BIOS 765 Homework 4

Leo Li
730031954

Problem 1(a).

In this question, we would like to compute the standard midrank scores for the nine categories of the response, survival time, as well as the logrank scores that place greater emphasis on group differences occurring at larger survival times. Specifically, the standardized midrank scores are calculated as,

$$a_j = \frac{\sum_{k=1}^{j-1} n_{.k} + (n_{.j}+1)/2}{n+1},$$

whereas the logrank scores are calculated as,

$$a_j = 1 - \sum_{k=1}^{j} \left(\frac{n_{.k}}{\sum_{m=k}^{r} n_{.m}}\right).$$

The results are presented in Table 1.

Table 1: Scores (a_j) for the Survival Time (Hours) of Mice to a Drug Challenge

Survival Time	Standard Midranks	Logranks
0-6	0.065	0.909
6-12	0.196	0.709
12-18	0.413	0.334
18-24	0.565	0.234
24-30	0.652	-0.099
30-48	0.761	-0.433
48-72	0.848	-0.933
72-96	0.913	-1.433
>96	0.957	-2.433

Problem 1 (b).

We would like to conduct the Wilcoxon Rank Sum Test to test the equivalence of the two drugs in terms of the survival time of mice to the drug challenge. The null hypothesis is that the two drugs are equivalent in terms of the survival time of mice to the drug challenge; and the alternative hypothesis is that the two drugs are not equivalent in terms of the survival time of mice to the drug challenge. The chi-square test statistic is 3.512, which asymptotically follows a χ_1^2 distribution, and the corresponding p-value is 0.061, which is greater than 0.05, and we fail to reject the null hypothesis. Our conclusion is that there is still insufficient evidence to reject the null hypothesis that the two drugs are equivalent in terms of the survival time of mice to the drug challenge.

Problem 1 (c).

We would like to conduct the Randomization Chi-square Test with logrank scores to test the equivalence of the two drugs in terms of the survival time of mice to the drug challenge. The null hypothesis is that the two drugs are equivalent in terms of the survival time of mice to the drug challenge; and the alternative hypothesis is that the two drugs are not equivalent in terms of the survival time of mice to the drug challenge. The test statistic is 4.057, which asymptotically follows a χ_1^2 distribution, and the corresponding p-value is 0.044, which is less than 0.05, so that we reject the null hypothesis. Our conclusion is that there is sufficient evidence to reject the null hypothesis that the two drugs are equivalent in terms of the survival time of mice to the drug challenge.

Problem 2 (a).

Table 2 presents the reference set of tables and their probabilities giving the permutation distribution of outcomes under the null hypothesis of no association between judges' origins and preferred skater.

Table 2: Reference Set of Tables and Corresponding Probabilities

y_{11}	y_{12}	y_{21}	y_{22}	Probabilities
3	0	1	5	0.048
2	1	2	4	0.357
1	2	3	3	0.476
0	3	4	2	0.119

Here, we would like to use the two-sided Fisher's exact test to test the null hypothesis that there is no association between the judges' origin and the preferred skater; and the alternative hypothesis is that there is some association between the judges' origin and the preferred skater. Then, the p-value for the two-sided Fisher's exact test is 0.048, which is less than 0.05, which lead us to reject the null hypothesis. Therefore, we conclude that there is sufficient evidence for us to reject the null hypothesis that there is no association between the judges' origin and the preferred skater.

In addition, we would also like to use the exact Pearson chi-square test to test the null hypothesis that there is no association between the judges' origin and the preferred skater; and the alternative hypothesis is that there is some association between the judges' origin and the preferred skater. Then, the p-value for the exact Pearson chi-square test is also 0.048, which is less than 0.05, which lead us to reject the null hypothesis. Therefore, we conclude that there is sufficient evidence for us to reject the null hypothesis that there is no association between the judges' origin and the preferred skater.

Problem 2 (b).

