Problem Statement

The pharmaceutical company Sun Pharma is manufacturing a new batch of painkiller drugs, which are due for testing. Around 80,000 new products are created and need to be tested for their time of effect (which is measured as the time taken for the drug to completely cure the pain), as well as the quality assurance (which tells you whether the drug was able to do a satisfactory job or not).

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.

b.) Calculate the required probability.

Solution-1.a: The test performed on the each drug can produce the following outcomes

- Satisfactory result(Success)
- Non-satisfactory result(fail)

The quality assurance check is an event where the probability of the questions (on drug test)

Whether it is no(fail) or yes(success) is same for all questions indicates this is a discrete variable.

for such kind of event(with fixed sample),we can find the probability using **Binomial distribution**

Following are the conditions that this distribution follows

1. Total number of trials is fixed at n
2. Each trial is binary, i.e., has only two possible outcomes - success or failure
3. Probability of success is same in all trials, denoted by p

Solution-1.b: The formula for finding binomial probability is given by -

$P(X=r) = nCr(p)r(1-p)n-r$

X = probability that drug is ineffective (non-satisfactory result)

4X= probability that drug produce a satisfactory result

X+4X =1

X=1/5=0.2=p

Number of trails =n=10

Theoretical probability that at most, 3 drugs are not able to do a satisfactory job

P(X<=3) = P(X=0) + P(X=1)+ P(X=2)+ P(X=3)

=[10C0*(0.2)^0*(0.8)^10]+[10C1*(0.2)^1*(0.8)^9]+[10C2*(0.2)^2*(0.8)^8]

+[10C3*(0.2)^3*(0.8)^7]

=0.87

P(X<=3) = 87%

Question 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.

b.)Find the required range.

Solution2.a:

Effectiveness is measured by “time of effect” which is a continuous variable

and the probability of this variable is measured in terms of intervals rather than

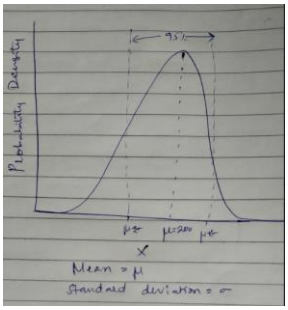
exact values.

Probability Density Function:PDF:This is used in which the area under the curve

Gives the cumulative probability

As the sample size is 100(n>30) this sampling distribution becomes

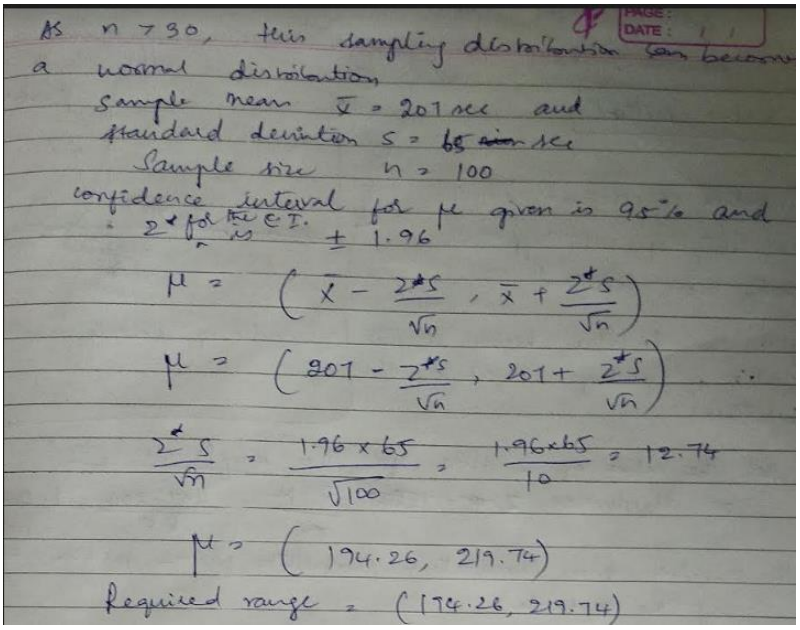
Normal distribution



Solution2.b: Standard deviation = 65 seconds

Sample mean = 207 sec

population mean with a 95% confidence level.



population mean lie in the range.(194.26,219.74)

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.

