Statistics Assignment → Inferential and Hypothesis Testing

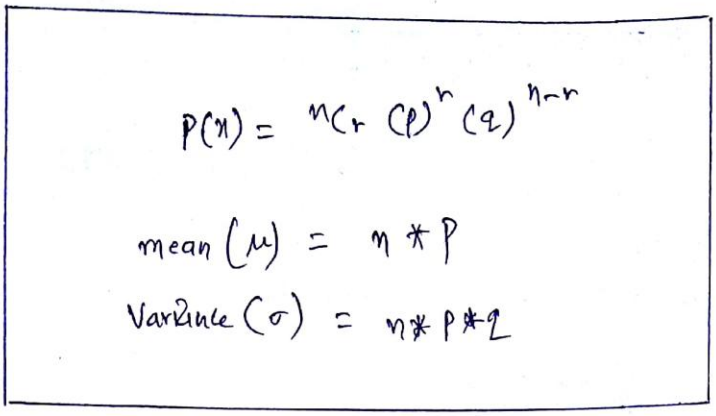
Solution: 1.

(a)

The type of probability distribution that would accurately portray the mentioned scenario could be proposed as the -- "**Binomial Distribution**".

And please find below list of the three conditions that this distribution follows:

1. The number of observations 'n' is fixed and each such trail/observation is independent.
2. Each trail is binary i.e. it represents either one of two outcomes -- success ('p') or failure ('q')
3. The probability for success is same for each trail and is denoted as 'P'. It can be calculated as:



(b)

Number of sample of drugs (n) = 10

P(x<=3)=?

$$P(x \leq 3) = ?$$

$$P(x \leq 3) = {}^{n}C_r (p)^r (q)^{n-r}$$

$$= {}^{10}C_0 (4/5)^{10} + {}^{10}C_1 (4/5)^9 (1/5)^1 + 10 (4/5)^8 (1/5)^2 + 10 (4/5)^7 (1/5)^3$$

$$= 1 (0.8)^{10} (0.2)^0 + 10 (0.8)^9 (0.2)^1 + 45 (0.8)^8 (0.2)^2 + 120 (0.8)^7 (0.1)^3$$

$$= 0.11 * 1 + 10 (0.13)(0.2) + 45 (0.17)(0.04) + 120 (0.21)(0.008)$$

$$= 0.11 + 0.26 + 0.31 + 0.20$$

$$\boxed{P(x \leq 3) = 0.88}$$

Solution: 2.

(a)

The methodology which would be used for this problem would be: **Central Limit theorem (CLT).**

Required properties are listed below:

- Mean of the population is approximately equal to the mean of the sampling distribution.

$$\mu_{\bar{x}} = \mu$$

- The standard deviation of sampling distribution also known as the standard error is equal to the population standard deviation divided by the sq. root of the sample size (n).

$$\text{Standard Error (SE)} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\text{sample mean}} = \frac{\sigma_{\text{Population}}}{\sqrt{n}}$$

- Therefore, if n > 30, then it will be considered as normal distribution. Greater the size, lower the standard deviations, inverse proportion as shown below.

$$\sigma_{\text{mean sample}} = \frac{\sigma_{\text{Population}}}{\sqrt{n}}$$

(b)

b) $\bar{x} = 207$

$n = 100$

$\sigma = 105$

Confidence Interval $= 95\% = 0.95$

$\therefore z^* = 1.96$ [from z-table]

$$\text{Required Range} = \left(\bar{x} + \frac{z^* \sigma}{\sqrt{n}} , \bar{x} - \frac{z^* \sigma}{\sqrt{n}} \right)$$

$$= \left(207 + \frac{(1.96 * 105)}{\sqrt{100}} , 207 - \frac{(1.96 * 105)}{\sqrt{100}} \right)$$

$$= [(207 + 12.74) , (194.26)]$$

$\therefore$ Range $= (219.74 , 194.26)$

The range is $(219.74, 194.26)$

Solution: 3.

(a)

3) a)

$n = 100$

$\mu_0 = 200$ ∵ Painkiller drugs needs effect of atmost 200sec.

