Statistics Assignment 26 August 2019

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.

Answer:
The probability distribution that would accurately portray the above scenario will be binomial distribution as this event is of type success or failure, i.e., either the drug is able to produce a satisfactory result or unsatisfactory result.

The three conditions that this distribution follows is:

1. Total number of trials is fixed at n.
2. Each trial is binary, i.e., there are only 2 possible outcomes – success or failure.
3. Probability of success is same for all the trials denoted by p.

b) Calculate the required probability.

Answer:
If x is defined as the number of drugs producing unsatisfactory result after testing 10 drugs, then x would follow a binomial distribution
Here n = 10
Let x be the probability for unsatisfactory result.

Since in this case, it is defined that it is 4 times more likely that a drug is able to produce a satisfactory result than not. Therefore,

So, x+4x = 1, as [Total Probability is always one]
5x = 1

Therefore, $x = 1/5 = 0.2$

Probability of drug produces unsatisfactory result is 0.2

Now we have to calculate the theoretical probability that at most, 3 drugs are not able to do a satisfactory job can be found out using the cumulative probability of X, denoted by F(x), which is the probability that the random variable X takes a value less than or equal to x.

Therefore, $F(x) = P(X<=x)$

$F(3) = P(X<=3)$

$F(3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

Formula for Binomial distribution is
$P(X=r) = nCr(p)^r * (1-p)^(n-r)$
Where, n = number of trials=10, p = probability of expected result = 0.2

As shown in below attached screenshot, a probability that at most 3 drugs are unable to do satisfactory job is 87.91%

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

$$P(X=0) = {}^{10}C_{0}\,(0.2)^{0}\,(0.8)^{10-0}$$

$$= \frac{\cancel{10!}}{\cancel{10!} * 0!} * 1 * 0.107374$$

$$= 0.107374$$

$$P(X=1) = {}^{10}C_{1}\,(0.2)^{1}\,(0.8)^{9}$$

$$= \frac{10 * \cancel{9!}}{\cancel{9!} * 1!} * 0.2 * 0.134218$$

$$= 0.268436$$

$$P(X=2) = {}^{10}C_{2}\,(0.2)^{2}\,(0.8)^{8}$$

$$= \frac{10 * 9 * \cancel{8!}}{\cancel{8!} * 2!} * 0.04 * 0.167772$$

$$= 45 * 0.04 * 0.167772$$

$$= 0.3019896$$

$$P(X=3) = {}^{10}C_{3}\,(0.2)^{3}\,(0.8)^{7}$$

$$= \frac{10 \times 9 \times 8 \times \cancel{7!}}{\cancel{7!} \times 3!} * 0.008 * 0.209715$$

$$= 0.2013264$$

$$\therefore P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= 0.107374 + 0.268436 + 0.301989 + 0.2013264$$

$$= 0.87912$$

$$= 87.91\ \%$$

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.

Answer:

The main methodology that we would be using Central Limit Theorem to estimate the population mean in the form of an interval.

Hence based on the Central Limit Theorem, for a sampling distribution, we can say that

1. Sampling distribution's mean ($\mu_{\bar{x}}$) = Population mean(μ) {unknown}
2. Sampling distribution's standard deviation

(Standard Error) = $\sigma/\sqrt{n}$ is approximately equal to $S/\sqrt{n}$

Since we know only the samples standard deviation (S), we approximate the population's standard deviation (σ) with that of sample. n is the sample size

3. For n > 30, sampling distribution becomes a normal distribution

Given the sample's size, mean and standard deviation, we can say that the confidence interval for μ lies in the range of (X^ − [($Z*S$)/√n], X^ +[($Z*S$) /√n])

Here z* is the z-score associated to 95% of the confidence level, X^ is the sample's mean and S is the sample's standard deviation.

b) Find the required range.

