Name (Print):

STAT509 Fall 2014 Practice Final November 21, 2014

Instructions The open book, open notes final exam will be held on 12/09/14 at 7:30pm. The purpose of this practice exam is to help you understand and practice some of the important topics in Topic 1, 2, 3 and 7. For the other three topics, please see practice exam for midterm 2 as a review. The final exam topics include everything covered in chapter 1 through chapter 7. Your homework and the lecture notes are always your best references for your exam. Good luck!

- 1. In **Topic 1**, understand the following terms: sample space, event, union, intersection, mutually exclusion (disjoint), conditional probability, independence, probability rule, fundamental theorem of counting, permutation and combination. Complete questions 2-8.
- 2. Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

		Shock resistance		
		high	low	
scratch	high	70	9	
resistance	low	16	5	

- (a) If a disk is selected at random, what is the probability that its scratch resistance is high and its shock resistance is high?
- (b) If a disk is selected at random, what is the probability that its scratch resistance is high or its shock resistance is high?
- (c) Consider the event that a disk has high scratch resistance and the event that a disk has high shock resistance. Are these two event mutually disjoint?
- (d) Given a disk has high shock resistance, what is the probability that its scratch resistance is also high?
- 3. Suppose 2% of cotton fabric rolls and 3% of nylon fabric rolls contain flaws. Of the rolls used by a manufacturer, 70% are cotton and 30% are nylon. What is the probability that a randomly selected roll used by the manufacturer contains flaw?

4. Samples of emissions from three suppliers are classified for conformance to air-quality specifications. The results from 100 samples are summarized as follows:

		conforms	
		yes	no
supplier	1	22	8
	2	25	5
	3	30	10

Let A denote the event that a sample is from supplier 1, and let B denote the event that a sample conforms to specifications.

- (a) Are events A and B independent?
- (b) Determine P(B|A).
- 5. How many possible license plates could be stamped if each license plate were required to have exactly 3 letters and 3 numbers?



- 6. With 6 names in a bag, randomly select a name. How many ways can the 6 names be assigned to 6 job assignments?
 - (a) If we assume that each person can only be assigned to one job, then we must select (or sample) the names without replacement. That is, once we select a name, it is set aside and not returned to the bag.
 - (b) What if the 6 names were sampled with replacement? That is, once we select a name, it is returned to the bag.

- 7. An artist has 9 paintings. How many ways can he hang 4 paintings side-by-side on a gallery wall?
- 8. Maria has three tickets for a concert. She'd like to use one of the tickets herself. She could then offer the other two tickets to any of four friends (Ann, Beth, Chris, Dave). How many ways can 2 people be selected from 4 to go to a concert?
- 9. In **Topic 2**, understand the following terms: discrete random variables, pmf, cdf, expectation and its properties, variance and its properties, standard deviation, Bernoulli trial, Binomial distribution, Geometric distribution, Hypergeometric distribution, Poisson distribution. Complete questions 10 14.
- 10. We flip the coin three times. Let X be the number of heads obtained.
 - (a) Complete the distribution table below:

X	0	1	2	3
p(x)				

- (b) What is the probability that we could obtain less than 3 and more than 0 heads?
- (c) What is the mean and variance for X, the number of heads?
- (d) What is the mean and variance for Y, the number of tails? Hint: Y is a function of X.
- 11. Suppose now we flip the fair coin 100 times, let X denote the number of heads we obtained.(a) What is the probability that we could see 60 heads?

(b) What is the mean and variance of X?

12. We still look at the flipping fair coin example.

- (a) What is the probability that the first head appears on the third flip?
- (b) Let X denote the number of flips until we see the first head, what is the mean of X?
- 13. A wallet contains 3 \$100 bills and 5 \$1 bills. You randomly choose 4 bills. What is the probability that you will choose exactly 2 \$100 bills?
- 14. If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.
- 15. In **Topic 3**, understand the following terms: continuous random variables, pdf, cdf, expectation, variance, standard deviation, Exponential distribution, Poisson-exponential relationship, memoryless property, constant failure rate mode, Weibull distribution, normal distribution, standardizing, z-table, inverse normal probabilities. Complete questions 16-18.
- 16. Suppose that a random variable X has the pdf $f(x) = e^{-(x-4)}$ for x > 4. Answer the following question:
 - (a) Determine the CDF of X.
 - (b) Find the mean and variance of X.

- (c) Find P(2 < X < 8).
- 17. The compressive strength of samples of cement can be modeled by a normal distribution with a mean of 6000 kilograms per square centimeter and a standard deviation of 100 kilograms per square centimeter.
 - (a) What is the probability that a sample's strength is less than 6250 Kg/cm^2 ?
 - (b) What is the probability that a sample's strength is between 5800 and 5900 Kg/cm²?
 - (c) What strength is exceeded by 95% of the samples?
- 18. The number of stork sightings on a route in South Carolina follows a Poisson process with a mean of 2.3 per year.
 - (a) What is the mean time between sightings?
 - (b) What is the probability that there are no sightings within next three months?
 - (c) What is the probability that the time until the first sighting exceeds six months?
 - (d) What is the probability of no sightings within three years?

- 19. For the review questions in Topic 4-6, please see practice midterm 2 and the notes.
- 20. In **Topic 7**, understand the following terms: regression, intercept, slope and their interpretations, inference on β_1 (t-test and ANOVA), interpretation of confidence and prediction interval at $x = x_0$, relationship between correlation and coefficient of determination, model diagnosis. Complete question 21.
- 21. The weight and systolic blood pressure of 26 randomly selected males in the age group 25 to 30 are given in the following R code. Answer the following questions.

weight<-c(165,137,180,155,212,175,190,210,200,149,159,169,170,172,159,168,174,183,215,195,180,143,240,235,192,187) BP<-c(130,133,150,128,151,146,150,140,148,125,133,135,150,153,128,132,149,158,150,163,156,124,170,165,160,159)

- (a) Find a regression line relating systolic blood pressure (BP) to weight in R. State your prediction model clearly.
- (b) Interpret the value of $\hat{\beta}_1$. Make sure you use the context of the question.
- (c) Construct a 95% CI for β_1 . The point estimate and standard error can be found via R.
- (d) Suppose we want to test the significance of β_1 , that is $H_0: \beta_1 = 0$ vs. $H_a: \beta_1 \neq 0$. Report the *p*-value from the t-test and make a conclusion and decision. Is there another approach that can do this?

Source	SS	df	MS	F
Regression	10234.7	⊕	2	3
Error	4	5	6	
Total	17057.6	25		

(e) Without doing ANOVA in R, complete the following table:

- $(1) = \underline{\qquad}; (2) = \underline{\qquad}; (3) = \underline{\qquad}; (4) = \underline{\qquad}; (5) = \underline{\qquad}; (6) = \underline{\qquad}; (6)$
- (f) Suppose we have a new male whose blood pressure is 156, but with unknown weight. What is the mean weight for this male?

(g) For the previous question, find a 95% confidence interval and prediction interval in R. Interpret both intervals.

- (h) Suppose the coefficient of determination is 0.6, how do we interpret this value? What is the correlation?
- (i) Construct a QQ plot for the residuals. Is the normality assumption satisfied?
- (j) Plot residuals vs. the fitted values. Are the assumptions $E(\epsilon) = 0$ and constant variance satisfied? If yes, explained why; if no, how do you suggest to correct these?