Assignment: Inferential and Hypothesis Testing

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.

Answer 1:

a.)

The binomial probability distribution is an important probability model that is used when there are two possible outcomes. Here in the above case, we have two outcomes: one is drug is able to produce satisfactory result and the second is drug is not able to produce satisfactory result. Binomial distribution uses the following formula:

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

Where **n** is **no. of trials**, **p** is **probability of success** and **r** is **no. of successes after n trials**.

Following are some **conditions** that need to be followed in order for us to be able to apply the binomial probability distribution formula.

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

As we can observe from the question, binomial probability distribution accurately portrays the above scenario and follows the conditions of binomial distribution. Hence, we will use binomial distribution to solve the question.

b.)

probability that the drug is unable to produce a satisfactory result: p= 1/5

probability that the drug is able to produce a satisfactory result: q= 4/5

sample size (n) = 10

r = no. of successes after n trials

the theoretical probability that at most, 3 drugs are not able to do a satisfactory job is as follows:

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

$$= {}^{10}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10} + {}^{10}C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^9 + {}^{10}C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^8 + {}^{10}C_3 \left(\frac{1}{5}\right)^3 \left(\frac{4}{5}\right)^7$$

$$= \left(\frac{4}{5}\right)^{10} + \frac{10}{5}\left(\frac{4}{5}\right)^9 + \frac{45}{25}\left(\frac{4}{5}\right)^8 + \frac{120}{125}\left(\frac{4}{5}\right)^7$$

$$= 0.87912$$

$$\therefore P(X \leq 3) = 0.87912$$

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.

Answer 2:

a.)

Instead of finding the mean and standard deviation for the entire population, it is sometimes beneficial to find the mean and standard deviation for only a small representative sample because of time and/or money constraints.

The sampling distribution, which is basically the distribution of sample means of a population, has some interesting properties which are collectively called the **central limit theorem**, which states that no matter how the original population is distributed, the sampling distribution will follow these three properties:

1. **Sampling distribution's mean** ($\mu_{\bar{X}}$) = **Population mean** (μ)
2. Sampling distribution's standard deviation (**Standard error**) = $\frac{\sigma}{\sqrt{n}}$, where σ is the population's standard deviation and n is the sample size
3. **For n > 30**, the sampling distribution becomes a **normal distribution**

b.)

sample size = n

sample mean = $\overline{X}$

standard deviation of sample = S

$n = 100$

$\bar{x} = 207$

$S = 65$

confidence level = 95%

z^* for 95% = 1.96

$$\text{confidence interval} = \bar{x} \pm \frac{z^* S}{\sqrt{n}}$$

$$= 207 \pm \frac{1.96 \times 65}{\sqrt{100}}$$

$$= 207 \pm 12.74$$

$$= (194.26,\ 219.74)$$

Hence, the range in which the population mean might lie — with a 95% confidence level is (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.

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.

Answer:

a.)

for the above scenario, we will consider the following Null Hypothesis:

H_0 = Painkiller drug have a time effect of ≤ 200

And the Alternate Hypothesis is

H_1 = Painkiller drug have a time effect of > 200

sample size = n = 100

sample mean = $\overline{X}/\mu$ = 207

standard deviation of sample = S = 65

significance level = 5%

cumulative probability of the critical point = 0.95

$$Z_c = 1.645$$

$$\text{critical value} = \mu \pm Z_c\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}} = \frac{65}{\sqrt{100}} = 6.5$$

$$\text{critical value} = 207 + (1.645 \times 6.5)$$
$$= 217.69$$

Now,

As (217.69 > 207)

That is- critical values > μ

We fail to reject the Null hypothesis

b.)

If the consequences of a type I error are serious or expensive, then a very small significance level (value of α) is appropriate. So, if we consider the first method where we are getting low probability of type I error. It is suitable if we are not treating severe pain where the painkiller needs to work promptly.

Similarly, If the consequences of a Type I error are not very serious (and especially if a Type II error has serious consequences), then a larger significance level is appropriate. For the second case where we are maintaining α and β at 0.15, it is suitable where type I or type II error is not going to make a significance difference. For some less severe pain, where the painkiller doesn't have to be fast in action this method would be suitable.

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.

Answer:

A/B split testing is a form of hypothesis testing involving two variants. A series of randomized experiments are performed involving these two variants, with all factors and elements the same except for where variations of one particular factor or element. A/B testing allows you to test the impact your changes might have on your site, temporarily presenting changes to part of your audience while still serving your unchanged content to others.

A/B Testing comprises of a set of processes that one must follow sequentially in order to arrive at a **realistic conclusion**.

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

Collect Data: Your analytics will often provide insight into where you can begin optimizing. It helps to begin with high traffic areas of the site or app, as that will allow you to gather data faster. Look for pages with low conversion rates or high drop-off rates that can be improved.

Identify Goals: Conversion goals are the metrics that you are using to determine whether or not the variation is more successful than the original version. Goals can be anything from clicking a button or link to product purchases and e-mail signups.

Generate Hypothesis: Once you've identified a goal you can begin generating A/B testing ideas and hypotheses for why you think they will be better than the current version. Once you have a list of ideas, prioritize them in terms of expected impact and difficulty of implementation.

Create Variations: Using A/B testing software, make the desired changes to an element of the website or mobile app experience. This might be changing the color of a button, swapping the order of elements on the page, hiding navigation elements, or something entirely custom. Many leading A/B testing tools have a visual editor that will make these changes easy. Make sure to QA the experiment to make sure it works as expected.

Run Experiment: start the experiment and wait for visitors to participate! At this point, visitors to the site or app will be randomly assigned to either the control or variation of your experience. Their interaction with each experience is measured, counted, and compared to determine how each performs.

Analyze Results: Once the experiment is complete, it's time to analyze the results. A/B testing software will present the data from the experiment and show the difference between how the two versions of the page performed, and whether there is a statistically significant difference.

See if you can apply learnings from the experiment on other pages of the site and continue iterating on the experiment to improve the results. If the experiment generates a negative result or no results, use the experiment as a learning experience and generate new hypothesis that can be tested.