MATU9D2: Practical Statistics **Spring 2017**

Solutions to Weekly Assignment 5

Question 1.

(i)

	Tory	Labour	LibDem	SNP	Other
Male	61	152	49	98	25
Female	92	125	62	105	31
Total	153	277	111	203	56
% Male	$\frac{100\times 61}{153}$				
	= 39.9%	54.9%	44.1%	48.3%	44.6%

% range from 40% - 55% suggesting that there may be a relationship between gender & political preference

(ii) H_0 : Gender & Political Preference Independent
H_1 : Gender & Political Preference Associated

Significance level 0.05

Test Statistic $X^2 = \sum \frac{(O-E)^2}{E} \sim \chi^2(df)$ under H_0
$df = (r-1)(c-1)$

Observed Test Statistic

Observed (O)

	Tory	Labour	LibDem	SNP	Other	
Male	61	152	49	98	25	385
Female	92	125	62	105	31	415
	153	277	111	203	56	800

Expected (E) $\left(= \frac{\text{Row Total} \times \text{Column Total}}{\text{Overall Total}}\right)$

	Tory	Labour	LibDem	SNP	Other	
Male	73.63	133.31	53.42	97.69	26.95	385
Female	79.37	143.64	57.58	105.31	29.05	415
	153	277	111	203	56	800

$$\chi^2 = \frac{(61-73.63)^2}{73.63} + \frac{(152-133.31)^2}{133.31} + \dots + \frac{(31-29.05)^2}{29.05}$$

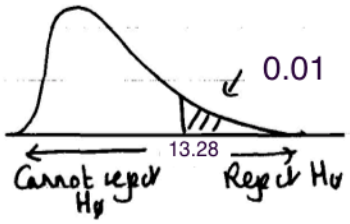
$$= 2.167 + 2.621 + 0.366 + 0.001 + 0.141 + 2.010 + 2.432 + 0.339 + 0.001 + 0.131 = 10.209 //$$

Rejection Region 0.01 ; One tailed ; $\chi^2(4)$

$df = (r-1)\times(c-1)$

$= (2-1)\times(5-1) = 4$

Critical Value = $\chi^2(4; 0.01) = 13.28$



Conclusion : Observed Test Statistic (10.209) is not in the Rejection Regions so we cannot reject Ho in favour of H1 at 1% level i.e. insufficient evidence to suggest that there is an association between gender and political preference.

(iii) 95% CI for θ, proportion of population who vote Labour

$$\hat{\theta} \pm 1.96\sqrt{\frac{\theta(1-\theta)}{n}} \qquad \hat{\theta} = \frac{277}{800} = 0.346$$