Here, we would like to use the Fisher's exact test to test the null hypothesis that there is no association between the judges' origin and the preferred skater; and the alternative hypothesis is that there is some association between the judges' origin and the preferred skater. In Table 3, we present the reference set of tables and their probabilities giving the permutation sample space under the null hypothesis of no association between judges' origins and preferred skater. According to Table 3, the Fisher's exact test for p-value for Table B is calculated as,

$$(0.032 + 0.016 + 0.024 + 0.008) = 0.079,$$

which is greater than 0.05, which cannot lead us to reject the null hypothesis. Therefore, we conclude that there is still insufficient evidence to reject the null hypothesis that there is no association between the judges' origin and the preferred skater.

Table 3: Reference Set of Tables and Corresponding Probabilities and Q_P

y_{11}	y_{12}	y_{21}	y_{22}	y_{31}	y_{32}	Probabilities	Q_P
3	0	1	3	0	2	0.032	5.963
3	0	0	4	1	1	0.016	6.975
2	1	2	2	0	2	0.143	2.250
2	1	1	3	1	1	0.190	1.238
2	1	0	4	2	0	0.024	6.300
1	2	3	1	0	2	0.095	3.263
1	2	2	2	1	1	0.286	0.225
1	2	1	3	2	0	0.095	3.263
0	3	4	0	0	2	0.008	9.000
0	3	3	1	1	1	0.063	3.938
0	3	2	2	2	0	0.048	4.950

Problem 2 (c).

In this question, we would like to use the exact Pearson Chi-square test to test the null hypothesis that there is no association between the judges' origin and the preferred skater; and the alternative hypothesis is that there is some association between the judges' origin and the preferred skater. In Table 3, we also present the Pearson Chi-square test statistic, and according to Table 3, the following tables in the reference set have a value of Q_P at least as large as that of the observed table:

- The table represented by the first row (observed table): $(3, 0, 1, 3, 0, 2)$;
- The table represented by the second row: $(3, 0, 0, 4, 1, 1)$;
- The table represented by the fifth row: $(2, 1, 0, 4, 2, 0)$;
- The table represented by the ninth row: $(0, 3, 4, 0, 0, 2)$.

Therefore, the Pearson Chi-square statistic (Q_P) exact p-value is computed as,

$$(0.032 + 0.016 + 0.024 + 0.008) = 0.079,$$

which coincides with the Fisher's exact test p-value in part (b). The p-value is greater than 0.05, which cannot lead us to reject the null hypothesis. Therefore, we conclude that there is still insufficient evidence to reject the null hypothesis that there is no association between the judges' origin and the preferred skater.

Problem 2 (d).

In this question, we would like to use the mean score test to test the null hypothesis that there is no association between the judges' origin and the preferred skater; and the alternative hypothesis is that there is some association between the judges' origin and the preferred skater. The mean score statistic for each table in the reference set can be calculated as that,

$$Q_S = \frac{(n-1)\sum_{i=1}^{2} n_{i.}(\bar{f}_i - \mu_a)^2}{n v_a},$$

where,

$$\bar{f}_i = \sum_{j=1}^{r} a_j y_{ij}/n_{i.},$$

$$\mu_a = \sum_{j=1}^{r} a_j n_{.j}/n,$$

$$v_a = \sum_{j=1}^{r} (a_j - \mu_a)^2 (n_{.j}/n).$$

Table 4 presents the value of Q_S, the mean score statistic for each table in the reference set. Then the exact p-value for Q_S is calculated as 0.032, which is lower than 0.05, so that we reject the null hypothesis. Therefore, we conclude that there is sufficient evidence to reject the null hypothesis that there is no association between the judges' origin and the preferred skater.

Table 4: Reference Set of Tables and Corresponding Probabilities and Q_S

y_{11}	y_{12}	y_{21}	y_{22}	y_{31}	y_{32}	Probabilities	Q_S
3	0	1	3	0	2	0.032	4.809
3	0	0	4	1	1	0.016	1.782
2	1	2	2	0	2	0.143	1.782
2	1	1	3	1	1	0.190	0.227
2	1	0	4	2	0	0.024	0.145
1	2	3	1	0	2	0.095	0.227
1	2	2	2	1	1	0.286	0.145
1	2	1	3	2	0	0.095	1.536
0	3	4	0	0	2	0.008	0.145
0	3	3	1	1	1	0.063	1.536
0	3	2	2	2	0	0.048	4.400

Problem 2 (e).