Solution3.a:

Claim: Status quo: Newer batch produces a satisfactory result and passed the quality test

Of satisfactory job that a painkiller drug does have a time of effect of atmost 200 sec

Hypothesis testing methods we will be discussing

1. Critical value method
2. P-value method

Null Hypothesis-Ho-Newer batch produces a satisfactory result

Ho<=200sec

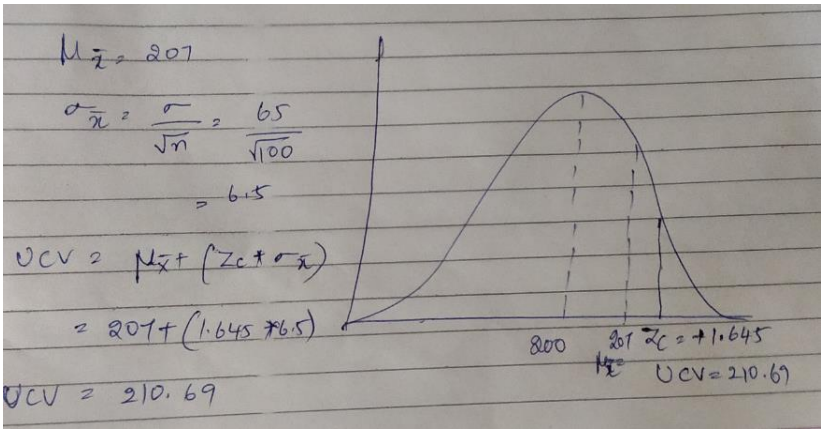
Alternate Hypothesis-H1- Newer batch produces a satisfactory result and

Doesnot pass the quality test

H1>200 sc

Significance level for the test is 5%=σ =0.05

Critical Value method(CVM):



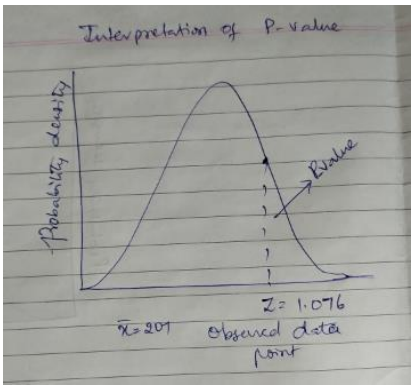
Sample mean doesnot fall in the critical region .

So **Decision**: Fail to reject the null hypothesis.

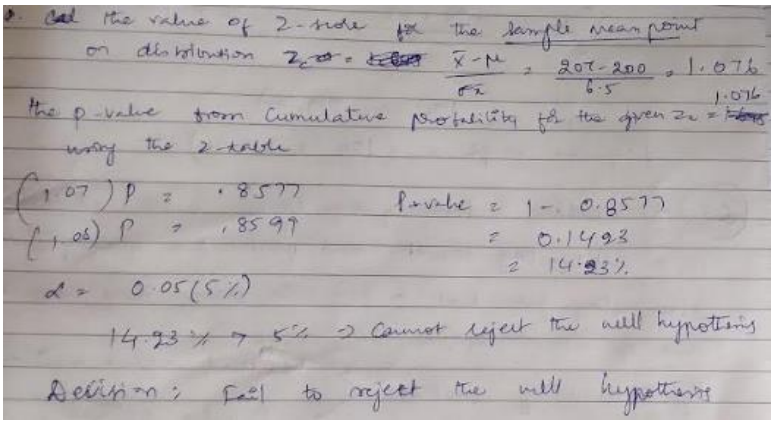
P-Value method:

P-value is defined as the probability of null hypothesis being not rejected.

Higher the p value higher is the probability of Failing to reject a null hypothesis



The value of Z score for the sample mean point on distribution



P-value = 0.1423=14.23%>5%

Decision: Fail to reject the null hypothesis.

Question 3.b

b) You know that two types of errors can occur during hypothesis testing — namely Type-I and Type-II errors — whose probabilities are denoted by α and β respectively. For the current hypothesis test conditions (sample size, mean, and standard deviation), the value of α and β come out to 0.05 and 0.45 respectively.

Now, a different sampling procedure is proposed so that when the same hypothesis test is conducted, the values of α and β are controlled at 0.15 each. Explain under what conditions would either method be more preferred than the other.

Solution3.b:

Type 1 error: Ho is true but is rejected

Newer batch produces a satisfactory result but it didn't pass the quality test

Type 2 error: Ho is not true but we fail to reject it

Newer batch produces non-satisfactory result but passes the quality test

In the given hypothesis test conditions

α = 0.05,Type 2 error(β)=0.45

when type 2 error is more ,more drugs reaches the patients

And in turn in-effective drugs are reached.

As this is new drugs reaching market it will create a bad reputation to the company

More damage than the revenue loss (which is caused by type 1 error)

For the different sampling procedure, values are controlled at

α = 0.15, Type 2 error(β)=0.15

Here type 1 error is increased and type 2 error is reduced,this clearly indicates

Company reputation is more prioritized than their revenue.

When we compare both the sampling procedures,for the launch of this drugs

Second sampling procedure is more preffered than the initial method.

