8.10

(a)

a. Table A

Controls	Cases: High	Cases: Low	Total
High	3	1	4
Low	3	1	4
Total	6	2	8

Table B.

Diet	Case	Control
High	1	1
low	0	0

Type 2: 3

	Case	Control
High	0	1
Low	1	0

Type II: 1

	Case	Control
High	1	0
low	0	1

Type III: 3

	Case	Control
High	0	0
Low	1	1

Type IV: 1

(b)

McNemar z^2: $\chi^2 = \frac{(n_{12}-n_{21})^2}{n_{12}+n_{21}} = 1$

As $\chi^2_{0.05,1} = 3.84$, so our test statistic is smaller than $\chi^2_{0.05,1}$. So we fail to reject H_0, two marginal probabilities are equal.

$E(y_{i1} \mid mar, H_0) = 1$ type I table. $Var(y_{i1} \mid mar, H_0) = 0$ type I & IV table

$= 0.5$ Type II or III $= \frac{1}{4}$ type II & III tables

$= 0$ Type IV.

$\chi^2 = \frac{(n_{12}-n_{21})^2}{n_{12}+n_{21}} = 1$ Same as McNemar statistic above.

(c)

If delete the data on case and control, which have same dist. equal to remove type I, type IV table.

$\chi^2_{CMH} = \frac{[n_{12}\times(1-0.5) + n_{21}\times(0-0.5)]^2}{n_{12}\times 0.25 + n_{21}\times 0.25} = 1$. Same as (b). ∴ Not changed.

(d)

$H_0: \pi_{12} = \pi_{21}$ $H_1: \pi_{12} < \pi_{21}$.

as. $n_{21} \mid n_{12}+n_{21} \sim Bin(n_{12}+n_{21}, p_{21}/(p_{12}+p_{21})) = Bin(4, \frac{1}{2})$.

Exact p-value $= P[n_{21}-n_{12} \geq 2 \mid H_0] = P[n_{21}=3 \mid H_0] + P[n_{21}=4 \mid H_0] = 0.3125$

8.14
(a)

The FREQ Procedure

Frequency Percent					
Table of resd16 by resd04					
	resd04				
resd16	1	2	3	4	Total
1	425 16.42	17 0.66	80 3.09	36 1.39	558 21.55
2	10 0.39	555 21.44	74 2.86	47 1.82	686 26.50
3	7 0.27	34 1.31	771 29.78	33 1.27	845 32.64
4	5 0.19	14 0.54	29 1.12	452 17.46	500 19.31
Total	447 17.27	620 23.95	954 36.85	568 21.94	2589 100.00

Statistics for Table of resd16 by resd04

Symmetry Test		
Chi-Square	DF	Pr > ChiSq
119.4321	6	<.0001

Interpretation: As p-value above is <0.0001 which is smaller than 0.05, so we reject H0, so the symmetry model is not applied.

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	0	0.0000	0.0000	0.0000	0.0000	.	.
x1	1	2.1375	0.2745	1.5995	2.6756	60.63	<.0001
x2	1	1.1115	0.2104	0.6992	1.5238	27.92	<.0001
x3	1	0.1570	0.1976	-0.2304	0.5444	0.63	0.4270
Scale	0	1.0000	0.0000	1.0000	1.0000		

Interpretation: for p-value of the β_1 is smaller than 0.05, so we reject H0, $\beta_1 = 0$ is not applied.
for p-value of the β_2 is smaller than 0.05, so we reject H0, $\beta_2 = 0$ is not applied.
for p-value of the β_3 is bigger than 0.05, so we reject H0, $\beta_3 = 0$ is applied.
Not all slopes are not equal to 0 so marginal symmetry model is not applied.

