribution, since it is (usually) the simple

hypothesis, testing against one specific value of the population parameter. Establishing the null and alternative hypotheses is sometimes considered the first step in hypothesis testing.

Type I and Type II Errors

Two types of errors can occur and there are three naming schemes for them. These errors cannot both occur at once. Perhaps a table will make it clearer.

Reject\Truth	**H_0 True**	**H_a True**
Reject H_a	no error	False positive, Type II, beta=P(Reject H_a|H_a true)
Reject H_0	False negative, Type I, alpha=P(Reject H_0|H_0 true)	no error

The term **false positive** for type II errors comes from perhaps a blood test where the test results came back positive, but it is not the case (false) that the person has whatever was being tested for. The term **false negative** for type I errors then would mean that the person does indeed have whatever was being tested for, but the test didn't find it. When testing for pregnancy, AIDS, or other medical conditions, both types of errors can be a very serious matter. Formally, alpha=P(Accept H_a|H_0 true), meaning the probability that we "accepted" H_a when in fact H_0 was true. **Alpha** is the term used to express the **level of significance** we will accept. For 95% confidence, alpha=0.05. For 99% confidence, alpha=0.01. These two alpha values are the ones most frequently used. If our **P-value,** the high unlikeliness of the H_0, is less than alpha, we can reject the null hypothesis. Alpha and beta usually cannot both be minimized---there is a trade-off between the two. Ideally, of course, we would minimize both. Historically, a **fixed level** of significance was selected (alpha=0.05 for the social sciences and alpha=0.01 or alpha=0.001 for the natural sciences, for instance). This was due to the fact that the null hypothesis was considered the "current theory" and the size of **Type I errors** was much more important than that of **Type II errors.** Now both are usually considered together when determining an adequately sized sample. Instead of testing against a fixed level of alpha, now the P-value is often reported. Obviously, the smaller the P-value, the stronger the evidence (higher significance, smaller alpha) provided by the data is against H_0.

Example: On July 14, 2005 we took 10 samples of 20 pennies set on edge and the table banged. The resultant mean of heads was 14.5 with a standard deviation of 2.12. Since this is a small sample, and the population variance is unknown, after we calculate a t value and obtain t=6.71=(14.5-10)/(2.12/$\sqrt{}$(10)), we apply the t-test and find a P-value of either 8.73×10^{-5} or 4.36×10^{-5} depending on whether we do a one-tailed or two-tailed test. In either case our results are **statistically significant** at the 0.0001 level.

> The ***P*-value** of a test is the probability that the test statistic would take a value as extreme or more extreme than that actually observed, assuming H_0 is true.

Power of a Test

The **power** of a test against the associated correct value is 1-beta. It is the probability that a Type II error is not committed. **There is a different value of beta for each possible correct value of the population parameter.**It also depends on sample size (*n*), thus increasing the sample size increases the power. Power is thus important in planning and interpretting tests of significance. It is easy to misspeak power (1-beta) and *P*-value (alpha).

Setting the level of significance will correspond to the probability that we are willing to be wrong in our conclusion if a type I error was committed. That probability will correspond to certain area(s) under the curve of a probability distribution. Those areas, known as the **region of rejection** is bounded by a **critical value** or **critical values** which are often computed. Alternatively, one might compare the **test statistic** with the corresponding point(s) on the probability curve. These are equivalent ways of viewing the problem, just different units of measure are being used. In a one-tailed test there is one area bounded by one critical value and in a two-tailed test there are two areas bounded by two critical values. Which tail (left or right) under consideration for a one-tailed test depends on the direction of the hypothesis.

Establishing the significance level and the corresponding critical value(s) is sometimes considered the second step in hypothesis testing. Presumeably we have determined how the statistic we wish to test is distributed. This sampling distribution is the underlying distribution of the statistic and determines which statistical test will be performed.

Once the hypotheses have been stated, and the criterion for rejecting the null hypothesis establish, we compute the **test.** The test statistic for testing a null hypothesis regarding the population mean is a *z*-score, if the population variance is known (yeah right!). We used a *t*-score above, which is computing similarly, due to the small size of our sample and the fact that we do not know the population variance. We will have to examine other such test statistics and their underlying distributions. However, the same basic procedure always applies. This is considered by some step 3 in hypothesis testing.

Making a decision about H_0

The last step is whether we reject or fail to reject the null hypothesis. Although it is common to state that we have a small chance that the observed test statistic will occur by chance if the null hypothesis true, it is technically more correct to realize that the statement should refer to a test statistic **this extreme or more extreme** since the area under any point on the probability curve is zero. It can also be said that the difference between the observed and expected test statistic is too great to be attributed to chance sampling fluctuations. That is 19 out of 20 times it is too great---there is that 1 in 20 chance that our random sample betrayed us (given an alpha=0.05). Again, should we fail to reject the null hypothesis we have to be careful to make the correct statement, such as: the probability that a test statistic of blah would appear by chance, if the population parameter were blah, is greater than 0.05. Stated this way the level of significance used is clear and we have not committed another common error (that with 95% probability, H_0 is true).

Question 4:

Now, once the batch has passed all the quality tests and is ready to be launched in the market, the marketing team needs to plan an effective online ad campaign to attract new customers. Two taglines were proposed for the campaign, and the team is currently divided on which option to use.

Explain why and how A/B testing can be used to decide which option is more effective. Give a stepwise procedure for the test that needs to be conducted.