Answer:

As shown in below attached screenshot the margin of error corresponding to 95% confidence level is [($Z*S$)/√n], i.e. (1.96 * 65)/√100) = 12.74 and the population mean lies between 194.26 seconds and 219.74 seconds

Here,

Sample size (n) = 100,

Sample mean ($\hat{x}$) = 207

Sample standard deviation (S) = 65

For confidence level of 95% Z^* (Z-score) is ± 1.96

Confidence interval is —

$$= \hat{x} - \cancel{Z^* S}\ \frac{Z^* S}{\sqrt{n}}, \hat{x} + \frac{Z^* S}{\sqrt{n}}$$

$$= 207 - \frac{1.96 * 65}{\sqrt{100}}, 207 + \frac{1.96 * 65}{\sqrt{100}}$$

$$= 207 - \frac{127.4}{10}, 207 + \frac{127.4}{10}$$

$$= 207 - 12.74, 207 + 12.74$$

$$= (194.26, 219.74)$$

Therefore, The range in which the population mean might lie with 95% confidence level is 194.26 seconds – 219.74 seconds

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:

Here, the Null and Alternate Hypothesis are
$H0$: $\mu \leq 200$ seconds, the time of effect that the drug needs to do satisfactory job
$H1$: $\mu > 200$ seconds, the time of effect that the drug needs to not do satisfactory job

The Type of Test can be deduced as:
Since > sign is used in alternate hypothesis, it would be a One-tailed test (upper-tailed test) and the rejection region would be on the right side of the distribution.

Sample size n = 100

Assumed Sample mean μ = 200

Sample mean $\mu\overline{x}$ = 207

Sample standard deviation $\sigma\overline{x}$ = 65. (Since only the samples standard deviation is given, we approximate the population standard deviation to samples standard deviation i.e. 65)

Significance level α = 5% i.e. 0.05

<u>First Method - Based on critical value test</u>
This is a one-tailed test. Therefore, for a significance level of 5%, there would be only one critical region that is on the right side with a total area of 0.05.

Therefore, the cumulative probability till the critical area would be 1 – 0.05 = 0.95

Since the number of samples is 100 (which is > 30), we will use the z-test.

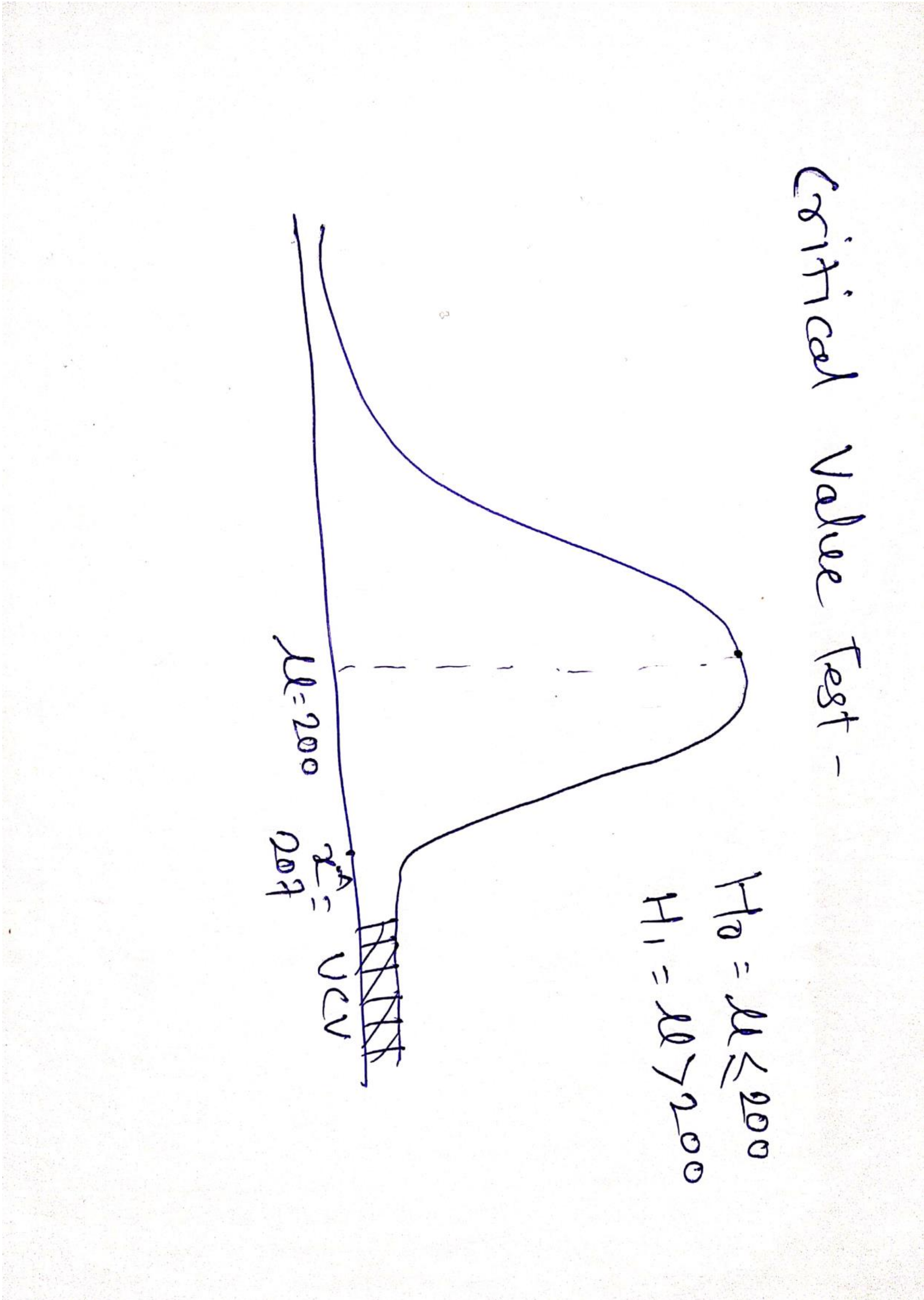
The Z-score for 0.95 from the z-table is 1.645 (average of 1.64 and 1.65 from z-table)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633

