Stats

Assumptions

- Have to answer in context of question
- Binomial Distribution
  - Trials are independent
  - Probability of sucess remains constant (not affected by results of other trial)
- Random Samples
  - Each element in population has equal chance of being selected
  - Selections are independent of each other (selection of 1 element does not affect the chance of any other element being selected)

Discrete Random Variables

- Variance of discrete random variable X is a measure of expected spread of a random variable about its mean. It is defined as $Var(X) = E(X-\mu)^2 = E(X^2) - [E(X)]^2$
- Standard deviation defined as $\sigma = \sqrt{Var(x)}$
- $Var(aX \pm bY) = a^2Var(X) + b^2Var(Y)$
- $Var(X+b) = Var(X)$ since horizontal shift does not affect variance

Normal Distribution

- Standardisation: Covert to $Z \sim N(0,1)$ where $Z \ \frac{X-\mu}{\sigma}$
- $aX \pm bY \ N(\alpha\mu_1 \pm b\mu_2, \alpha^2\mu_1^2 + b^2\sigma_2^2)$

Sampling

- Sample mean $\overline{x} = \frac{\sum x}{n}$ is an unbiased estimate of μ
- Unbiased estimate of σ^2 is $s^2 = \frac{n}{n-1}[\frac{\sum(x-\overline{x})^2}{n}] = \frac{1}{n-1}[\sum x^2 - \frac{(\sum x)^2}{n}]$. Is equal to product of $\frac{n}{n-1}$ and sample variance.
- Sample mean $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$, Sample sum $\sum X = n\overline{X} \sim N(n\mu, n\sigma^2)$
- Central Limit Theorem: Since $n = 50 > 30$ is large, by Central Limit Theorem, $n \sim N()$ approximately

Hypothesis Testing

- Level of Signficance of a hypothesis test α is defined by probability of rejecting H_0 when (given that) H_0 is true
- Test statistic $\overline{X}$ is random variable used to make decision whether to reject H_0 by measuring degree of agreement between sample and null hypothesis.
- Critical Region refers to range of sample mean (variable), (context), such that there is sufficient evidence at $\alpha\%$ signficiance level to conclude that the mean (variable) is not μ (variable).
- p-value is the probability of observing a sample, a test statistic value "at least as extreme as" the observed value computed under the assumption the null hypothesis is true. It is the smallest level of significance at which H_0 can be rejected.
- z-test: Since p-value = $p \leq\geq \alpha$, we reject/accept H_0 and conclude there is sufficient/insufficient evidence at $\alpha\%$ level of significance, (that the mean is not μ_0 / to support the claim that mean is not μ_0)