MATH 4740/MSSC 5740
Chapter 7 Problem Solving # 14, 22

## 7.10 Summary

| **Number of Groups, Outcome: Parameter** | Test Statistic, $n<30$ | Test Statistic, $n\geq 30$ |
|---|---|---|
| Two independent samples, continuous: $\mu_1 = \mu_2$ | $t = \dfrac{\bar{X}_1 - \bar{X}_2}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$, $df=n_1+n_2-2$ <br> (Technically assumes populations are normal, independent within and between populations, and population variances equal.) | $z = \dfrac{\bar{X}_1 - \bar{X}_2}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$, <br> $S_P = \sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ |
| Two matched samples, continuous: $\mu_d$ | $t = \dfrac{\bar{X}_d}{s_d/\sqrt{n}}$, $df=n-1$ <br> (Technically assumes data is normal, pairs independent of each other, variances equal between populations.) | $z = \dfrac{\bar{X}_d}{s_d/\sqrt{n}}$, <br> $\bar{X}_d = \dfrac{1}{n}\sum_{i=1}^{n} d_i$, <br> $s_d = \sqrt{\dfrac{\Sigma d^2 - (\Sigma d)^2/n}{n-1}}$ |
| Two independent samples, dichotomous: $p_1=p_2$ | Multinomial Test <br> (Not taught in this class.) | $z = \dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$, <br> $\hat{p}_1 = \dfrac{X_1}{n_1}$, $\hat{p}_2 = \dfrac{X_2}{n_2}$, $\hat{p} = \dfrac{X_1+X_2}{n_1+n_2}$ |
| More than two samples, continuous: $\mu_1 = \ldots = \mu_k$ | Same as large sample. <br> (Technically assumes populations are normal, independent within and between populations, and population variances equal.) | $F = \dfrac{MSB}{MSE}$, $df_1=k-1$, $df_2=N-k$ <br> $MSB = \dfrac{1}{k-1}\sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X})^2$ <br> $MSE = \dfrac{1}{N-k}\sum_{j=1}^{k}\sum_{i=1}^{n_j} n_j(X_i - \bar{X}_j)^2$ |
| Two or more samples, Categorical and Ordinal: $p_{11},\ldots,p_{rc}$ | Multinomial Test <br> (Not taught in this class.) | $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\dfrac{(O_{ij}-E_{ij})^2}{E_{ij}}$, <br> $df=(r-1)(c-1)$ |

MATH 4740/MSSC 5740
Chapter 7 Problem Solving # 14, 22

MATH 4740/MSSC 5740

Chapter 7 Problem Solving # 14, 22

## 7.11 Practice Problems

14. Suppose a hypertension trial is mounted and 18 participants are randomly assigned to one of the comparison treatments. Each participant takes the assigned medication and his or her systolic blood pressure is recorded after 6 months on the assigned treatment. The data are shown in Table 7.58.

| Standard Treatment (1) | Placebo (2) | New Treatment (3) |
|---|---|---|
| 124 | 134 | 114 |
| 111 | 143 | 117 |
| 133 | 148 | 121 |
| 125 | 142 | 124 |
| 128 | 150 | 122 |
| 115 | 160 | 128 |

Is there a difference in mean systolic blood pressure among treatments? Run the appropriate test at $\alpha = 0.05$.

Answer:

Step 1. Set up hypotheses and determine level of significance.

$H_0: \mu_1 = \mu_2 = \mu_3$

$H_1$: Means are not all equal. $\alpha = 0.05$

Step 2. Select the appropriate test statistic.

$$F = \frac{MSB}{MSE}, \quad df_1 = k-1,\ df_2 = N-k$$

$$MSB = \frac{1}{k-1}\sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X})^2$$

$$MSE = \frac{1}{N-k}\sum_{j=1}^{k}\sum_{i=1}^{n_j} n_j(X_i - \bar{X}_j)^2$$

Step 3. Set up decision rule.

$df_1 = k - 1 = 3 - 1 = 2$ and $df_2 = N - k = 18 - 3 = 15$.

$df_1$ — $\alpha = 0.05$

| $df_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 236.8 | 238.9 | 240.5 | 241.9 | 248.0 | 250.1 | 251.1 | 251.8 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 | 19.45 | 19.46 | 19.47 | 19.48 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 | 8.66 | 8.62 | 8.59 | 8.58 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.80 | 5.75 | 5.72 | 5.70 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.56 | 4.50 | 4.46 | 4.44 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | 4.06 | 3.87 | 3.81 | 3.77 | 3.75 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | 3.64 | 3.44 | 3.38 | 3.34 | 3.32 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | 3.35 | 3.15 | 3.08 | 3.04 | 3.02 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | 3.14 | 2.94 | 2.86 | 2.83 | 2.80 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.77 | 2.70 | 2.66 | 2.64 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | 2.85 | 2.65 | 2.57 | 2.53 | 2.51 |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | 2.75 | 2.54 | 2.47 | 2.43 | 2.40 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 | 2.67 | 2.46 | 2.38 | 2.34 | 2.31 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | 2.60 | 2.39 | 2.31 | 2.27 | 2.24 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 | 2.33 | 2.25 | 2.20 | 2.18 |

Reject $H_0$ if $F \geq 3.68$.