In part a, b, c, and d, we essentially use different type of approaches to test the association between judges' origin and preferred skater. The result from each test can be summarized as follow:

- If we collapse Table B into a 2×2 table by combining WE and NA into a single category, then the p-value for the two-sided Fisher's exact test as well as the exact Pearson chi-square test are both 0.048, which is lower than 0.05, which lead us to reject the null hypothesis, and our conclusion is that there is sufficient evidence to reject the null hypothesis that the judges' origin is not associated with the preferred skater.

- If we do not collapse Table B and just leave it as a 3×2 table to conduct the analysis, then the p-value for the Fisher's exact test as well as the Pearson Chi-square statistic (Q_P) exact p-value are both 0.079, which is greater than 0.05, which cannot lead us to reject the null hypothesis, and our conclusion is that there is still insufficient evidence to reject the null hypothesis that the judges' origin is not associated with the preferred skater.

- If we treat region as an ordinal variable having table scores (1 =EE, 2 =WE, 3 =NA), then the exact p-value for the Q_S is 0.032, which is lower than 0.05, which lead us to reject the null hypothesis, and our conclusion is that there is sufficient evidence to reject the null hypothesis that the judges' origin is not associated with the preferred skater.

Based on the above results, I would conclude that there is some geopolitically-based bias in figure skating judging. Among the three different ways of looking at the data, two of them will result in significant results to reject the null hypothesis that the judges' origin is not associated with the preferred skater. Even for the only setting in which the null hypothesis is not rejected, the resulted p-value is also very small.

There are also some assumptions of the statistical methods that lead to the conclusion:

- In the contingency table, we assume that both row and column totals are regarded as fixed, and the probability of observing a particular table is given by the multiple hypergeometric distribution, which can be derived from finite sampling arguments, or by conditioning on column marginal totals subsequent to product-multinomial sampling based on fixed row totals.
- If the method in part a is used, then we need to assume that the geopolitical ideas among the judges from WE and NA are very similar compared to those from EE.
- If the method in part d is used, then we need to assume that the table scores and appropriate ordering are appropriate.

Problem 3 (a).

We perform the analysis using the conditional logistic regression model, and the results of the model fitting are reported in Table 5.

Table 5: Analysis of Conditional Maximum Likelihood Estimates

Effect	Estimate	SE	Wald χ^2	p-value
rainsuit	-2.079	1.061	3.844	0.050

The conditional maximum likelihood estimate of the odds ratio for the association of NOT wearing a rainsuit and GTS is 8.000, and the corresponding 95% confidence interval is $(1.001, 63.963)$.

Problem 3 (b).

Now, we apply exact inference to the model in part a, and the exact conditional maximum likelihood estimate of the odds ratio is 8.000, and the corresponding "exact" 95% confidence interval is $(1.073, 354.981)$.

Problem 3 (c).

In this question, we consider the following stratified logistic regression:

$$\log(\frac{\pi_{ij}}{1-\pi_{ij}}) = \alpha_i + x_{ij1}\beta_1 + x_{ij2}\beta_2$$

where $i = 1, \ldots, 18$, and $j = 1, 2$; π_{ij} is the probability that the j^{th} subject from the i^{th} matched pair had a diagnosis of GTS; $x_{ij1} = 1$ if that person reported wearing a rainsuit, and $x_{ij1} = 0$ if he did not wear a rainsuit; $x_{ij2} = 1$ if that person understood English, and

Table 6: Analysis of Conditional Maximum Likelihood Estimates

Effect	Estimate	SE	Wald χ^2	p-value
rainsuit	-2.287	1.156	3.912	0.048
English	-1.322	0.813	2.643	0.104

$x_{ij2} = 0$ if he did not understand English. The results of the model fitting are reported in Table 6.

The conditional maximum likelihood estimate of the odds ratio for the association of NOT wearing a rainsuit and GTS is 9.843, and the corresponding 95% confidence interval is $(1.021, 94.906)$. The conditional maximum likelihood estimate of the odds ratio for the association of understand English and GTS is 0.266, and the corresponding 95% confidence interval is $(0.054, 1.312)$.

Problem 3 (d).

