STAT 545 MIDTERM EXAM

October 5, 2015

Instructions

1. You have 75 minutes to finish the exam.
2. This exam is closed-book. However, you may use your own two-page (both sides) cheat sheet prepared before the exam.
3. Please put your name on each page.
4. Try your best to answer each question. You may get partial credits for incomplete answers that are sufficiently relevant.
5. You may or may not need the following quantities in your answers to certain problems in the exam: $Prob(X < 1.645) = .95$ and $Prob(X < 1.960) = 0.975$ assuming $X \sim N(0, 1)$.
6. You do not need to do any calculations that require a calculator. That is, you can leave all your answers in appropriate expressions/formula with all relevant numbers plugged in.

1. **[30 points]** Consider the following 2×2 table of counts:

	Y		
X	1	2	Total
1	2	3	5
2	3	8	11
Total	5	11	16

(1) Calculate the sample odds ratio, difference in proportions, and relative risk, and interpret. When needed, treat Y as the response variable in the calculations and interpretations.

(2) Calculate the 95% confidence intervals for the odds ratio, difference in proportions, and relative risk.

(You do not need to worry about the continuity correction in this problem.)

2. **[30 points]** Consider the following 2×2 table with multinomial sampling:

	Y		
X	1	2	Total
1	n_{11}	n_{12}	n_{1+}
2	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	n

(a) Give the Pearson and likelihood ratio χ^2 statistics and corresponding degrees of freedom for testing the independence between X and Y.

(b) Suppose $n_{11} = 1$, $n_{12} = 3$, $n_{21} = 3$, and $n_{22} = 1$. Let $\theta = n_{11}n_{22}/(n_{12}n_{21})$ be the odds ratio. To test H_0: $\theta = 1$ against H_1: $\theta < 1$, would you use the test statistics from (a)? Why or why not? If not, describe a test and compute the p-value. (If needed, you can treat the row and column totals as fixed.)

3. **[20 points]** For a 2×2 table with cell probabilities π_{ij}, $i, j = 1, 2$, consider testing the null hypothesis

$$H_0 : \pi_{11} = \theta^2, \pi_{12} = \pi_{21} = \theta(1 - \theta), \pi_{22} = (1 - \theta)^2,$$

where $0 \leq \theta \leq 1$ is unknown. The sample $\{n_{11}, n_{12}, n_{21}, n_{22}\}$ is assumed to follow a multinomial $(n, \{\pi_{ij}\})$ distribution with a fixed n.

(a) Show that the maximum likelihood estimate of θ under H_0 is given by $\hat{\theta} = (p_{1+} + p_{+1})/2$, where $p_{1+} = n_{1+}/n$ and $p_{+1} = n_{+1}/n$ with $n_{1+} = n_{11} + n_{12}$ and $n_{+1} = n_{11} + n_{21}$.

(b) Derive the χ^2 goodness-of-fit test for H_0. Give the degree of freedom.

4. **[20 points]** (a) Find the normalization factor, denoted as $r(\theta)$, in the discrete probability distribution $f_Y(y; \theta) = r(\theta) \exp(\theta y)/y!$, $y = 1, 2, \ldots$. (You may use the fact that for a Poisson random variable X with mean λ, $Pr(X = x) = \exp(-\lambda)\lambda^x/x!$, $x = 0, 1, \ldots$.)

(b) Show that $f_Y(y; \theta)$ has the following exponential-family form:

$$f_Y(y; \theta, \phi) = \exp\{(y\theta - b(\theta))/a(\phi) + c(y, \phi)\}$$

for some functions $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$. Find the cumulant function $b(\theta)$ and hence derive the likelihood equation for $\hat{\theta}$ based on a sample of independent and identically distributed observations $Y_1, \ldots, Y_n$.