ST744, Spring 2009

Midterm Exam, 10:15-11:30 AM, Thursday, 2/26/2009

Name: **Student ID:** **Signature:**

Honor Pledge: "By signing my name I swear that I have neither given nor received unauthorized aid in any form on this exam."

95% quantiles of χ^2 distributions

df	1	2	3	4	5	6	7	8
$\chi^2_{0.05,df}$	3.841	5.991	7.815	9.488	11.070	12.592	14.067	15.507

1. (25 pts) Suppose we have the following 2×2 contingency table for random categorical variables X and Y obtained from a multinomial sampling

	Y	
X	1	2
1	n_{11}	n_{12}
2	n_{21}	n_{22}

Denote the underlying probability structure for the joint distribution of X and Y is given by

	Y	
X	1	2
1	π_{11}	π_{12}
2	π_{21}	π_{22}

and we are interested in testing the hypothesis: $H_0 : \pi_{21} = 2\pi_{11}, \pi_{22} = 2\pi_{12}$. Do the following:

(a) (5 pts) Show that X and Y are independent under H_0.

(b) (10 pts) Find the MLE's of π_{ij}'s under H_0.

(c) (10 pts) Construct a large sample Pearson χ^2 test of H_0 at level $\alpha = 0.05$ for the following data (That is, calculate the Pearson χ^2 test statistic, compare it to an appropriate quantile from an appropriate distribution, etc.)

	Y	
X	1	2
1	24	16
2	38	25

2. (25 pts) The following table presents data on two ordinal categorical variables X and Y:

	Y		
X	1	2	3
1	12	14	2
2	2	1	1

and the exact conditional distributions of the Pearson χ^2 statistic, likelihood ratio test statistic (G^2 and the Pearson correlation coefficient r (using scores 1,2 for 2 levels of X and scores 1,2,3 for 3 levels of Y) given the margins under the independence assumption between X and Y are given in the following table:

subtable	1	2	3	4	5	6	7
χ^2	5.878	1.916	4.354	0.479	1.698	10.231	1.567
G^2	3.681	1.109	2.873	0.425	0.732	4.405	0.948
r	-0.386	-0.239	-0.092	-0.092	0.055	0.202	0.055
p	0.028	0.152	0.03	0.266	0.114	0.008	0.177
subtable	8	9	10	11	12	13	14
χ^2	1.567	8.882	23.51	5.181	3.962	10.057	23.467
G^2	0.655	2.871	8.454	3.358	2.641	4.257	8.383
r	0.202	0.349	0.496	0.202	0.349	0.496	0.643
p	0.123	0.018	0.000	0.038	0.038	0.009	0.000

where p is the table probability and Table 5 is the observed table. Do the following:

(a) (2 pts) Show the calculation of the table probability for the observed table.

(b) (5pts) Find the p-value and mid p-value of the Fisher's exact test on the null hypothesis that X and Y are independent.

(c) (5pts) Find the p-value and mid p-value of the exact Pearson χ^2 test on the null hypothesis that X and Y are independent.

(d) (5pts) Find the p-value and mid p-value of the exact likelihood ratio test G^2 on the null hypothesis that X and Y are independent.

(e) (5pts) Find the p-value and mid p-value of the exact Mantal-Haenszel test on the null hypothesis that X and Y are independent.

(f) (3pts) Which test is more appropriate?

3. (25 pts) The following data for random categorical variables X and Y is obtained from a multinomial sampling

	Y	
X	1	2
1	12	16
2	14	8

and the underlying probability structure for the joint distribution of X and Y is assumed to be given by

	Y	
X	1	2
1	π_{11}	π_{12}
2	π_{21}	π_{22}

We are interested in testing H_0: X and Y have the same marginal distribution. Do the following:

(a) (5 pts) Denote $\delta = \pi_{21} - \pi_{12}$. Show that H_0 is equivalent to $H_0 : \delta = 0$.

(b) (10 pts) Given data, find the MLE of δ and its asymptotic variance and conduct a large-sample Wald test of H_0.

(c) (10 pts) Construct a likelihood ratio test of $H_0 : \delta = 0$. What asymptotic distribution does it follow under H_0?

4. (25 pts) Assume X and Y are both ordinal categorical random variables with following probability structure

	Y	
X	1	2
1	π_{11}	π_{12}
2	π_{21}	π_{22}

where "2" is considered higher than "1".

(a) (10 pts) Show that γ is related to the odds ration through

$$\gamma = \frac{\theta - 1}{\theta + 1}$$

where θ is the true odds-ratio.

(b) (5 pts) Suppose we have data

	Y	
X	1	2
1	20	12
2	14	24

Estimate θ and construct a 95% large sample CI for θ.

(c) (10 pts) Estimate γ with the above data, and use the delta method to find the asymptotic variance of $\widehat{\gamma}$ and then construct a 95% large-sample CI for γ.