Solution to HW10

Problem 8.10 (5 pts)

(a) The 2×2 cross-classification table of diet for the case against diet for the control is:

		Diet for the case (Y_2)	
		High (1)	Low (0)
Diet for the control (Y_1)	High (1)	3	1
	Low (0)	3	1

Denote $Y = 1/0$ for case/control, $x = 1/0$ for high/low level of red meat. Then partial tables are:

		Y							
		1	0	1	0	1	0	1	0
X	1	1	1	0	1	1	0	0	0
	0	0	0	1	0	0	1	1	1
		$n_{11} = 3$		$n_{12} = 1$		$n_{21} = 3$		$n_{22} = 1$	

(b) The McNemar's test statistic z^2 for Table A is:

$$z^2 = \frac{(1-3)^2}{1+3} = 1.$$

The CMH χ^2 for Table B is:

$$\frac{\{3 \times (1-1) + 1 \times (0-1/2) + 3 \times (1-1/2) + 1 \times (0-0)\}^2}{0 + 1 \times (1/4) + 3 \times (1/4) + 0} = \frac{1}{1} = 1.$$

(c) The CMH χ^2 for Table B after deleting tables where case and control had the same diet (table 1 & table 4) is:

$$\frac{1 \times (0-1/2) + 3 \times (1-1/2)}{1 \times (1/4) + 3 \times (1/4)} = \frac{1}{1} = 1.$$

It is unchanged.

(d) Under marginal homogeneity $\pi_{12} = \pi_{21}$,

$$n_{21} | n^* = n_{12} + n_{21} \sim \text{Bin}(n^*, \pi_{21}/(\pi_{12} + \pi_{21})) = \text{Bin}(4, 0.5).$$

Since larger n_{21} represents the case has a higher probability of having higher level red meat diet than control. Therefore, the one-sided exact P-value will be

$$\begin{aligned} \text{P-value} &= P(n_{21} \geq 3 | n_{12} + n_{21} = 4) \\ &= P(n_{21} = 3 | n_{12} + n_{21} = 4) + P(n_{21} = 4 | n_{12} + n_{21} = 4) \\ &= 0.25 + 0.0625 = 0.3125. \end{aligned}$$

The exact mid P-value is:

$$\begin{aligned} \text{mid P-value} &= 0.5P(n_{21} = 3|n_{12} + n_{21} = 4) + P(n_{21} = 4|n_{12} + n_{21} = 4) \\ &= 0.5 \times 0.25 + 0.0625 = 0.1875. \end{aligned}$$

Note: We can also use the exact CMH test for $H_0 : \pi_{12} = \pi_{21}$. But the exact P-value and mid P-value are two-sided:

```
data prob8_10;
   input contdiet y1 y2;
   datalines;
   1 3 1
   0 3 1
   ;

data prob8_10; set prob8_10;
  array temp {2} y1-y2;

  do j=1 to 2;
    count=temp(j);
    casediet = 2-j;
    output;
  end;
run;

data newdata; set prob8_10;
  retain pair;
  if _n_=1 then pair=0;

  do i=1 to count;
    pair = pair + 1;

    do cancer=0 to 1;
      if cancer=0 then
        x = contdiet;
      else
        x = casediet;
      output;
    end;
  end;
run;

title "Exact CMH test";
proc logistic descending;
    class pair / param=ref;
    model cancer = pair x / link=logit;
    exact x;
run;
```

```
**********************************************************************

                        The LOGISTIC Procedure

                      Exact Conditional Analysis

                       Exact Conditional Tests

                                                  --- p-Value ---
          Effect   Test          Statistic    Exact      Mid

          x        Score            1.0000   0.6250   0.5000
                   Probability      0.2500   0.6250   0.5000
```

Problem 8.14 (5 pts)

(a) The probability that residence at age 16 is the same as that in 2004 is

$$\widehat{P}(\text{Same residence}) = (425 + 555 + 771 + 452)/2589 = 0.85.$$

Under the symmetry model, other estimated probabilities are:

$\widehat{P}$(Northeast to Midwest) = $\widehat{P}$(Midwest to Northeast) = $(17 + 10)/(2 \times 2589) = 0.005$

$\widehat{P}$(Northeast to South) = $\widehat{P}$(South to Northeast) = 0.017

$\widehat{P}$(Northeast to West) = $\widehat{P}$(West to Northeast) = 0.008

$\widehat{P}$(Midwest to South) = $\widehat{P}$(South to Midwest) = 0.02

$\widehat{P}$(Midwest to West) = $\widehat{P}$(West to Midwest) = 0.01

$\widehat{P}$(South to West) = $\widehat{P}$(West to South) = 0.01

The SAS program for fitting quasi-symmetry model

$$\log(\pi_{ij}/\pi_{ji}) = \beta_i - \beta_j$$

and part of the output are:

```
data prob8_14;
  input resid16 n1-n4;
  datalines;
   1 425  17  80  36
   2  10 555  74  47
   3   7  34 771  33
   4   5  14  29 452
;

data prob8_14; set prob8_14;
    array temp {4} n1-n4;

  do resid04=1 to 4;
    count=temp(resid04);
    output;
  end;
run;

title "Test symmetry";
proc freq;
  weight count;
  tables resid16*resid04 / norow nocol;
  test agree;
run;

data prob8_14; set prob8_14;
  if resid16=resid04 then delete;

  if resid16<resid04 then do;
    y=1; ind1=resid16; ind2=resid04;
  end;
  else do;
    y=0; ind1=resid04; ind2=resid16;
  end;

  array x {4};
  do k=1 to 4;
    if k=ind1 then
      x[k]=1;
    else if k=ind2 then
      x[k]=-1;
  else
    x[k]=0;
  end;
```

```
  drop n1-n4 k;
run;

proc sort;
  by ind1 ind2 descending y;
run;

proc print;
run;

title "Quasi-symmetry model";
proc genmod descending;
  freq count;
  model y = x1 x2 x3 / dist=bin link=logit aggregate noint;
run;
```

```
************************************************************
```

Statistics for Table of resid16 by resid04

Test of Symmetry	
Statistic (S)	119.4321
DF	6
Pr > S	<.0001

Quasi-symmetry model

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	3	3.9324	1.3108
Pearson Chi-Square	3	3.9030	1.3010

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square
Intercept	0	0.0000	0.0000	0.0000	0.0000	.
x1	1	2.1375	0.2745	1.5995	2.6756	60.63
x2	1	1.1115	0.2104	0.6992	1.5238	27.92
x3	1	0.1570	0.1976	-0.2304	0.5444	0.63
Scale	0	1.0000	0.0000	1.0000	1.0000	

The estimated β's are: $\widehat{\beta}_1 = 2.14 > \widehat{\beta}_2 = 1.11 > \widehat{\beta}_3 = 0.16 > \widehat{\beta}_4 = 0$, implying

$$\begin{aligned}
\frac{\widehat{\pi}_{1j}}{\widehat{\pi}_{j1}} &= e^{\widehat{\beta}_1 - \widehat{\beta}_j} > 1 \quad j = 2, 3, 4 \\
\frac{\widehat{\pi}_{2j}}{\widehat{\pi}_{j2}} &= e^{\widehat{\beta}_2 - \widehat{\beta}_j} > 1 \quad j = 3, 4 \\
\frac{\widehat{\pi}_{3j}}{\widehat{\pi}_{j3}} &= e^{\widehat{\beta}_2 - \widehat{\beta}_j} > 1 \quad j = 4
\end{aligned}$$

Therefore, people whose residence was in Northeast at age 16 were more likely to move to other places than people whose residence in other places at age 16 to move to Northeast, etc.

(b) The Pearson χ^2 goodness-of-fit for the symmetric model indicates that we reject the symmetry model (almost at any level).

The LRT test (calculated by hand) is $G^2(S) = 134.46$, with $df = 6$. The same conclusion as the Pearson χ^2 goodness-of-fit test. Note, this G^2 is also the deviance of the symmetry model.