(b)

$$G^2 = 2\sum\left\{n_{ij}\log\left(\frac{2n_{ij}}{n_{ij}+n_{ji}}\right) + n_{ji}\log\left(\frac{2n_{ji}}{(n_{ij}+n_{ji})}\right)\right\} = 2\times\left(17\times\log\frac{2\times17}{17+10}\right) +$$

$$36\times\log\left(\frac{2\times36}{36+5}\right) + 80\times\log\left(\frac{2\times80}{80+7}\right) + 74\times\log\left(\frac{2\times74}{74+34}\right) + 47\times\log\left(\frac{2\times47}{47+14}\right)$$

$$+ 33\log\left(\frac{2\times33}{33+29}\right) + 10\times\log\left(\frac{2\times10}{17\times10}\right) + 5\times\log\left(\frac{2\times5}{36+5}\right) + 7\times\log\left(\frac{2\times7}{80+7}\right)$$

$$+ 34\times\log\left(\frac{2\times34}{74+34}\right) + 14\times\log\left(\frac{2\times14}{47+14}\right) + 29\times\log\left(\frac{2\times29}{33+29}\right)\Big]. \approx 134$$

$df = \frac{4\times(4-1)}{2} = 6$ $\qquad \chi^2_{6,0.05} = 12.592$ $\qquad$ As ~~test~~ $G^2 > \chi^2_{6,0.05}$

so we reject H_0, so we reject symmetry model.

(c)

Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	3	3964.90	<.0001
time	3	170.13	<.0001
Residual	0	.	.

As p-value here is smaller than 0.05, so we reject null hypothesis, so there is difference between them, so the homogeneity is not applied.

8.20

(a)

Statistics for Table of rater1 by rater2

Statistic	DF	Value	Prob
Chi-Square	9	64.7524	<.0001
Likelihood Ratio Chi-Square	9	69.1626	<.0001
Mantel-Haenszel Chi-Square	1	51.4239	<.0001
Phi Coefficient		0.6592	
Contingency Coefficient		0.5504	
Cramer's V		0.3806	

WARNING: 38% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

As p-value above is <0.0001 smaller than 0.05, so we reject H_0, so they are not independent.

Observation Statistics						
Observation	Raw Residual	Pearson Residual	Deviance Residual	Std Deviance Residual	Std Pearson Residual	Likelihood Residual
1	13.194631	2.6492589	2.4549372	4.4276709	4.7781451	4.6732018
2	-5.926174	-1.792835	-2.00876	-2.76001	-2.463332	-2.624664
3	-3.248322	-1.80231	-2.548852	-3.15498	-2.230908	-2.867973
4	-4.020134	-1.79425	-2.193936	-2.776698	-2.270845	-2.598223
5	6.5033557	1.2634035	1.2163854	2.2258746	2.3119135	2.2865583
6	-0.671141	-0.196452	-0.198382	-0.276553	-0.273864	-0.275251
7	-0.469799	-0.252208	-0.258247	-0.324326	-0.316742	-0.321572
8	-5.362416	-2.315689	-3.274879	-4.205276	-2.973579	-3.768897
9	-9.731544	-2.190792	-2.422895	-4.193838	-3.792088	-3.930747
10	5.3087248	1.8007294	1.6526357	2.1792245	2.3745061	2.2642566
11	2.4161074	1.5030693	1.3301106	1.5800908	1.7855553	1.6426151
12	2.0067113	1.0041985	0.9339751	1.1344417	1.2197377	1.162607
13	-9.966443	-2.767768	-3.339217	-5.497801	-4.556948	-4.924998
14	1.2885906	0.5391918	0.5206144	0.6529925	0.6762936	0.6615774
15	1.3020134	0.9991906	0.9005486	1.0175801	1.1290412	1.042756
16	7.3758389	4.5531942	3.4647942	4.0030542	5.2605385	4.3527562

From the residues above, we can get the same conclusion as above due to the residuals difference is not large enough to make it noticed compared to the independent conditions.

(b)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	1.0767	0.2126	0.6600	1.4934	25.64	<.0001
rater1	1	1	-0.3860	0.2331	-0.8428	0.0708	2.74	0.0977
rater1	2	1	1.8705	0.1509	1.5749	2.1662	153.75	<.0001
rater1	3	1	0.6871	0.1357	0.4210	0.9531	25.62	<.0001
rater1	4	0	0.0000	0.0000	0.0000	0.0000	.	.
rater2	1	1	0.5432	0.1897	0.1714	0.9151	8.20	0.0042
rater2	2	1	0.8782	0.1773	0.5306	1.2257	24.52	<.0001
rater2	3	1	-1.2987	0.3015	-1.8897	-0.7077	18.55	<.0001
rater2	4	0	0.0000	0.0000	0.0000	0.0000	.	.
qi	1	1	2.4037	0.2265	1.9597	2.8476	112.61	<.0001
qi	2	1	-1.4275	0.1474	-1.7164	-1.1385	93.74	<.0001
qi	3	1	1.1444	0.3223	0.5127	1.7761	12.61	0.0004
qi	4	1	1.2259	0.2350	0.7654	1.6864	27.22	<.0001
qi	5	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

Quasi-independent model: $\delta_1 = 2.4037, \delta_2 = -1.4275, \delta_3 = 1.1444, \delta_4 = 1.2259$

It seems more on agreement than disagreement between neurologists.