Now, we apply exact inference to the model in part c, and the exact conditional maximum likelihood estimate of the odds ratio for the association of NOT wearing a rainsuit and GTS is 8.482, and the corresponding 95% confidence interval is $(1.029, 410.792)$. The conditional maximum likelihood estimate of the odds ratio for the association of understand English and GTS is 0.279, and the corresponding 95% confidence interval is $(0.029, 1.438)$.

Problem 3 (e).

Based upon the exact inference results in parts b and d, we can see that the 95% confidence interval of the odds ratio for the association of NOT wearing a rainsuit and GTS is $(1.073, 354.981)$ without adjustment for understanding English, and $(1.029, 410.792)$ with adjustment for understanding English. Since the odds ratio of 1 would indicate no association, and 1 is not included in both 95% confidence intervals. Therefore, there is a statistically significant association not wearing a raincoat and GTS. Specifically, those workers who did not wear rainsuit are much more likely to have GTS regardless of whether we adjust for the effect of understanding English.

Problem 4 (a).

We consider the stratified logistic regression model in problem 3a,

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha_i + x_{ij}\beta.$$

Then the unconditional likelihood function is that,

$$\begin{aligned} L(\alpha_i, \beta) &= \prod_{i=1}^{18}\prod_{j=1}^{2}\{[\pi_{ij}(x_{ij})]^{y_{ij}}[1-\pi_{ij}(x_{ij})]^{1-y_{ij}}\} \\ &= \frac{\exp[\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}(\alpha_i + x_{ij}\beta)]}{\prod_{i=1}^{18}\prod_{j=1}^{2}[1+\exp(\alpha_i + x_{ij}\beta)]}. \end{aligned}$$

In the above unconditional likelihood function, we can see that, $T_i = \sum_{j=1}^{2} y_{ij}$ is a sufficient statistic for α_i, and $\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}$ is a sufficient statistic for β. Then the conditional likelihood can be derived as,

$$\begin{aligned} L_c(\beta) &= \frac{\exp[\sum_{i=1}^{18} t_i\alpha_i + \sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]}{\sum_R \exp[\sum_{i=1}^{18} t_i\alpha_i + \sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]} \\ &= \frac{\exp[\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]}{\sum_R \exp[\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]}, \end{aligned}$$

where $R = \{(y_{ij}), i = 1, \ldots, 18; j = 1, 2 : \sum_{j=1}^{2} y_{ij} = t_i, \forall i\}$. In the context of matched pairs in question 3a, the conditional likelihood can be further reduced as that,

$$\begin{aligned} L_c(\beta) &= \frac{\exp[\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]}{\sum_R \exp[\sum_{i=1}^{18}\sum_{j=1}^{2} y_{ij}x_{ij}\beta]} \\ &= \prod_{i=1}^{18} \frac{\exp[y_{i1}x_{i1}\beta + y_{i2}x_{i2}\beta]}{\sum_{R_i} \exp[y_{i1}x_{i1}\beta + y_{i2}x_{i2}\beta]}, \end{aligned}$$

where $R_i = \{(y_{i1}, y_{i2}) : y_{i1} + y_{i2} = 1\}$. For the i^{th} population, the likelihood that $y_{i1} = 1$ given $y_{i1} + y_{i2} = 1$ is that,

$$\frac{\exp[(x_{i1} - x_{i2})\beta]}{\exp[(x_{i1} - x_{i2})\beta] + 1},$$

so that the conditional likelihood can be further simplified as,

$$\begin{aligned} L_c(\beta) &= \prod_{i=1}^{18} \frac{\exp[(x_{i1} - x_{i2})\beta]}{\exp[(x_{i1} - x_{i2})\beta] + 1} \\ &= (\frac{1}{2})^9 (\frac{\exp(-\beta)}{1 + \exp(-\beta)})^8 (\frac{\exp(\beta)}{1 + \exp(\beta)}) \\ &= \frac{\exp(\beta)}{512(1 + \exp(\beta))^9}, \end{aligned}$$

which is in the simplest form.

Problem 4 (b).