Now, the Upper Critical Value (UCV) can be calculated as

$\mu + ZC * (\sigma/\sqrt{n})$
UCV = 200 + 1.645 * (65/√100)
= 200 + (1.645 * (65/10))
= 200 + (1.645 * 6.5)
= 200 + 10.69
= 210.69
Upper Critical Value (UCV) = 210.69 seconds

As shown in below screen shot, since the sample mean 207 seconds is less than the Upper critical value 210.69 seconds, we fail to reject the null hypothesis which states that the drug needs a time of effect <= 200 seconds to do a satisfactory job



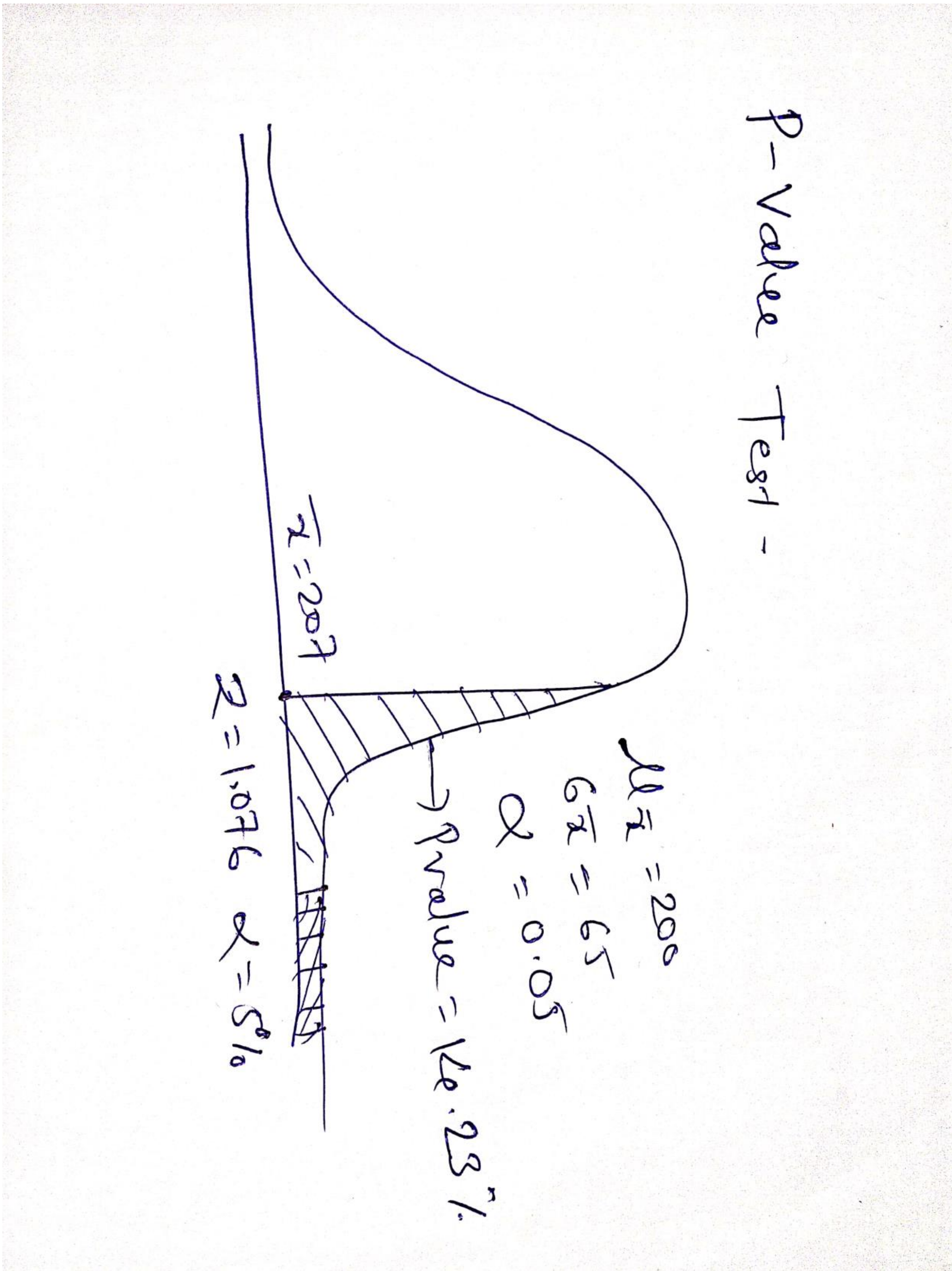
Second Method -Based on P-value test

Z-Score for sample mean 207 can be given by $(\bar{X} - \mu) / (\sigma/\sqrt{n})$
Z = (207 – 200) / (65/√100)

$Z = (207 - 200) / (65/10)$
$Z = 7/6.5$
$Z = 1.076$
Since the sample mean lies on the right side of the hypothesized mean of 200 seconds, the z-score comes out to be positive.
The p-value for the z-score of 1.076 (corresponding to the sample mean of 207) will be 0.8577

Since the sample mean is on the right side of the distribution and this is a one-tailed test, the p-value would be (1 - 0.8577) = 0.1423 (14.23%)

As shown in below screen shot, since the p-value is greater than the significance level ($0.1423 > 0.05$), we fail to reject the null hypothesis which states that the drug needs a time of effect <= 200 seconds to do a satisfactory job



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 sample conditions (sample size, mean, and standard deviation), the value of α and β come out to be 0.05 and 0.45 respectively.

