MATH 204 - SOLUTIONS 5

1. (a) Fitted values:

	Smoking Status			Total
	Non-smoker	Occasional	Regular	
Cough	339.137	357.096	44.767	741
No cough	963.863	1014.904	127.233	2106
Total	1303	1372	172	2847

(b) For the Chi-squared statistic

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{3} \frac{(n_{ij} - \widehat{n}_{ij})^2}{\widehat{n}_{ij}} = 64.247$$

(c) Here $(r-1)(c-1) = 2$, and thus we compare with the Chisquared(2) distribution: for $\alpha = 0.05$, from Tables in McClave and Sincich

$$\text{Chisq}_{0.05}(2) = 5.99 < X^2$$

and thus we **reject** H_0 at $\alpha = 0.05$; in fact the p-value is miniscule (1.12×10^{-14}). Hence we conclude that there **is** evidence for association.

(d) Using the $\widehat{n}_{ij}$ from above

$$LR = 2\sum_{i=1}^{2}\sum_{j=1}^{3} n_{ij} \log(n_{ij}/\widehat{n}_{ij}) = 61.013$$

and thus we again **reject** H_0 at $\alpha = 0.05$

2. Using the formulae given

$$\log\widehat{\psi} = \log\left(\frac{90 \times 307}{255 \times 84}\right) = 0.690 \qquad \text{s.e.}(\log\widehat{\psi}) = \sqrt{\frac{1}{90} + \frac{1}{255} + \frac{1}{84} + \frac{1}{307}} = 0.180$$

and

$$Z = \frac{\log\widehat{\psi}}{\text{s.e.}(\log\widehat{\psi})} = \frac{0.690}{0.179} = 3.837$$

The two-sided test at the $\alpha = 0.05$ significance level has critical values ± 1.96, so as $Z > 1.96$, we **reject** the null hypothesis of no association, which here corresponds to

$$H_0 \ : \ \psi = 1 \qquad \text{or equivalently} \qquad H_0 \ : \ \log\psi = 0$$

Note: The **odds on** an event E, where $P(E) = p$, are given by

$$\frac{p}{1-p}.$$

Here, ψ is the odds-ratio measuring the **change** in odds on having Hodgkin's disease in the two groups, that is, if p_1 and p_2 are the probability of having the disease in the Tonsillectomy and No Tonsillectomy groups respectively, then

$$\psi = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$$

In this example

$$\widehat{\psi} = \left(\frac{90 \times 307}{255 \times 84}\right) = 1.994$$

so the odds on having Hodgkin's disease in the Tonsillectomy group are **almost twice as high** as for the No Tonsillectomy group. Hence it appears that there is a **positive association** between the Tonsillectomy factor and the disease.

3. The data and ranks are summarized below;

Group	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
Obs.	0.73	0.80	0.83	1.04	1.38	1.45	1.46	1.64	1.89	1.91	0.74	0.88	0.9	1.15	1.21
Rank	1	3	4	7	10	11	12	13	14	15	2	5	6	8	9

Thus the Wilcoxon statistic is

$$W = R_2 = 2 + 5 + 6 + 8 + 9 = 30$$

and the Mann-Whitney statistic is

$$U = R_2 - \frac{n_2}{2}(n_2 + 1) = 30 - \frac{5 \times 6}{2} = 15.$$

The Z statistic is thus

$$Z = \frac{U - \dfrac{n_1 n_2}{2}}{\sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{15 - \dfrac{10 \times 5}{2}}{\sqrt{\dfrac{10 \times 5 \times 16}{12}}} = -1.225$$

Comparing this with the standard normal distribution, the critical values of the two-sided test at the $\alpha = 0.05$ significance level are ± 1.96. Thus there is **no evidence to reject H_0** of equal population medians.

Note : An exact test using the Mann-Whitney-Wilcoxon table on page 832 of McClave and Sincich can be carried out. Using the test statistic $W = R_2$, we look up in the table for $n_1 = 10$ and $n_2 = 5$; the table gives

$$T_L = 24 \qquad T_U = 56.$$

As

$$T_L < W < T_U$$

we have again **no evidence to reject H_0** in either a one-sided or two-sided test.