Stat 135 Fall 2013 **FINAL EXAM** **December 18, 2013**

Name: ____________________________ **SID:** ____________________________

Person on right ____________________________ Person on left ____________________________

- There will be **one, double sided, handwritten**, 8.5in x 11in page of notes allowed during the exam.
- The exam is **closed book** and will be **2 hours and 50 minutes** long.
- You are allowed the use of a non-programmable calculator, that cannot communicate with any other device.
- Relevant tables are provided.
- The test begins on the next page. Please do not write below this line.
- ALWAYS SHOW YOUR REASONING. No work, no credit.

Page	Maximum score	Score achieved
2	10	
3	10	
4	10	
5	10	
6	15	
7	10	
8	15	
9	10	
Total	**90**	

Name: ____________________________

1. One year, there were about 3,000 institutions of higher learning in the US, and as part of a study, the Carnegie Commission took a simple random sample of 400 of these. The average enrollment in the 400 sample schools was 3,700; and the sample standard deviation was 6,500.

   (a) True or false: An approximate 68% confidence interval for the average enrollment of all 3,000 schools runs from about 3,400 to 4,000. (2 points)

   (b) True or false: If a statistician takes a simple random sample of 400 institutions out of 3,000, and goes 1 SE in either direction from the sample average enrollment, there is about a 68% chance that the resultant interval will contain the average enrollment of all 3,000 schools. (Here SE indicates the estimated standard error of the sample mean). (2 points)

   (c) True or false: It is estimated that about 68% of the 3,000 schools enrolled between 3,000 and 4,000 students. (2 points)

   (d) Sketch (roughly) the following distributions, and approximately mark the mean and SD of the distribution. (2 points each)

      i. The sampling distribution of the average enrollment of the 400 sample schools.

      ii. The enrollments of all 3,000 schools.

Name: ____________________________

2. Lost Springs is a town in Converse County, WY. As of the 2010 census, the population of Lost Springs was 4. The ages of the 4 residents are {50,54,65,80}. We want to take a sample of size 2 from this population. Suppose, instead of a simple random sample, the census official decides to sample from the following pairs, with each pair being equally likely: {50,54},{54,80},{65,80},{54,65}

   (a) Write down the distribution of the sample mean. (3 points)

   (b) Is the sample mean unbiased? (2 points)

3. Two large populations are independently surveyed using simple random samples of size n, and two proportions are estimated: $\hat{p}_1 = 0.15$, and $\hat{p}_2 = 0.12$. How large should n be so that the standard error of the difference will be less than 0.05? (3 points)

4. According to a recent Gallup poll, 51% of Americans disapprove of the Affordable Care Act ("Obamacare"). The poll also states that the results for this Gallup poll are based on telephone interviews conducted Dec. 11-12, 2013, on the Gallup Daily tracking survey, with a random sample of 1,011 adults, aged 18 and older, living in all 50 U.S. states and the District of Columbia, and that the margin of sampling error is ±3 percentage points at the 95% confidence level. How is this margin of error (±3) arrived at? (2 points)

Name: ____________________________

5. Statisticians A and B obtain independent samples $X_1, X_2, \ldots X_{10}$ and $Y_1, Y_2, \ldots, Y_{17}$ respectively, both from a $N(\mu, \sigma^2)$ distribution, with both μ and σ^2 unknown. They estimate (μ, σ^2) by $(\bar{X}, S_X^2)$ and $(\bar{Y}, S_Y^2)$ respectively. Given that $\bar{X} = 5.5$ and $\bar{Y} = 5.8$ respectively, which statistician's estimate of σ^2 is more probable to have exceeded the true value by more than 50%? Explain! (5 points)

6. A random sample of 59 people from the planet Krypton yielded the following results:

		Eye Color	
		Blue	Brown
Gender	Male	19	10
	Female	9	21

Professor A had believed that sex and eye-color are independent factors on Krypton. After doing a χ^2 test of his hypothesis against the alternative $H_1 : \sum_i \sum_j \pi_{ij} = 1$, he finds that he has to reject his null hypothesis at the 5% significance level.