The Pearson χ^2 goodness-of-fit P-value for the quasi-symmetry model is $P(\chi^2_3 \geq 3.9030) = 0.27$, indicating a reasonable fit of the quasi-symmetry model to the data.

The deviance $G^2(QS) = 3.9324$ for the quasi-symmetry model indicates the goodness-of-fit P-value $= P(\chi^2_3 \geq 3.9324) = 0.27$, also indicating a reasonable fit of the quasi-symmetry model to the data.

Given the quasi-symmetry model, the LRT test for the symmetry model is

$$LRT = 134.46 - 3.9324 = G^2(S) - G^2(S) = 130.53 \text{with} df = 6 - 3 = 3.$$

So we reject the symmetry model assuming the quasi-symmetry model.

(c) SAS program for marginal homogeneity is based on the difference of marginal proportions is

```
title "Testing Marginal Homogeneity";
proc catmod data=prob8_14;
  weight count;
  response marginals;
  model resid16*resid04 = _response_;
  repeated time 2;
run
```

```
*****************************************************************
                        Analysis of Variance

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

The test statistic is $\chi^2 = 170$ with $df = 3$. Therefore, we reject the marginal homogeneity. That is, the residence at age 16 is different from the residence in 2004.

Problem 8.20 (5 pts)

(a) Under the independence model, the standardized Pearson residuals are

		rater 2			
		1	2	3	4
rater1	1	4.48	-2.46	-2.23	-2.27
	2	2.31	-0.27	-0.32	-2.97
	3	-3.79	2.37	1.79	1.22
	4	-4.56	0.68	1.13	5.26

From the above table, we know that these two raters agree with each for category 1 and 4 much more than expected by the independence model. The disagreement occurs most in row 1 or column 1. The first row indicates that when rater 1 rates category 1, rater 2 disagrees with rater 1 much less than expected by the independence model. The first column indicates

that when rater 2 rates category 1, rater 2 disagrees with rater 1 much more or much less than expected by the independence model.

(b) We can use a quasi-independence model to describe the pattern and strength of the agreement. The estimated δ's are:

```
data prob8_20;
  input rater1 n1-n4;
  datalines;
    1 38 5 0 1
    2 33 11 3 0
    3 10 14 5 6
    4 3 7 3 10
   ;

data prob8_20; set prob8_20;
  array temp {4} n1-n4;

  do rater2=1 to 4;
    count=temp(rater2);
    if rater1=rater2 then
      qi=rater1;
    else
      qi=5;
    output;
  end;
run;

title "Problem 8.20(b) - quasi-independence model";
proc genmod data=prob8_20;
  class rater1 rater2 qi;
  model count = rater1 rater2 qi / dist=poi link=log;
run;
```

```
*******************************************************************************
               Analysis Of Maximum Likelihood Parameter Estimates

                              Standard       Wald 95%             Wald
Parameter     DF  Estimate       Error   Confidence Limits  Chi-Square  Pr > ChiSq

qi         1   1    2.2231    0.5097    1.2241    3.2221       19.02      <.0001
qi         2   1   -1.0385    0.4516   -1.9236   -0.1535        5.29      0.0215
qi         3   1    0.8601    0.6556   -0.4248    2.1450        1.72      0.1895
qi         4   1    2.4197    0.5846    1.2738    3.5656       17.13      <.0001
qi         5   0    0.0000    0.0000    0.0000    0.0000           .           .
Scale          0    1.0000    0.0000    1.0000    1.0000
```

$\widehat{\delta}_1 = 2.2231, \widehat{\delta}_2 = -1.0385, \widehat{\delta}_3 = 0.8601, \widehat{\delta}_4 = 2.4197.$

So for any a, b except $a = 2, b = 3$ or $a = 3, b = 2$, $\widehat{\tau}_{ab} = exp(\widehat{\delta}_a \widehat{\delta}_b) > 1$, that is, the two raters agreed with each other more than disagreed.

However, for $a = 2, b = 3$ or $a = 3, b = 2$,

$$\widehat{\tau}_{23} = exp(-1.0385 + 0.8601) = 0.83.$$

That is, given that rater 1 and rater 2 put two patients in categories 2 or 3, the odds that they put the patients in the same categories than different categories is 0.81. So for these two categories, they disagreed more than agreed.

(c) The simple κ estimate is $\widehat{\kappa} = 0.21$. That is, the difference the observed agreement and that expected under independence is only about 21% of the maximum possible difference.

The weighted κ estimate is $\widehat{\kappa}_w = 0.38$, indicating that the observed weighted agreement and that expected under independence is only about 38% of the maximum possible difference.

Problem 8.24 (5 pts)

(a) The output of the Bradley-Terry model is:

