Tutorial letter 202/2/2018

Statistical Inference II

STA2602

Semester 2

Department of Statistics

Solutions to Assignment 2

QUESTION 1 **[10]**

(a) yield of maize per plot. **(2)**

(b) rate of application of fertiliser and watering frequency. **(2)**

(c) one-acre plots. **(2)**

(d) province or research station because the different environmental conditions in the province or at the stations may also affect yield. **Justify your answer.** **(4)**

QUESTION 2 **[35]**

(a) **(10)**

Identify:

(i) mileage per liter; **(2)**

(ii) brand of petrol; **(2)**

(iii) cars; and **(2)**

(iv) type of car because the differences among types of cars may also affect mileage per litre. **(4)**

(b) **(25)**

(i) **(8)**

ANOVA Table I

Source	df	SS	MS	F
Brand	**2**	13.80	**6.900**	**1.0970**
Error	**6**	**35.55**	**5.925**	
Total	**8**	49.34		

ANOVA Table II

Source	df	SS	MS	F
Brand	**2**	13.80	**6.9000**	**3.5375**
Car	**2**	28.20	**14.1000**	**7.6735**
Error	**4**	**7.35**	**1.8375**	
Total	**8**	49.34		

(ii) In $ANOVA$ Table I the estimate (5.925) contains the between types of car variation which is removed in $ANOVA$ Table II. **(3)**

(iii) The **appropriate** table is $ANOVA$ Table II.
The test statistic is $F = 3.5374$.
The decision rule is "Reject H_0 if $F > F_{0.05} = 6.9443$ (**obtained from the** $F - tables$ **using 2 and 4 numerator and denominator degrees of freedom, respectively**)". (3)
Since $F = 3.5375 < F_{0.05} = 6.9443$ we fail to reject H_0 and conclude, at the 0.05 level of significance, that there are no significant differences among the brand means. (2)

(iv) Let $\hat{\mu}_{III} - \hat{\mu}_I = 2.9$ be the difference between the III and I brand sample means. (1)

The estimate of the standard error of $\hat{\mu}_{III} - \hat{\mu}_I$ is given by

$$s_{\hat{\mu}_{III}-\hat{\mu}_I} = \sqrt{MSE\left(\frac{2}{3}\right)} = \sqrt{\frac{1.8375}{2}} = 1.1068.$$

(3)

The 95% confidence interval for the difference between the III and I brand means is given by

$$\hat{\mu}_{III} - \hat{\mu}_I \pm t_{0.025} \times s_{\hat{\mu}_{III}-\hat{\mu}_I} = 2.9 \pm 2.7764 \times 1.1068 = (-0.1730; 5.9730),$$

where $t_{0.025} = 2.7764$ is obtained from the $t - tables$ using 4 degrees of freedom. (5)

QUESTION 3 **[15]**

(a) **(4)**

The $paired - t$ test in order to remove the effect of pair variation in the starting annual salaries of male and female graduates.

(b) **(11)**

Let $\hat{\mu}_D = 400$ be the difference between the mean starting annual salaries of male and female graduates and $s_D^2 = 188888.89$ be the variance of the pairwise differences between the starting annual salaries of the male and female graduates.

The estimate of the standard error of $\hat{\mu}_D$ is given by

$$s_{\hat{\mu}_D} = \sqrt{\frac{s_D^2}{10}} = \sqrt{\frac{188888.89}{10}} = 137.44.$$

(3)

The test statistic is $t = \frac{\hat{\mu}_D}{s_{\hat{\mu}_D}} = \frac{400}{137.44} = 2.91.$ (3)

The decision rule is "Reject H_0 if $t > t_{0.05} = 1.833$ (**obtained from the $t - tables$ using 9 degrees of freedom**)". (3)

Since $t = 2.91 > t_{0.05} = 1.8333$ we reject H_0 and conclude, at the 0.05 level of significance, that the mean starting annual salary for males exceeds the mean starting annual salary for females. (2)

QUESTION 4 **[25]**

(a) **(15)**

Source	df	SS	MS	F
Brand	2	6.43	3.22	0.4503
Error	12	85.80	7.15	
Total	14	92.23		

(b) **(5)**

The test statistic is $F = 0.4503$.

The decision rule is "Reject H_0 if $F > F_{0.05} = 3.8853$ (**obtained from the** $F - tables$ **using 2 and 12 numerator and denominator degrees of freedom, respectively**)".

(3)

Since $F = 0.4503 < F_{0.05} = 3.8853$ we fail to reject H_0 and conclude, at the 0.05 level of significance, that there are no significant differences among the brand mean distances.

(2)

(c) The errors are independent and normally distributed with mean 0 and constant variance σ^2. **(5)**

TOTAL: [85]