Answer:

To address these examples through A/B testing, we will first formulate your hypothesis of interest.

For both examples, the hypothesis is that the conversion rate for one of the three new page designs (ρ_N) is equal to the conversion rate for the current design (ρ_C). This called a null hypothesis and is denoted by H0. Therefore:

H0: $\rho_N = \rho_C$

Then you formulate your alternative hypothesis H1. The alternative hypothesis can be one of the three possible forms:

H1: ρ_N not equal ρ_C (less or greater)

or

H1: ρ_N greater than ρ_C (right-sided) – one of the variations will cause an uplift in conversions

or

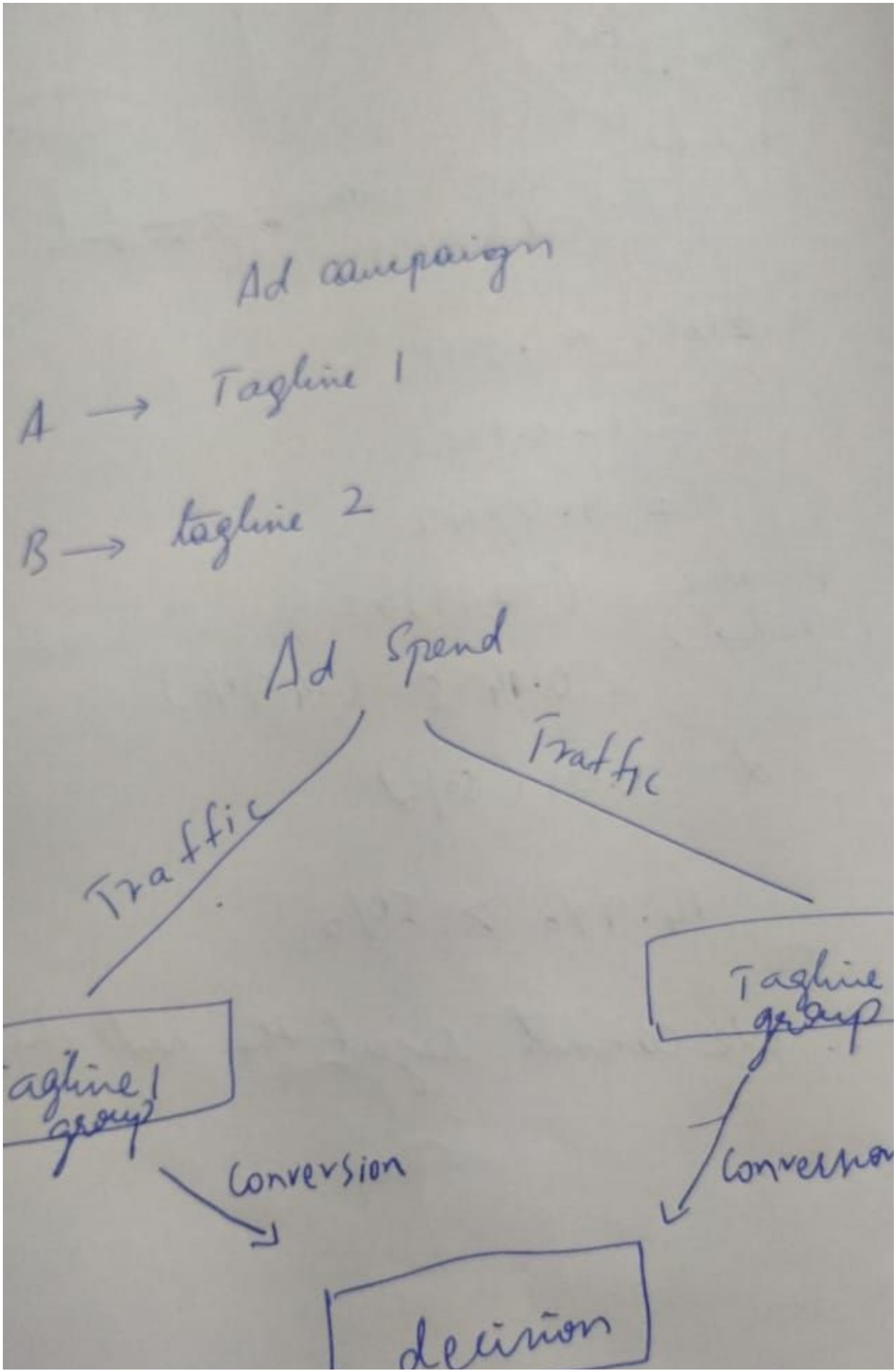
H1: ρ_N less than ρ_C (left-sided) – one of the variations will cause a drop in conversions

In the first case, you do not assume that the new design will generate an increase or decrease in conversions. The first type of test is called a two-tailed test.

In the last two cases, you assume the impact of the new design. This is referred to as a one-sided test (left-or right-sided)

Question 4:

A/B Testing Process


Ad campaign
A → Tagline 1
B → tagline 2
Ad Spend
Traffic
Traffic
Tagline group
Conversion
decision

The following is an A/B testing framework we can use to start running tests:

- Collect Data:
- Identify Goals:
- Generate Hypothesis:
- Run Experiment:
- Analyze Results: Once your experiment is complete, it's time to analyze the results. Your A/B testing software will present the data from the experiment and show you the difference between how the two versions of your page performed, and whether there is a statistically significant difference.

If your variation is a winner, congratulations we have selected which tagline to go with!