MATH 476 Statistics

Fred J. Hickernell **Spring 2011**
Makeup Test 2 **Wednesday, April 20**

Instructions:

i. *This test consists of* FOUR *questions. The scores of your* BEST THREE *attempts will count.*

ii. *The time allowed for this test is 75 minutes.*

iii. *This test is closed book, but you may use 2 double-sided letter-size sheets of notes.*

iv. *Calculators, even of the programmable variety, are allowed. Computers using JMP or MATLAB, are also allowed. No internet access, web browsing, email, chat, etc. is allowed.*

v. *Show all your work to justify your answers. Answers without adequate justification will not receive credit.*

1. (35 marks)
For each statement below, give a brief answer with explanation

a) If Y_n is a consistent estimator for θ, must Y_n be unbiased, i.e., $E(Y_n) = \theta$?

Answer: No. The estimator Y_n may be only asymptotically unbiased, i.e., $\lim_{n\to\infty} E(Y_n) = \theta$.

b) For hypothesis testing which one or more of the following are reasonable null and alternative pairs, and which are unreasonable:

$$\begin{array}{lll} H_0 : \mu = 10 & H_0 : \bar{x} = 10 & H_0 : \mu \ge 10 \\ H_a : \mu \ne 10 & H_a : \bar{x} \ne 10 & H_a : \mu \le 15 \end{array}$$

Answer: The first pair is fine. Both hypotheses concern a population quantity and both cannot be true at the same time. The second pair is unreasonable because they do not concern population quantities. The third pair is unreasonable because they both can be true at the same time.

c) A hypothesis test for $H_0 : \theta = 4$ and $H_a : \theta = 5$ has Type I error of 5% and Type II error of 20%. What is the value of the power function, power(θ), for $\theta = 4$ and $\theta = 5$.

Answer: Since power(θ) *is the probability of rejecting the null hypothesis given that θ is the true value, so*

$$\text{power}(4) = \textit{type I error} = 5\%, \qquad \text{power}(5) = 1 - \textit{type II error} = 80\%,$$

2. (35 marks)
Let $X_1, \ldots, X_n$ be i.i.d. data with a geometric distribution, i.e., the probability mass function of X_i is

$$f(x) = \theta^x(1-\theta), \quad x = 0, 1, 2, \ldots$$

a) Let $Y = X_1 + \cdots + X_n$. Show that Y is a sufficient statistic for estimating θ.

Answer: The likelihood function is

$$\begin{aligned} L(x_1, \ldots, x_n|\theta) &= f(x_1) \cdots f(x_n) = \theta^{x_1}(1-\theta) \cdots \theta^{x_n}(1-\theta) \\ &= \theta^{x_1+\cdots+x_n}(1-\theta)^n = \underbrace{\theta^y(1-\theta)^n}_{K_1(y,\theta)} \times \underbrace{1}_{K_2(x_1,\ldots,x_n)}, \end{aligned}$$

where $y = x_1 + \cdots + x_n$. The factoring of the likelihood function establishes that Y is a sufficient statistic.

b) Derive the maximum likelihood estimator for θ.

Answer: The log likelihood is

$$\ell(\theta) = \ln(L(x_1, \ldots, x_n|\theta)) = y\ln(\theta) + n\ln(1-\theta)$$

Taking the derivative with respect to θ and setting this derivative to zero implies

$$\begin{aligned} 0 = \ell'(\theta) = \frac{y}{\theta} - \frac{n}{1-\theta} &\iff \frac{y}{\theta} = \frac{n}{1-\theta} \\ &\iff \bar{x}(1-\theta) = \frac{y(1-\theta)}{n} = \theta \\ &\iff \bar{x} = (1+\bar{x})\theta \\ &\iff \theta = \frac{\bar{x}}{1+\bar{x}} \end{aligned}$$

Since

$$\ell''(\theta) = -\frac{y}{\theta^2} - \frac{n}{(1-\theta)^2} < 0,$$

it follows that $\bar{x}/(1+\bar{x})$ is the maximum likelihood estimator.

3. (35 marks)
Dot is investigating the amount of time students spend on Facebook per week. She records the usage of 15 females and 15 males.

Female	6	3.5	2.25	0	4.75	1.5	1.25	5.5	7.25	3.75	2.5	4	3.5	0	2.25
Male	5.75	1	4.5	3	2.5	3.5	2.25	4.75	1.5	3.5	4.25	7	3.75	2	4

a) The Dean of Students claims that female students spend on average no more than 2 hours per week on Facebook. Does the data support or contradict that claim?

Answer: Performing a hypothesis test of $H_0 : \mu_F \leq 2$, $H_a : \mu_F > 2$ on the female student data. yields a p-value of 2.26%. Thus, there is strong evidence to reject the null hypothesis and the claim of the Dean of Students.

b) The Provost claims that female and male students on average spend the same time on Facebook. Does the data support or contradict that claim?

Answer: Performing a hypothesis test of $H_0 : \mu_F = \mu_M$, $H_a : \mu_F \neq \mu_M$ on the data. yields a p-value of 61.34%. Thus, there is no evidence to reject the null hypothesis and the claim of the Provost.

4. (35 marks)

Ryan is investigating whether those who voted for Rahm Emmanuel for mayor or more supportive of his choice to head the Chicago Public Schools than those voted against him. He surveys a number of voters and records the following counts:

	Voted For	Voted Against
Support Choice	56	45
Do Not Support Choice	23	31

Does this data support the notion that people's votes for or against Rahm Emmanuel are related to whether they support his choice to head the schools?

Answer: After performing an analysis, the p-value is 12.7%*, so there is no strong evidence to support this notion.*