Step 4. Compute the test statistic.

| Standard | Placebo | New Treatment |
|---|---|---|
| $n_1 = 6$ | $n_2 = 6$ | $n_3 = 6$ |
| $\bar{X}_1 = 122.7$ | $\bar{X}_2 = 146.2$ | $\bar{X}_3 = 121.0$ |

If we pool all $N = 18$ observations, the overall (or grand) mean is

$$\bar{X} = (124+111+133+125+128+115+134+143+148+142+150+160+114+117+121+124+122+128)/18 = 130.0$$

We can now compute $SSB = \sum_{j=1}^{3} n_j\ (\bar{X}_j - \bar{X})^2$

$$SSB = 6(122.7 - 130.0)^2 + 6(146.2 - 130.0)^2 + 6(121.0 - 130.0)^2 = 2380.4.$$

Next, $SSE = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_i - \bar{X}_j)^2$

| Standard | $(X - 122.7)$ | $(X - 122.7)^2$ |
|---|---|---|
| 124 | 1.30 | 1.7 |
| 111 | –11.7 | 136.9 |
| 133 | 10.3 | 106.1 |
| 125 | 2.3 | 5.3 |
| 128 | 5.3 | 28.1 |
| 115 | -7.7 | 59.3 |
| $\bar{X}_1 = 122.7$ | | 337.3 |

Thus, $\sum_{i=1}^{6}(X_i - \bar{X}_1)^2 = 337.3$.

| Placebo | $(X - 146.2)$ | $(X - 146.2)^2$ |
|---|---|---|
| 134 | –12.2 | 148.8 |
| 143 | –3.2 | 10.2 |
| 148 | 1.8 | 3.2 |
| 142 | –4.2 | 17.6 |
| 150 | 3.8 | 14.4 |
| 160 | 13.8 | 190.4 |
| $\bar{X}_2 = 146.2$ | | 384.8 |

Thus, $\sum_{i=1}^{6}(X_i - \bar{X}_2)^2 = 384.8$.

| New treatment | $(X - 121.0)$ | $(X - 121.0)^2$ |
|---|---|---|
| 114 | –7 | 49.0 |
| 117 | –4 | 16.0 |
| 121 | 0 | 0.0 |
| 124 | 3 | 9.0 |
| 122 | 1 | 1.0 |
| 128 | 7 | 49.0 |
| $\bar{X}_3 = 121.0$ | | 124.0 |

Thus, $\sum_{i=1}^{6}(X_i - \bar{X}_3)^2 = 124.0$

$$SSE = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_i - \bar{X}_j)^2 = 337.3+384.84+124.0=846.1.$$

We can now construct the ANOVA table.

| Source of Variation | Sums of Squares (*SS*) | Degrees of Freedom (*df*) | Mean Squares (*MS*) | *F* |
|---|---|---|---|---|
| Between Treatments | 2380.4 | 2 | 1190.2 | 21.1 |
| Error or Residual | 846.2 | 15 | 56.4 | |
| Total | 3226.6 | 17 | | |

Step 5. Conclusion.

We reject $H_0$ because $21.1 \geq 3.68$. We have statistically significant evidence at $\alpha = 0.05$ to show that there is a difference in mean systolic blood pressure among treatments.

MATH 4740/MSSC 5740

Chapter 7 Problem Solving # 14, 22

22. Use the data shown in Problem 21 and test if there is an association between mother's BMI and the child's obesity status (i.e., normal versus overweight/obese)? Run the test at a 5% level of significance.

| Mother: | Child: Normal | Child: Overweight/Obese | Total |
|---|---|---|---|
| Normal | 40 | 16 | 56 |
| Overweight | 15 | 14 | 29 |
| Obese | 7 | 8 | 15 |
| Total | 62 | 38 | 100 |

Answer:

Step 1. Set up hypotheses and determine level of significance.

$H_0$: Mother's BMI and child's obesity status are independent.

$H_1$: $H_0$ is false. $\alpha = 0.05$

Step 2. Select the appropriate test statistic.

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}},\ df=(r-1)(c-1)$$

The condition for appropriate use of this test statistic is that each expected frequency is at least 5. In Step 4 we will compute the expected frequencies and we will ensure that the condition is met.

Step 3. Set up decision rule.

$df = (3-1)(2-1) = 2$ and the decision rule is: Reject $H_0$ if $\chi^2 \geq 5.99$. (Table below.)

| df | .10 | .05 | .025 | .01 | .005 |
|---|---|---|---|---|---|
| 1 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 |
| 2 | 4.61 | 5.99 | 7.38 | 9.21 | 10.60 |

Step 4. Compute the test statistic.

We now compute the expected frequencies using the formula:

*Expected Frequency* = (*Row Total* * *Column Total*)/*N*

The top number in each cell of the table is the observed frequency and the bottom number is the expected frequency (shown in parentheses).

| Mother: | Child: Normal | Child: Overweight/Obese | Total |
|---|---|---|---|
| Normal | 40 (34.7) | 16 (21.3) | 56 |
| Overweight | 15 (18.0) | 14 (11.0) | 29 |
| Obese | 7 (9.3) | 8 (5.7) | 15 |
| Total | 62 | 38 | 100 |

The test statistic is computed as follows:

$$\chi^2 = \frac{(40-34.7)^2}{34.7}+\frac{(16-21.3)^2}{21.3}+\frac{(15-18.0)^2}{18.0}+\frac{(14-11.0)^2}{11.0}+\frac{(7-9.3)^2}{9.3}+\frac{(8-5.7)^2}{5.7}$$

$\chi^2 = 0.81 + 1.32 + 0.5 + 0.82 + 0.57 + 0.93 = 4.95.$

Step 5. Conclusion.

Do not reject $H_0$ because $4.95 < 5.99$. We do not have statistically significant evidence at $\alpha = 0.05$ to show that $H_0$ is false or that mother's BMI and child's obesity status are not independent.