Now, a different sampling procedure(with different sample size, mean, and standard deviation) 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, i.e. give an example of a situation where conducting a hypothesis test having α and β as 0.05 and 0.45 respectively would be preferred over having them both at 0.15. Similarly, give an example for the reverse scenario - a situation where conducting the hypothesis test with both α and β values fixed at 0.15 would be preferred over having them at 0.05 and 0.45 respectively. Also, provide suitable reasons for your choice(Assume that only the values of α and β as mentioned above are provided to you and no other information is available).

Answer:

Situation 1 : A situation where conducting the hypothesis test with α as 0.05 and β as 0.45 is preferred. In other word a decrease in Type 1 error is preferred.
Example:
A business man is looking out to start new branches for his catering business. He is on the look out to shortlist cities for the same. He plans to conduct survey among the residents of few cities and decides to open the branch only in cities where there is an evidence of high demand. The null hypothesis in this case would the demand isn't high enough. The alternate hypothesis in this case would be the demand is high enough. He decides to get the survey done by 2 different agencies say Agency 1 and Agency 2. Based on the survey, the type 1 and type 2 error are as given below:
Agency 1 Survey:

- Type 1 Error - Has 15% chance (α = 0,15) of showing evidence that he chooses a city where the demand isn't actually high.
- Type 2 Error - Has 15% (β = 0.15) chance of showing evidence that he doesn't choose a city where the demand is actually high.

Agency 2 Survey:

- Type 1 Error - Has 5% chance (α = 0,05) of indicating evidence that he chooses a city where the demand isn't actually high.
- Type 2 Error - Has 45% (β = 0.45) chance of indicating evidence that he doesn't choose a city where the demand is actually high.