$\sigma = 65$

$\bar{x} = 207$

$\alpha = 5\% = 0.05$

so, $(H_0 : \mu \leq 200)$ → could be the null hypothesis

and, $(H_1 : \mu > 200)$ → Alternate hypothesis

Now, using critical value theorem (CLT) :

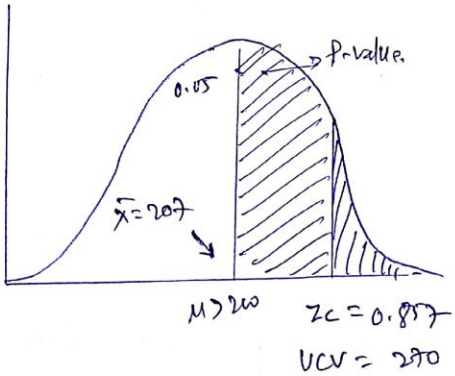
$$Z_c = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$= \frac{207 - 200}{6.5}$$

$$= 1.07 \quad \text{(Now look for 1.07 value in z table)}$$

$$\left[\mu_{\bar{x}} = \mu \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{65}{10} = 6.5 \right]$$

$\therefore Z_c = 0.857$



$UCV = 200 + (1.07 \times 6.5)$

$= 270$

Sampling distribution of $\bar{x}$

∴ Using p-value method :

Since its a 'one tailed test' → $P = 1 - Z_c = 1 - 0.857$

$\boxed{P = 0.143}$

If $P < 0.05$ i.e. significance value

then we can reject the null hypothesis

But, $0.143 > 0.05$, so it is clear that we

⇒ failed to reject the null hypothesis

(b)

b) Condition I :

$n = 100$

$\bar{X} = 207$

$\sigma = 65$

$\alpha = 0.05$ } - diff. values

$\beta = 0.45$ }

Condition II :

$n = 100$

$\bar{X} = 207$

$\sigma = 65$

$\alpha = 0.15$ } → Controlled at same value

$\beta = 0.15$ }

So for the current hypothesis please find below the type I and type II errors as mentioned in the problem statement:

(i) Type I error (α): For this, here the painkiller drugs have a time of effect <= 200 seconds but rejected.

(ii) Type II error (β): Here the painkiller drug won't have the same time of effect i.e. the time of effect is > 200 seconds; but is accepted.

Condition 1: "The pain is really mild to normal and time of effect is given to be 200 seconds. "

Therefore, here in this condition the probability of 'β' would be preferred with more and greater time effect which can be given to the patients with no such need of painkillers.

Conditions 2: "The pain is critical and time of effect cannot be greater more than 200 seconds."

Hence, in this condition the probability of 'α' would be preferred so that the painkiller drugs having effect > 200 seconds should not be given to any patients having critical conditions.

Solution: 4.

A/B testing is an effective process and measure in which we compare the two versions of assets/objects of same marketing, which can be done for example as the web page exposed to similar visitors or here in this case the online ad campaign. On the basis of better conversion rate we can decide and conclude which one to choose finally.

A/B testing makes sure that the companies or individuals make changes to the offering according to the user experiences after dealing with the data, as a result. It could use that modification to optimize the desired outcome that in turn ensures marketing campaign to be more effective and profitable.

Stepwise procedure (A/B testing) that can be followed:

1. Get the marketing taglines (both of them) ready to launch, finalized.
2. Deploy/launch the taglines to some other tools or websites (like Google Ads, Google Analytics, and other relevant platforms).
3. Observe the visitor behavior of the customers towards both taglines.
4. Formulate the Hypothesis.
5. Test the formulated hypothesis (Creating variations…)
6. Analyze the test data and accordingly enhance the conclusion to get the most efficient one.

Above mentioned are some brief steps that can be followed for A/B testing, please find below some of the steps explained briefly:

- **Collect the data:** As mentioned above the very first step is to get the required data for analysis. It helps us to begin with important high areas of traffic of different applications and web sites as that will allow gathering the required information more quickly and faster for analysis.

- **Goals Identify:** The most important metric here is the 'conversion goals' that could determine whether a particular variation ('B') is more successful or not than the original one ('A'). This is the metric that will decide which to opt for.

- **Generating Hypothesis:** After the goals are identified, then we can begin with generating the required A/B testing hypothesis, stating, how and why they can be better than the original version (Like A is better that B, so on and so forth).

- **Create Variation:** Like it was mentioned in the course we can make use of tools like Optimizely, and can make changes the website or application experience for users. And here we can make changes to the tag lines which can be shown to different users exposing them to the taglines to get the numbers to compare.

- **Running Experiment:** So after plotting all these, at this time we just wait for the users to participate to our websites/applications/taglines which would be randomly assigned to either control or variations of experience.

  The interaction of the users with the control and variation is taken into considerations and after measuring the effect of both of them; we can go for the better performing tagline after calculation.

- **Analyzing the Result:** So after the experiment is completed, now the time has come to analyze results where we can check which version(A/B) performed the better out of the two taglines which can be seen comparing the numbers coming out from the users conversions while performing above mentioned steps.

- **Keep testing:** Just because we were successful in finding the improved conversion rate, it doesn't mean we found the best landing tagline. More tests have the potential to produce bigger gains.

This is how the A/B testing is performed.