The conditional maximum likelihood estimate $\hat{\beta}$ can be determined by maximizing the conditional likelihood in part a. First, we can derive the conditional log likelihood as that,

$$l_c(\beta) = \beta - \log(512) - 9\log(1 + \exp(\beta)).$$

Then, the conditional MLE can be obtained as the following:

$$\frac{d}{d\beta} l_c(\beta) = 1 - 9\frac{\exp(\beta)}{1 + \exp(\beta)} = 0.$$

Then,

$$\begin{aligned} 9\exp(\hat{\beta}) &= 1 + \exp(\hat{\beta}), \\ \hat{\beta} &= \log(1/8) = -2.079, \end{aligned}$$

which is the same as our results in problem 3a.

Appendix: SAS Code

```
**********************************************
           Question 1
**********************************************;

data q1b;
input drug $ rank;
cards;
b 1
b 2
a 3.5
b 3.5
b 5.5
b 5.5
a 8
b 8
b 8
a 10
a 13.5
a 13.5
a 13.5
a 13.5
a 13.5
b 13.5
a 18.5
a 18.5
a 18.5
b 18.5
a 21.5
b 21.5
;
run;

title '1 (b): Wilcoxon Rank Sum Test';
proc npar1way wilcoxon data=q1b;
  class drug;
  var rank;
run;

data q1c;
input drug $ surv count logrank;
cards;
A 1 1 0.90909
A 2 3 0.70909
A 3 5 0.33409
A 4 1 0.23409
A 5 1 -0.09924
A 6 0 -0.43258
A 7 1 -0.93258
A 8 0 -1.43258
A 9 0 -2.43258
B 1 1 0.90909
B 2 1 0.70909
B 3 1 0.33409
```

```
B 4 0 0.23409
B 5 2 -0.09924
B 6 2 -0.43258
B 7 1 -0.93258
B 8 1 -1.43258
B 9 1 -2.43258
;
run;

title '1 (c): Randomization Chi-square Test with Logrank Scores';
proc freq data=q1c;
  format surv;* fsev.;
  weight count;
  tables drug*logrank/nopercent nocol cmh2;
run;

***********************************************
           Question 2
***********************************************;

data q2a;
input judge $ skater $ count;
cards;
E S 3
E H 0
A S 1
A H 5
;
run;

title "2 (a): Fisher's exact and Pearson chi-square test results";
proc freq data=q2a;
  weight count;
  tables judge*skater/chisq nocol;
  exact chisq/midp;
run;


***********************************************
           Question 3
***********************************************;

data q3;
input pair gts english rainsuit;
notrainsuit = 1-rainsuit;
cards;
1 1 1 1
1 0 1 0
2 1 1 0
2 0 1 0
3 1 0 0
3 0 1 0
4 1 0 0
4 0 1 0
5 1 0 0
5 0 1 0
6 1 1 0
```

```
6 0 0 1
7 1 0 0
7 0 1 0
8 1 1 0
8 0 0 1
9 1 1 0
9 0 0 0
10 1 0 0
10 0 1 0
11 1 0 0
11 0 1 1
12 1 0 0
12 0 1 1
13 1 0 0
13 0 1 0
14 1 0 0
14 0 1 1
15 1 0 0
15 0 0 1
16 1 0 0
16 0 0 1
17 1 1 0
17 0 1 1
18 1 1 1
18 0 1 1
;
run;

title '3 (a): Parameter estimates';
proc logistic data = q3 descending;
 strata pair;
 model gts = rainsuit;
run;

title '3 (a): OR';
proc logistic data = q3 descending;
 strata pair;
 model gts = notrainsuit;
run;

title '3 (b): OR';
proc logistic data = q3 descending exactonly;
 class pair/param=ref;
 model gts = pair notrainsuit;
 exact 'rainsuit' notrainsuit / estimate=parm;
run;

title '3 (c): Parameter estimates';
proc logistic data = q3 descending;
 strata pair;
 model gts = rainsuit english;
run;

title '3 (c): OR';
proc logistic data = q3 descending;
 strata pair;
 model gts = notrainsuit english;
```

```
run;

title '3 (d): OR';
proc logistic data = q3 descending exactonly;
 class pair/param=ref;
 model gts = pair notrainsuit english;
 exact 'rainsuit' notrainsuit / estimate=parm;
 exact 'English' english / estimate=parm;
run;
```