Once the drug captures the market with the trust of customers,

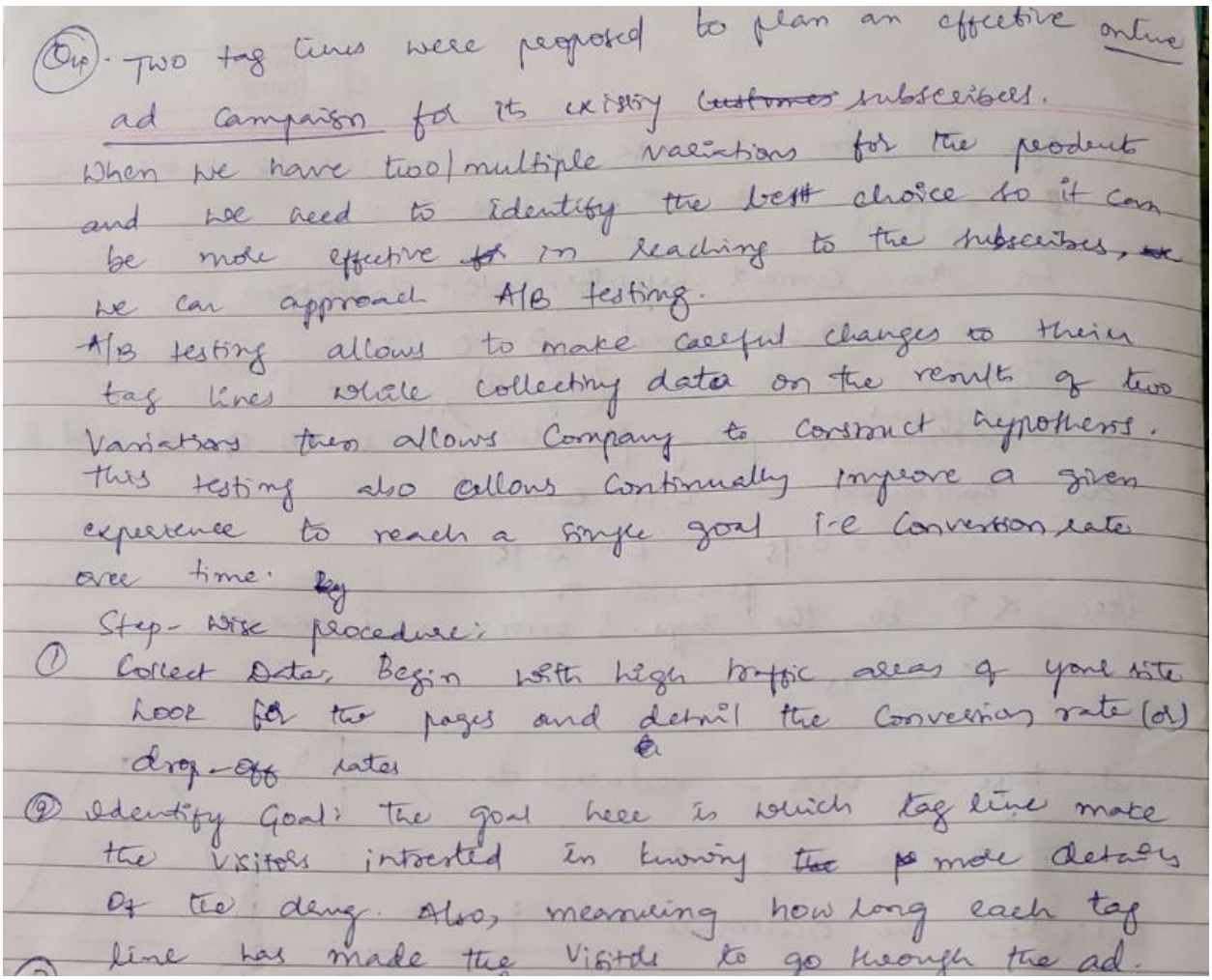
Type 1 and type 2 errors can be further re-designed.

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 for its existing subscribers. 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.

Solution 4:



(3) Generate Hypothesis: Begin generating A/B testing ideas and hypothesis i.e how many variations of tag lines are proposed and finalized for A/B testing and hypothesis for why you think they will be better than each other.

(4) Create Variations: i.e create the tag line and the ad campaign for the both the variations

(5) Run Experiment: Kick off the experiment to the traffic by dividing the two subscribers and wait for them to participate. Their interaction with each variation is measured.

(6) Analyze Results: Once Experiment completes, analyze the results in defining understanding which tag line has performed better and some times these results also helps us ~~[illegible]~~ getting better insights of the variations.