①

ASSIGNMENT: Inferential and Hypothesis Testing.

Q-1

ⓐ The type of probability distribution that would portray the above scenario is BINOMIAL DISTRIBUTION.

Conditions for Binomial Probability Distribution:

① Total no. of trials is fixed at n.

② Each trial is binary i.e. only two possible outcomes success or failure.

③ Probability of success is same in all trials, denoted by p.

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

n = no. of trials p = prob. of success r = no. of successes after n trials.

Examples of Binomial Distribution (Applicability):

① Tossing a coin 20 times to see how many tails occur.

② Asking 200 randomly selected people if they are older than 21 or not.

③ Drawing 4 red balls from a bag putting each ball back after drawing it.

ⓑ calculating the required probability:

Let P(not producing satisfactory result) = x.

then as per statement, It is 4 times likely for drug to produce satisfactory job.

so, $4x + x = 1$.

$5x = 1$

$x = 1/5$

P(not producing satisfactory result) = 1/5

$(1-p)$ [satisfactory result) $= 1 - p = 1 - \frac{1}{5} = \frac{4}{5}$

Sample drugs (n) = 10

To find the probability that atmost 3 drugs are not able to do a satisfactory result, Using Binomial Distribution.

we can find, $P(X=x) = {}^{10}C_x \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{10-x}$

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

$P(X=0) = {}^{10}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10} = 1 \times 1 \times 0.107 = 0.107$

$P(X=1) = {}^{10}C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^9 = 10 \times 0.2 \times 0.134 = 0.268$

$P(X=2) = {}^{10}C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^8 = 45 \times 0.04 \times 0.168 = 0.302$

$P(X=3) = {}^{10}C_3 \left(\frac{1}{5}\right)^3 \left(\frac{4}{5}\right)^8 = 120 \times 0.008 \times 0.2097 = 0.201$

$P(X \le 3) = 0.107 + 0.268 + 0.302 + 0.201$
$= 0.878$

$\therefore$ The repuired probability is 0.878

Q 2

(a) Central limit theorem is the methodology that would be used to approach this problem.

Properties of CLT:

1. Sampling distribution's mean$(\mu_{\bar{x}})$ = Population mean (μ)
2. ~~sampling~~ Sampling distribution's standard deviation (standard error) = $\frac{\sigma}{\sqrt{n}}$

σ = population standard deviation.
n = sample size.

3. for $n > 30$, the sampling distribution becomes Normally distributed.

(b) sample drugs(n) = 100
sample mean, $\mu_{\bar{x}} = 207$ sec.
sample standard deviation $(\sigma_{\bar{x}})$ = 65sec , assuming population Std dev = Sample Std dev = 65sec.
Confidence level : 95%

Confidence Interval $= \cancel{207}\ \mu_{\bar{x}} \pm \left(z * \frac{\sigma}{\sqrt{n}}\right)$

$= 207 \pm \left(1.96 * \frac{65}{\sqrt{100}}\right)$

$= 207 \pm (1.96 * 6.5)$

$= 207 \pm (12.74)$

$= (219.74, 194.26)$

As sample mean lie in Range so one can say that It's effective.

5

Q-3A

In this question, we are given that:

Sample data (n) = 100

$H_0 = \mu \leq 200$ [Null Hypothesis]

$H_1 = \mu > 200$ [Alternate Hypothesis]

Sample mean given as 207 sec.

Sample standard deviation = 65 sec

Significance level = 5%.

① Critical value method.

as per null and alternate hypothesis:

Area of test is one tailed test i.e. Upper tailed test.

(rejection region on right side of distribution)

$\sigma_{\bar{x}} = \sigma = 65$ sec [Assuming population stand.dev = Sample Std dev]

$Z_c = 1 - 0.05 = 0.950$

$UCV = \mu + (Z_c \times \sigma_{\bar{x}})$

$= 200 + \left(1.645 \times \frac{65}{\sqrt{100}}\right)$

$\left(\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$

(Value of Z_c at 0.950 is 1.645 (Z-Table))

$= 200 + 10.6925$

$= 210.6925$

So, 207 (sample mean) falls in acceptance region so Null Hypothesis cannot be rejected. [Fail to reject Null hypothesis]

② P- Value Test:

as per null and alternate hypothesis:

Area of test is one tailed test i.e. Upper tailed test

$\mu_{\bar{x}} = \mu = 200$ $n = 100$

$\bar{x} = 207$

$\sigma_{\bar{x}} = \sigma = 65$ (Assuming pop std.dev same as sample std dev)

$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{207 - 200}{\frac{65}{\sqrt{100}}} = \frac{7}{6.5} = 1.0769$

Prob. Z value of 1.0769 = 0.8599

p value = 1 − 0.8599 = 0.14 so 0.14 > 0.05

As p value is greater than significance level, so fail to reject null Hypothesis

Q-3B

Null Hypothesis, H_0: Drug produces satisfactory result.
H_1: Drug doesnot produce satisfactory result.

There are two types of error that can happen in hypothesis testing:

Type 1 error is when we reject a true null hypothesis and is denoted by α

Type 2 error is when we fail to reject a false null hypothesis and is denoted by β.

In given case, Type 1 error would be:
- Drug doesnot produce satisfactory result even when It does.

Type 2 error would be:
- Drug produces satisfactory result even when It doesnot.

So we have 2 cases here:

	Case 1	Case 2
α (Type 1)	0.05	0.15
β (Type 2)	0.45	0.15

① In Case 1, we have α as 0.05 and β as 0.45. Lower value of α makes It harder to reject null hypothesis., So If null hypothesis is false It may be difficult to reject with low value of α. Low value of α means high probability of Type 2 error, β. As both α and β are inversely proportional to eachother. Meanwhile In case 2 probability of α and β are same which is a rare case, so In this any type of error can happen.

② Case 1 has more consequencies compared to Case 2, as due to high probability of β, Null hypothesis is not getting rejected even when Drugs are not producing satisfactory result. So to have stability, prob of type 1 should be increased. In Case 2, we don't have much consequences just to be safer side we can increase α.

③. In terms of consumer both α and β can be hazardous. like high dosage of painkiller means α should be reduce and If painkiller doesnot have any side effects than we can have β more probability to α error as compared to β. With type 2 error being more, company is selling drugs which would not give satisfactory results so that might have bad reviews so proper test should be done before manufacture of large quantity else all will posses same error.

④ In this case, Type 2 error would be more dangerous so It should be reduced by increasing the type 1 error.

⑤ Company can use Case 2 for conducting test. as It doesnot have much consequences.

Q-4

A/B testing is done for testing two different versions at the same time.

It's called A/B testing because It have A version and B version.

Following are the steps which one should follow while doing A/B testing:

1. Pick a variable to test
2. Identify a goal
3. Create a control and a challenger
4. Split your sample groups equally and randomly.
5. Determine your sample size.
6. Decide how significant your results need to be.
7. Make sure you're only running one test at a time or any campaign.

Applying the steps in an example:

- In a mobile fighting game, It was seen that Only 15% users were able to pass the 4rt level.
- So A/B testing is done to check different versions of 4rt level.
- Two test cases were decided to check how many users will pass the 4rt level.
- Testcase 1: Original as It is.
  Test case 2: 10% fewer enemies : 10% easier.
- Next step is to distribute this version to 2 group of players so 50% players will play the original version and 50% will play new version.
- Then we get our result. that:
  15% completed original level 4.
  70% completed new version level 4 which 10% easier.

So with the help of A/B testing, It was decided which version is best for game i.e new version with 10% less enemies.