STA2002 Assignment 5

Chen Ang (118010009)

1

	# of Concussions				
	0	**1**	**2**	**≥3**	**Total**
Soccer	45(59.91)	25(17.06)	11(8.34)	10(5.69)	91
N-S Athletes	68(63.20)	15(18.00)	8(8.80)	5(6.00)	96
Non-athletes	45(34.89)	5(0.94)	3(4.86)	0(3.31)	53
Total	158	45	22	15	240

Numbers in parentheses are the expected frequency counts under null hypothesis, with estimators

$$\begin{aligned}\hat{p}_0 &= 158/240 = 0.6583\\ \hat{p}_1 &= 45/240 = 0.1875\\ \hat{p}_2 &= 22/240 = 0.0917\\ \hat{p}_{\geq 3} &= 15/240 = 0.0625\end{aligned}$$

The degree of freedom for the chi-square distribution is

$$(3-1)(4-1) = 6$$

Compute the test statistic:

$$t = 22.029 > \chi^2_{.05}(6) = 12.592$$

Hence we reject homogeneity at 5%.

2

		Writing Surface			
		1	**2**	**3**	$\bar{\mathbf{x}}_{\mathbf{row}}$
Brand of Pen	**1**	$\overset{684.0}{709, 659}$	$\overset{719.5}{713, 726}$	$\overset{652.5}{660, 645}$	685.333
	2	$\overset{676.5}{668, 685}$	$\overset{731.0}{722, 740}$	$\overset{706.0}{692, 720}$	704.500
	3	$\overset{672.0}{659, 685}$	$\overset{675.0}{666, 684}$	$\overset{714.0}{678, 750}$	687.000
	4	$\overset{674.0}{698, 650}$	$\overset{685.0}{704, 666}$	$\overset{709.5}{686, 733}$	689.500
	$\bar{\mathbf{x}}_{\mathbf{col}}$	676.625	702.625	695.500	691.583

The number overset in each cell is the mean of the cell.

(a)

	SS	df	MS	F
Row	1387.50	3	462.50	0.676
Column	2888.08	2	1444.04	2.109
Interaction	8100.25	6	1350.04	1.912
Error	8216.00	12	684.67	
Total	20591.83	23		

(b)

For row effect,

$$F_{\text{row}} = 0.676 < F_{0.05}(3, 12) = 3.49$$

Hence we fail to reject the row null hypothesis at 5%.

(c)

For column effect,

$$F_{\text{col}} = 2.109 < F_{0.05}(2, 12) = 3.89$$

Hence we fail to reject the column null hypothesis at 5%.

(d)

For interaction,

$$F_{\text{int}} = 1.912 < F_{0.05}(6, 12) = 3.00$$

Hence we fail to reject the interaction null hypothesis at 5% as well.

3

(a)

Using Excel,

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Treatment	2.5E+09	2	1.3E+09	338.384	0	2.99602
Error	1.2E+11	31428	3765705			
Total	1.2E+11	31430				

(b)

The realized test statistic can be read off the table:

$$f = 338.384 > F_{0.05}(2, 31428) = 2.996$$

Here the degrees of freedom of the F distribution are $m - 1 = 3 - 1 = 2$, for three groups of weathers, and $n - m = 31430 - 2 = 31428$, for the in-group error.

As the test statistic suggests, we reject the hypothesis of equal means of traffic volume for three weather groups at 5%.