Consequence of Type 1 error is that he chooses a city where the demand isn't actually high while the consequence of the Type 2 error is that he doesn't choose a city where the demand is actually high. The business man feels that the type 1 error is more costly for his business as he would not want to incur losses by opening a branch in a city which doesn't have much demand and hence opts for the survey of Agency 2 (α = 0.05, β = 0.45) with a lower significance level of α which reduces the probability of type 1 error.

Situation 2 : A situation where conducting the hypothesis test with both α and β values controlled or fixed at 0.15 is preferred. In other words a decrease in Type 2 error is preferred.
Example:
A woman patient visits a doctor for pregnancy test. The doctor believes that she is not pregnant until a medical test is performed and it proves otherwise. The doctor is testing for the null hypothesis that the woman patient is not pregnant. The alternate hypothesis in this case would be that the woman patient is pregnant. Say there are two tests available for pregnancy, Test A and Test B.
Test A:

- Type 1 Error - Has 5% chance (α = 0,05) of showing evidence for pregnancy when the woman patient is infact not pregnant.

- Type 2 Error - Has 45% chance (β = 0.45) of indicating evidence for no pregnancy when in fact the woman patient is pregnant.

Test B:

- Type 1 Error - Has 15% chance (α = 0,15) of showing evidence for pregnancy when the woman patient is infact not pregnant.
- Type 2 Error - Has 15% chance (β = 0.15) of indicating evidence for no pregnancy when in fact the woman patient is pregnant.

The consequence of type 1 error is, diagnosing a woman patient as pregnant when in fact she is not, whereas the consequence of type 2 error is that the woman patient is told she is not pregnant when in fact she is. If the doctor believes that the type 2 error has more serious consequence since it risks the life of the woman patient and the baby than the type 1 error, then the Test B (α = 0.15, β = 0.15) should be preferred, as using a higher significance level increases the probability of Type 1 error, but decreases the probability of Type 2 error.

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:

A/B testing is a way to compare two versions of a single variable and determine which one performs better than other in order to improving the sales and revenue. In other words, A/B testing provides a way to test two different versions of the same element and see which one performs better. As there are two taglines proposed, we can show one tagline for a specific set of audience (controlled version) and another tagline for another set of audience (variant version). As different set of audience are

shown different set of ads, their responses to these 2 ads can be measured statistically and thereby used to determine which ad campaign gets converted into sales and generates revenue.
Why A/B Testing:
Below given are few of the reasons why A/B testing should be done from a generic perspective:

- Enables to solve visitor pain points. This can be done using data gathered through visitor behaviour analysis tools such as heat maps, Google analytics, and surveys to solve the visitors' pain points.
- Get more conversion while investing less. ROI from A/B testing can be significant with minor changes resulting in a significant increase in conversions
- A/B testing is completely data driven with no room for guesswork and hence helps to easily determine a "winner" and a "loser" based on various metrics.
- A/B testing allows for maximum output with minimal modifications, resulting in increased ROI.
- Any changes on product prices should be done after A/B testing as it will let know the customers receptiveness to the same.

Stepwise Procedure:
1) Performing Research:

Make observations to identify the problem area using surveys, analytics tools etc. in the conversion funnel and find out what is stopping visitors from converting once they view the ads.

2) Formulate Hypothesis:

Based on the above research insights, a hypothesis should be built. The hypothesis can be arrived at by determining what should be the final result, statistics on the user behaviour to which type of ad etc. The hypothesis in this case should be built with the main purpose of increasing conversions.

3) Determine the test sample size:

An advance calculation of how many views need to be generated should be determined so that the test results are statistically significant.

4) Determine the duration of the experiment:

The test should last at least a week even if we have achieved the sample size in 2 days. The recommended testing time generally is around 10 – 14 days. Once the duration is decided, the test shouldn't be stopped before the duration ends. In this case since there are 2 banners, the test duration can be calculated keeping in mind the monthly visitors, current conversion rate, and the expected change in the conversion rate.

5) Testing, Analysing results and drawing conclusions:

Flag off the test and wait for the stipulated time for achieving statistically significant result. Once the testing is done, analyse the test results and, if it succeeds, deploy the winning variation. If the test remains inconclusive, draw insights from it and implement these in subsequent test.