Question 1:

Quality assurance checks on the previous batches of medications found that it is four times more likely that a medicine is able to produce a satisfactory result than not.

Given a small sample of 10 medicines, you are required to find the theoretical probability that, at most, 3 medicines are unable to do a satisfactory job:

1. Propose the type of probability distribution that would accurately portray the above-mentioned scenario and list out the three conditions that this distribution follows.
2. Calculate the required probability.

Answer:

1. This distribution follows "**Binomial Theorem**".
   Conditions:
   - Total number of trials should be fixed or considered as "n".
   - Each trail should be binary, i.e. it has only two outcomes- success or failure.
   - Probability of success is denoted by p, should be the same in all the trails.

   Formula = **nCr(p)^r(1-p)^n-r**.

2. Let, probability of getting unsatisfactory result = p
   Then, probability of getting satisfactory result = 4*p (Acc. to the question)

   So, p+4p=1 , p=0.2

   Given, n=10, r<=3, p=0.2, 1-p=0.8

P(r<=3) =p(r=0) +p(r=1) +p(r=2) +p(r=3)

= {10C0(0.2)^0(0.8)^10-0}+ {10C1(0.2)^1(0.8)^10-1}+ {10C2(p)^2(0.8)^10-2}+{10C3(0.2)^3(0.8)^10-3}

= {0.107}+{0.268}+{0.301}+{0.201}=0.877

So, the required probability is **0.877 = 87.7%**

Question 2:

For the effectiveness test, a sample of 100 medicines 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 interval in which the population mean might lie – with a 95% confidence level:

1. 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.
2. Find the required interval.

Answer:

1. We will be using Central limit theorem to solve this problem.it states that for n > 30, the sampling distributions become normally distributed.

The following properties are:

- Sampling distribution's mean = Population mean
- Sampling distribution's standard deviation (standard error) = std. dev./sqrt(no. of samples)
- For n > 30, the sampling distribution becomes a normal distribution

2.

Population

$\sigma = 65$

~~Ho $\mu = 207$~~

~~Ha $\mu \neq 207$~~

(two-tailed test)

Sample

$\mu_x = 207$

$\sigma_x = ??$

$n = 100$

$LOS = 100 - 95$

$= 5\%$

$Z = 1.96$

$$\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{65}{\sqrt{100}} = 6.5$$

Upper critical Value $= \mu_x + (Z_c \times \sigma_x)$

Lower Critical Value $= \mu_x - (Z_c \times \sigma_x)$

$UCV = 207 + (1.96 \times 6.5) = 219.74$

$LCV = 207 - (1.96 \times 6.5) = 194.26$

So, the required interval is $(194.26, 219.74)$.

Question 3:

1. The painkiller 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) as that in the previous question, test the claim that the newer batch produces a satisfactory result and passes the quality assurance test. Utilise two hypothesis testing methods to take a 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.
2. You know that two types of errors can occur during hypothesis testing – 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 (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.

Under what conditions would either method be more preferred than the other? Give an example of a situation where conducting the hypothesis test with α and β as 0.05 and 0.45, respectively, would be preferred over conducting the same hypothesis test with α and β at 0.15 each. Similarly, give an example for the reverse scenario, where conducting the same hypothesis test with α and β at 0.15 each would be preferred over having them at 0.05 and 0.45, respectively.

For each example, give suitable reasons for your particular choice using the given values of α and β only. (Assume that no other information is available. Additionally, the hypothesis test that you are conducting is the same as mentioned in the previous question; you need to test whether the newer batch produces satisfactory results.)

Answer: 1.

JUNE

(a) Population $\mu = 200$

Sample $\mu_x = 207$, $\sigma_x = 6.5$, $n = 100$, LOS = 5%.

$H_0: \mu \leq 200$

$H_a: \mu > 200$

(Right tail test)

(Critical Value Method)

$Z_c = 1.64$

Upper critical Value:-

$= \mu + (Z_c \times \sigma_x)$

$= 200 + (1.64 \times 6.5)$

$= 210.66$

Lower critical value:-

$= \mu - (Z_c \times \sigma_x)$

$= 200 - (1.64 \times 6.5)$

$= 189.34$

SUNDAY 9 JUNE

So, As we can see the sample mean is lying between UCV & LCV. So, we failed to reject the null Hypothesis.

10 JUNE

P-value ~~for~~ method.

$\frac{\mu_x - \mu}{\sigma_x} = \frac{207 - 200}{6.5}$

$= 7/6.5 = 1.08$

$Z(1.08) = 0.86$

P-value $= 1 - 0.86$

$= 0.14$

$0.14 > 0.05$

P-value $> \alpha$, we will accept the null hypothesis

2. **Situation 1**: α = 0.05, β = 0.45

In cases where a Type I error (rejecting a true null hypothesis) is more critical than a Type II error (failing to reject a false null hypothesis), you would prefer the original method.

Example Reasoning: If producing a satisfactory result is crucial for patient safety, and the cost of producing a new batch is high, you might want to avoid the risk of rejecting a truly effective batch (Type I error). Here, the higher Type II error (β = 0.45) can be tolerated to minimize the risk of a faulty batch reaching the market.

Situation2: α = β = 0.15

In cases where Type I and Type II errors are equally critical, you might prefer the adjusted sampling procedure to balance the risks.

Example Reasoning: If the cost of producing a new batch is relatively low and you want to balance the risks of both types of errors, the adjusted sampling procedure with equal α and β values could be preferred. This ensures a more balanced approach to errors.

Question 4:

Once one 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 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 testing is entirely based on the two-sample proportion test & it helps us to choose between 2 variations of the same thing. Basically, we use A/B testing helps us to choose when the data is subjective and where we cannot use any statistical calculation or procedures.

Here are the steps to perform A/B testing.

1. Firstly, we divide the population in two groups i.e. one group for each variation.
2. Then we take their conversion rate into consideration.
3. Then we formulate null and alternate hypothesis.
    - Ho: The slogan 1 is more effective than slogan 2.
    - Ha: The slogan 2 is more effective than slogan 1.
4. Decide the level of significance.
5. Then we will perform statistical test on the conversion rate of both groups and calculate the p-vales.
6. Then we will compare it with LOS.
    - P-value> > LOS, we failed to reject the null hypothesis.
    - P-value> < LOS, we reject the null hypothesis.