STAT 245 FORMULA SHEET

Theophilus Nyasha Nenhanga

June 2023

1 Introduction

1.1 Mean

$$\text{Mean} = \frac{\text{Sum of observations}}{\text{Number of Observations}}$$

$$\mu = \text{Population Mean}$$
$$\overline{x} = \text{Sample Mean}$$

$$\text{Population Mean, } \mu = \frac{\sum x_i}{N}$$
$$\text{Sample Mean, } \overline{x} = \frac{\sum x_i}{n}$$

1.2 Median

$$\text{Median, } M$$
$$\text{Position of Median} = \frac{n+1}{2}$$

1.3 Measures of Spread

1.3.1 Range

$$\text{Range, } R$$
$$R = \text{largest value} - \text{Smallest value}$$

1.3.2 Mean Deviation

$$\text{Mean Deviation, } MD$$
$$\text{For Populations, } \frac{1}{N}\sum_{i=1}^{N}|x_i - \mu|$$

$$\text{For Samples, } \frac{1}{n}\sum_{i=1}^{N}|x_i - \overline{x}|$$

$$\text{N} = \text{Population size}$$
$$\mu = \text{Population mean}$$
$$n = \text{Sample size}$$
$$\overline{x} = \text{sample mean}$$

1.3.3 Variance

$$\text{For Population Variance}\sigma^2$$
$$\sigma^2 = \frac{1}{N}\left(\sum x_i^2 - \frac{1}{N}\left(\sum x_i\right)^2\right)$$

$$\text{For Sample Variance}s^2$$
$$s^2 = \frac{1}{N}\left(\sum x_i^2 - \frac{1}{n-1}\left(\sum x_i\right)^2\right)$$

1.3.4 Standard Deviation

$$\text{For Population Variance}\sigma$$
$$\sigma = \sqrt{\sigma^2}$$
$$\text{For Sample Variance}s$$
$$s = \sqrt{s^2}$$

1.4 Dealing with large Samples

$$\text{Population Size, } N = \sum f_i$$

$$\text{Mean, } \mu = \frac{\sum x}{N} = \frac{\sum(x_i f_i)}{\sum f_i} = \frac{\sum(x_i f_i)}{N}$$

$$\text{Variance, } \sigma^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N} = \frac{\sum({x_i}^2 f_i) - \frac{\sum(x_i f_i)^2}{\sum f_i}}{\sum f_i} = \frac{\sum({x_i}^2 f_i) - \frac{\sum(x_i f_i)^2}{N}}{N}$$

2 Bell Shaped Distributions

2.1 The z-score

The z score is a multiple of how many standard deviatiosn we are away from the mean.
$z = \dfrac{x-\mu}{\sigma}$

3 Margin of Error

$$ME = z_{\frac{\alpha}{2}}\sqrt{\frac{p \cdot q}{n}}$$
$$p = \text{proportion}$$
$$n = \text{sample size}$$
$$z = \text{critical value}$$
$$\alpha = \text{confidence interval}$$

4 Confidence Intervals

4.1 When Mean Unknown, Standard Deviation and Sample Size Given

$$2 \cdot z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

4.2 Sample Mean and Sample Standard Deviation Given

$$\overline{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\begin{aligned} \overline{x} &= \text{sample mean} \\ n &= \text{sample size} \\ s &= \text{sample standard deviation} \\ \alpha &= \text{confidence interval} \end{aligned}$$

4.3 Binomial Populations

$$\hat{p} - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}} \leq \mu \leq \hat{p} + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\begin{aligned} \hat{p} &= \text{proportion} \\ \hat{q} &= 1 - \hat{p} \\ n &= \text{sample size} \\ \alpha &= \text{confidence interval} \end{aligned}$$

5 Two Sample Problems

5.1 Difference Between Population Means

$$\overline{x}_1 - \overline{x}_2 - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \overline{x}_1 - \overline{x}_2 + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\begin{aligned} \overline{x}_1 &= \text{sample mean}_1 \\ \overline{x}_2 &= \text{sample mean}_2 \\ \sigma_1^2 &= \text{variance}_1 \\ \sigma_2^2 &= \text{variance}_2 \\ \alpha &= \text{confidence interval} \\ n_1 &= \text{sample size}_1 \\ n_2 &= \text{sample size}_2 \end{aligned}$$

5.2 Difference Between Population Proportions

$$\hat{p}_1 - \hat{p}_2 - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1 \cdot \hat{q}_1}{n_1} + \frac{\hat{p}_2 \cdot \hat{q}_2}{n_2}} \leq \hat{p}_1 - \hat{p}_2 \leq \hat{p}_1 - \hat{p}_2 + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1 \cdot \hat{q}_1}{n_1} + \frac{\hat{p}_2 \cdot \hat{q}_2}{n_2}}$$

$$\begin{aligned} \hat{p}_2 &= \text{sample proportion}_1 \\ \hat{p}_1 &= \text{sample proportion}_2 \\ \hat{q}_1 &= 1 - \hat{p}_1 \\ \hat{q}_2 &= 1 - \hat{p}_2 \\ \alpha &= \text{confidence interval} \\ n_1 &= \text{sample size}_1 \\ n_2 &= \text{sample size}_2 \end{aligned}$$

6 Estimating The Population Mean

6.1 Estimating The Population Mean

$$n \geq \Big(\Big)$$

7 Pooled Variables

7.1 Pooled Variance

$$\frac{(n_1-1)\cdot s_1^2+(n_2-1)\cdot s_2^2}{n_1+n_2-2}$$

8 Hypothesis Testing

8.1 When to reject

8.1.1 Critical Value Approach

1. If it is a two tailed test, Critical Values are $-z_{\frac{\alpha}{2}}$ and $z_{\frac{\alpha}{2}}$
   - Reject H_0 if TS $\leq -z_{\frac{\alpha}{2}}$ or TS $\geq z_{\frac{\alpha}{2}}$
   - Do not reject H_0 if $-z_{\frac{\alpha}{2}} <$ TS $< z_{\frac{\alpha}{2}}$

2. If it is a left tailed test, Critical Value is $-z_\alpha$
   - Reject H_0 if TS $\leq -z_\alpha$
   - Do not reject H_0 if TS $> -z_{\frac{\alpha}{2}}$

3. If it is a right tailed test, Critical Value is z_α
   - Reject H_0 if TS $\geq -z_\alpha$
   - Do not reject H_0 if TS $< z_{\frac{\alpha}{2}}$

8.1.2 P Value Approach

- if it is a two-tailed test, then

$$PValue = P(z \leq -|TS|) + P(z \geq |TS|) = 2 \cdot P(z \leq -|TS|)$$

- If it is a left-tailed test, then

$$PValue = P(z \leq TS)$$

- If it is a left-tailed test, then

$$PValue = P(z \leq TS)$$

And For all three types of test,

- Reject H_0 if $PValue \leq \alpha$
- Do not reject H_0 if $PValue > \alpha$

9 Calculating Test Statistics

9.1 Testing One Mean

$$\text{TS} = \frac{x-\mu_0}{\frac{s}{\sqrt{n}}}$$

$$\text{TS} = \text{Test Statistic}$$

9.2 Testing One Proportion

$$\text{TS} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}}$$