$\tau_{12} = e^{2.4037-1.4275} = 2.654$ $\tau_{13} = e^{2.4037+1.1444} = 34.747$ $\tau_{14} = e^{2.4037+1.2259} = 37.698$

$\tau_{23} = e^{-1.4275+1.1444} = 0.753$ $\tau_{24} = e^{-1.4275+1.2259} = 0.817$ $\tau_{34} = e^{1.1444+1.2259} = 10.701$

Interpretation: For any pair of subjects, the conditional odds ratio on agreements to disagreements is τ_{XY} which neurologist may put in category X and the other in category Y.

(c)

Kappa Statistics				
Statistic	Estimate	Standard Error	95% Confidence Limits	
Simple Kappa	0.2079	0.0505	0.1091	0.3068
Weighted Kappa	0.3797	0.0517	0.2785	0.4810

Test of H0: Kappa = 0				
Estimate	H0 Std Err	Z	Pr > Z	Pr > |Z|
0.2079	0.0456	4.5594	<.0001	<.0001

Test of H0: Weighted Kappa = 0				
Estimate	H0 Std Err	Z	Pr > Z	Pr > |Z|
0.3797	0.0530	7.1620	<.0001	<.0001

Sample Size = 149

From the output above, we can see the p-value correspond to Kappa = 0 is smaller than 0.05, so we reject H_0 and the agreement is not random. And the Kappa weight is 0.3797 so it means that the 37.97% difference between observed and expected agreements conditional on their diagnoses.

8.24

(a)

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	0	0.0000	0.0000	0.0000	0.0000	.	.
x1	1	0.1674	0.5960	-1.0007	1.3355	0.08	0.7788
x2	1	-0.2795	0.5796	-1.4156	0.8565	0.23	0.6296
x3	1	-0.4575	0.7249	-1.8782	0.9632	0.40	0.5279
x4	1	0.5592	0.6957	-0.8044	1.9228	0.65	0.4215
Scale	0	1.0000	0.0000	1.0000	1.0000		

From the output above, the $\beta_1 = 0.1674$, $\beta_2 = -0.2795$, $\beta_3 = -0.4575$, $\beta_4 = 0.5592$
So the rank is S.williams > Clijsters > V.williams > Davenport > Pierce

(b) probability that Serena Williams beats Venus Williams: $\frac{e^{\beta_4-\beta_5}}{1+e^{\beta_4-\beta_5}} = 0.6363$
sample proportion: 2/(2+2)=1/2
Compared the result above, the probability that Serena Williams beats Venus Williams is higher than the sample proportion of them.

(c)

Estimated Covariance Matrix				
	Prm2	Prm3	Prm4	Prm5
Prm2	0.35521	0.19101	0.19048	0.15389
Prm3	0.19101	0.33598	0.16154	0.16596
Prm4	0.19048	0.16154	0.52544	0.16395
Prm5	0.15389	0.16596	0.16395	0.48405

$$var(\beta_4 - \beta_5) = var(\beta_4) + var(\beta_5) - 2cov(\beta_4, \beta_5) = 0.48405$$

90% CI($\beta_4 - \beta_5$): $0.5592 \mp 1.645 * \sqrt{0.48405} = (-0.5852, 1.7037)$

So for $\Pi = \left(\frac{e^{-0.5852}}{1+e^{-0.5852}}, \frac{e^{1.7037}}{1+e^{1.7037}}\right)$ =(0.3577,0.8640)

Interpretation: We are 90% confident the probability that SerenaWilliams beats VenusWilliams is in (0.3577,0.8640).

(d)

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.5611	4	0.6337
Score	2.5096	4	0.6429
Wald	2.4082	4	0.6611

As p-value for Wald test is 0.6611 which is bigger than 0.05, so we fail to reject H_0, there is no difference among players.