```
data prob8_24;
  input winner player $ n1-n5;
  datalines;
  1 Clijsters  . 6 3 0 2
  2 Davenport  2 . 0 2 4
  3 Pierce     1 2 . 0 1
  4 S.Williams 2 2 2 . 2
  5 V.Williams 3 2 2 2 .
  ;

data prob8_24; set prob8_24;
  array temp {5} n1-n5;

  do loser=1 to 5;
    count=temp(loser);
    output;
  end;
run;

data prob8_24; set prob8_24;
  if winner=loser then delete;

  if winner<loser then do;
     y=1; ind1=winner; ind2=loser;
  end;
  else do ;
     y=0; ind1=loser; ind2=winner;
  end;

  array x {5};
  do k=1 to 5;
    if k=ind1 then
      x[k]=1;
    else if k=ind2 then
      x[k]=-1;
    else
      x[k]=0;
  end;
  drop n1-n5 k;
run;

proc sort;
    by ind1 ind2 descending y;
run;

proc print;
run;

title "Bradley-Terry Model for Woman Tennis Matches";
proc logistic descending covout;
  freq count;
  model y = x1 x2 x3 x4 / link=logit aggregate scale=none noint;
run;

*******************************************************************************

                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

                  Analysis of Maximum Likelihood Estimates

                                    Standard          Wald
```

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
x1	1	0.1674	0.5960	0.0789	0.7788
x2	1	-0.2795	0.5796	0.2325	0.6296
x3	1	-0.4574	0.7249	0.3982	0.5280
x4	1	0.5590	0.6957	0.6457	0.4217

The parameter estimates are: $\widehat{\beta}_1 = 0.1674$, $\widehat{\beta}_2 = -0.2795$, $\widehat{\beta}_3 = -0.4575$, $\widehat{\beta}_4 = 0.5592$, $\widehat{\beta}_5 = 0$. So $\widehat{\beta}_1 > \widehat{\beta}_4 > \widehat{\beta}_5 > \widehat{\beta}_2 > \widehat{\beta}_3$. This gives the ranking of these 5 players:

Clijsters, S.Williams, V.Williams, Davenport, Pierce.

(b) The probability estimate is $\widehat{\pi}_{45}$:

$$\widehat{\pi}_{45} = \frac{e^{0.5592}}{1 + e^{0.5592}} = 0.636.$$

The sample proportion is $2/4 = 0.5$. If we only used the sample proportion, we cannot rank S.Williams and V.Williams. However, if we use the Bradley-Terry model, we can borrow the information from other matches to rank them.

(c) A 90% Wald CI for β_4 is $\widehat{\beta}_4 \pm 1.645 \times SE(\widehat{\beta}_4) = 0.5592 \pm 1.645 \times 0.6957 = [-0.585, 1.704]$. So a 90% Wald CI for π_{45} is

$$\left[\frac{e^{-0.585}}{1 + e^{-0.585}}, \frac{e^{1.704}}{1 + e^{1.704}}\right] = [0.358, 0.846].$$

We are 90% confident that S. Williams would beat V. Williams with probability at least 0.358 and at most 0.846.

(d) The LRT for $H_0 : \beta_i = 0$ is $G^2 = 2.6$ with $df = 4$, resulting a p-value=0.63. Therefore, it is plausible that these players just won games by chance. Of course, the LRT may not be stable due to the small sample size.