$$0.346 \pm 1.96\sqrt{\frac{0.346(1-0.346)}{800}} \qquad n = 800$$

$$0.346 \pm 1.96 \times 0.0168$$

ie 0.346 ± 0.033

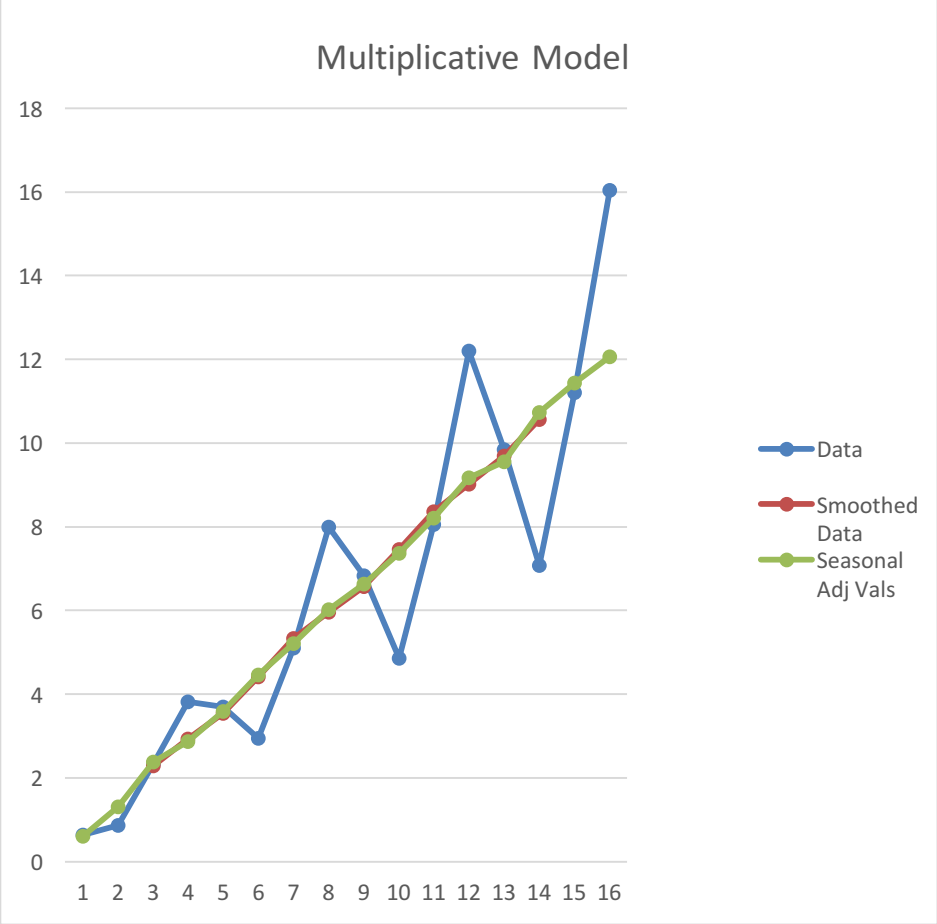
ie $(0.313, 0.379)$

ie true proportion of population who vote Labour lies between 0.31 & 0.379 with 95% confidence.

Question 2.

(b) The most appropriate model is a multiplicative model as the peak-to-trough is not constant and is a 'multiple' of the trend.

(a) The plot of the data (as well as the smoothed data and the seasonally adjusted values) is:



(c)

Multiplicative model

year	Quarter	Model	4pt smooth	Trend	Deviation	SAV
1	1	0.63				0.612
	2	0.86				1.303
			1.910			
	3	2.33		2.293	1.016	2.378
			2.675			
	4	3.82		2.935	1.302	2.872
			3.195			
2	1	3.69		3.541	1.042	3.583
			3.888			
	2	2.94		4.410	0.667	4.455
			4.933			
	3	5.1		5.325	0.958	5.204
			5.718			
	4	8		5.958	1.343	6.015
			6.198			
3	1	6.83		6.566	1.040	6.631
			6.935			
	2	4.86		7.460	0.651	7.364
			7.985			
	3	8.05		8.361	0.963	8.214
			8.738			
	4	12.2		9.015	1.353	9.173
			9.293			
4	1	9.84		9.688	1.016	9.553
			10.083			
	2	7.08		10.563	0.670	10.727
			11.043			
	3	11.21				11.439
	4	16.04				12.060

(d) Using the following table (with the deviation from the trend in the above table) to calculate trend/quarter index

	Qtr 1	**Qtr 2**	**Qtr 3**	**Qtr 4**	
Year 1			1.016	1.302	
Year 2	1.042	0.667	0.958	1.343	
Year 3	1.040	0.651	0.963	1.353	
Year 4	1.016	0.670			
Average	1.033	0.663	0.979	1.333	4.007
Adjustment	1.002	1.002	1.002	1.002	
Quarter Index	1.031	0.662	0.977	1.330	
	1.030	0.660	0.980	1.330	

(e) The graph shows an upward trend, with each quarter increasing year-on-year (first quarter highest, 2nd quarter lowest).

Residuals influences small as the trend and seasonally adjusted values close together.