Professor B has always believed the much stronger hypothesis that $\pi_{ij} = 1/4$ for all i and j. After doing a χ^2 test of his hypothesis against H_1, he finds that he does not need to reject it at the 5% level. Please check their computations. They both are doing χ^2 tests, so why do they get different results? (5 points)

Name: ____________________________

7. Suppose that X is a discrete random variable with:

$$\begin{aligned} P(X=0) &= \frac{2}{3}\theta \\ P(X=1) &= \frac{1}{3}\theta \\ P(X=2) &= \frac{2}{3}(1-\theta) \\ P(X=3) &= \frac{1}{3}(1-\theta) \end{aligned}$$

where $0 \leq \theta \leq 1$ is a parameter. We have 10 independent observations from this distribution:$(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$.

(a) Find the method of moments estimate of θ. (5 points)

(b) Find an approximate standard error for your estimate. (5 points)

(c) Find the maximum likelihood estimate for θ. (5 points)

(d) What is an approximate standard error of the maximum likelihood estimate? (10 points)

Name: ____________________________

8. Let $Y_i = \beta_0 + \beta_1 x_i + e_i,\ i = 1, 2, \ldots, n$ be the simple linear model for the response Y, given the predictor x. Let $\hat{\beta}_0$ and $\hat{\beta}_1$ be the least squares estimates for β_0 and β_1 respectively. Show that the residuals $\hat{e}_i = Y_i - \hat{Y}_i$ sum to 0. (The $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ are the fitted values.) (5 points)

9. At a large engineering school, students are required to take aptitude tests in various subjects, including mathematics and physics. Students who do well on the mathematics test also tend to score high on the physics tests. On both tests, the average score is 60, and the spreads are about the same for both tests. The data are homoscedastic, and the distributions of the scores for both tests are approximately normal. Of the students scoring about 75 on the mathematics test:

   (a) just about half scored over 75 on the physics test.

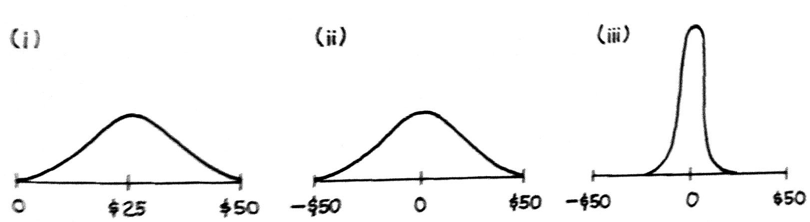
   (b) more than half scored over 75 on the physics test.

   (c) les than half scored over 75 on the physics test.

   Choose one option, and explain. (2 points)

10. Several different regression lines are used to predict the price of a stock (from different independent variables). Histograms for the residuals from each line are sketched below. Match the description with the histogram, and explain your choices. (3 points)

    (a) $s = \$5$ (b) $s = \$15$ (c) something's wrong



Name: ____________________________

11. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed Bernoulli(p) random variables. That is, for $x = 0$, 1

$$f(x;p) = p^x(1-p)^{1-x}$$

and $f(x;p)$ is 0 elsewhere.

(a) Find the best critical region (corresponding to the best test) for testing $H_0 : p = \frac{1}{2}$ against $H_1 : p = \frac{1}{3}$ is of the form $\left\{\sum X_i \leq c\right\}$. (7 points)

(b) Use the Central Limit Theorem to find n and c so that the significance level is approximately 0.1, and the power is approximately 0.8. (8 points)

Name: ____________________________

12. (a) The following numbers give the percent extension under a given load for two random samples of lengths of yarn, the first sample being taken before washing, and the second after six washings.

Before washing:	12.3	13.7	10.4	11.4	14.9	12.6	
After 6 washings:	15.7	10.3	12.6	14.5	12.6	13.8	11.9

Do they provide any justification for concluding that extensibility is affected by washing? (You may assume that the standard deviation is unaltered by the washings.) (5 points)

(b) In another experiment with the same type of yarn, six lengths of yarn were selected at random and each length cut into two halves. One of the halves was tested for extension without washing, and the other after six washings, giving the following percent extensions:

Length:	A	B	C	D	E	F
Before washing:	13.9	12.5	11.0	11.8	10.8	14.6
After 6 washings:	14.7	12.1	13.2	13.6	11.5	15.4

What evidence do these provide as to the effect of washing on extensibility? Why should this experiment be analysed differently from the previous one? (5 points)