TEST #2
STA 5326 **Name:**____________________________________
November 10, 2016

Please read the following directions.
DO NOT TURN THE PAGE UNTIL INSTRUCTED TO DO SO

Directions

- The exam is closed book and closed notes. You will be supplied with scratch paper, and a copy of the Table of Common Distributions from the back of our textbook.
- During the exam, you may use ONLY what you need to write with (pens, pencils, erasers, etc) and (if you wish) an ordinary scientific calculator (TI-86 or below is fine).
- Partial credit is available. (If you know part of a solution, write it down. If you know an approach to a problem, but cannot carry it out – write down this approach. If you know a useful result, write it down.)
- For all problems (except the multiple choice on the last page), you must show and explain your work (including your calculations). **No credit is given without work.** But don't get carried away! Give enough explanation and work so that what you have done is clearly understandable.
- Make sure that the grader can easily see how you get from one step to the next. If you needed scratch paper to work something out, make sure to transfer your work to the exam.
- You should give only one answer to each problem. **Circle your answer** if there is any chance for confusion.
- All the work on the exam should be your own. No "cooperation" is allowed.
- Numerical (decimal) answers should be given for problems requiring normal approximations.
- Simplify your answers when it is easy to do so. But, except for normal approximations (as noted above), more difficult arithmetic does **not** have to be done completely. Answers can be left as fractions or products. You do not have to evaluate large binomial coefficients, factorials or powers. Answers can be left as summations (unless there is a simple closed form such as when summing a geometric or exponential series).
- Do **not** quote homework results. If you wish to use a result from homework in a solution, you must prove this result.
- The exam has **9** problems and **11** pages. There are a total of **100** points.

Problem 1. (12%) For a continuous random variable, the median m satisfies

$$P(X \le m) = P(X \ge m) = 1/2\,.$$

Find the median of the distribution with density given by (with $\beta > 0$):

$$f(x) = \begin{cases} \dfrac{\beta}{x^{\beta+1}} & \text{for } x \ge 1 \\ 0 & \text{for } x < 1 \end{cases}$$

This is similar to exercise 2.17.

Problem 2. (12%) Three gamblers eat lunch together every day. They decide which one of them will pay for lunch by playing the following game. They each toss a fair coin. If exactly one of the gamblers gets a tail, then that gambler pays for lunch. But if more than one of the gamblers gets a tail or nobody gets a tail, then they play another round (each tossing a coin, etc.). They keep playing rounds until exactly one of them gets a tail (and pays for lunch). What is the expected number of rounds they play?

This is somewhat similar to Exercise 3.4(a). In that problem, sampling keys with replacement leads to a Geometric distribution because it gives rise to iid Bernoulli trials. Any situation where we are performing iid Bernoulli(p) trials and stopping at the first success will result in performing a Geometric(p) number of trials with a mean of $1/p$ *trials. In this problem, the rounds of the game are iid Bernoulli(*$p = 3/8$*) trials where* $p = 3/8$ *is the probability of exactly one tail when three coins are tossed. The number of rounds has a Geometric(*$p = 3/8$*) distribution with a mean of* $1/p = 8/3$*. So the answer is 8/3.*

Problem 3. A street has n street lights. Each light has an independent exponential lifetime with a mean of β years. Formerly, lights were replaced as soon as they failed. But now, due to Trump's victory, someone will be sent out to replace the lights only when **4** of them have burned out (where $4 < n$).

(a) (10%) Suppose the rate of violent crime on this street increases by an average of c crimes per year for every light that burns out. (In other words, when j lights have failed, the crime rate is jc crimes per year higher than when all the lights are working.) What is the expected number of additional crimes which will occur on this street because of Trump's victory?

This is similar to exercise C4. It relies on the discussion in the lecture notes in pages 9–12 of notes7.pdf

Let $T_1 < T_2 < T_3 < T_4$ denote the times at which the first four lights burn out. The failed lights are replaced at time T_4. Let U_i denote the length of time during which i lights are out. Then we have $U_0 = T_1$ and $U_i = T_{i+1} - T_i$ for $i = 1, 2, 3$ since all the lights are working up to time T_1 and exactly i lights are out between times T_i and T_{i+1}. According to the discussion in lecture given in pages 9–12 of notes7.pdf (where different variable names are used), the random variables U_i are independent exponential rv's with means $EU_i = \beta/(n-i)$ for $i = 0, 1, 2, 3$. If we knew the length of time U_i in which i lights are out, we would expect (on average) an additional icU_i crimes to occur during this time. Thus, during the time periods U_1, U_2, U_3 we expect an additional $c(U_1 + 2U_2 + 3U_3)$ crimes to occur on average. Taking the expected value of this expression, the expected number of additional crimes is $c\beta(1/(n-1) + 2/(n-2) + 3/(n-3))$.

[**Problem 3 continued**]

(b) (10%) Suppose all the lights are working today. Let X denote the length of time until someone is next sent to replace failed lights (that is, the time until 4 of the n lights have failed). What is the moment generating function (mgf) of X?

$X = T_4 = U_0 + U_1 + U_2 + U_3$ *can be expressed as the sum of four independent exponential rv's with means* β/n, $\beta/(n-1)$, $\beta/(n-2)$, $\beta/(n-3)$. *So the mgf of* X *is the product of the mgf's of these four exponential rv's. The mgf of an exponential rv with mean* β *is* $1/(1-\beta t)$ *for* $t < 1/\beta$. *So the answer is*

$$\frac{1}{(1-\frac{\beta}{n}t)(1-\frac{\beta}{n-1}t)(1-\frac{\beta}{n-2}t)(1-\frac{\beta}{n-3}t)} \quad \textit{for } t < \frac{n-3}{\beta}.$$

Deduct one point if students neglect to state where the mgf is finite.

Problem 4. (12%) Suppose the random variable X has mgf

$$M_X(t) = \frac{p}{e^t - (1-p)e^{-t}} \quad \text{for } t > \frac{1}{2}\log(1-p).$$

Find the mean and variance of X.

This is very similar to the example on page 9 of notes6.pdf in which we find the mean and variance of the geometric distribution from the mgf.

The answers are $EX = \frac{p-2}{p}$ *and* $Var(X) = \frac{4(1-p)}{p^2}$.

Another way to do the problem is as follows: Let $Y \sim Geometric(p)$. *We know the mgf and mean and variance of* Y *from the appendix. Using the scaling properties of mgf's, we can show that* $-2Y+1$ *has exactly the mgf of* X *given above. Therefore* X *has the same distribution as* $-2Y+1$. *Thus* $EX = -2EY + 1 = -2/p + 1$ *and* $Var(X) = 4\,Var(Y) = 4(1-p)/p^2$. *This approach is tricksy, but you might suspect to look for something like this because of the similarity between the mgf of* X *and the Geometric mgf.*

Extra Work Space (if needed)

[Problem 4 continued]

Problem 5. (12%) Let A, B, C be arbitrary events. For any event D, let I_D be the corresponding indicator random variable. Simplify the following expression:

$$E(I_A - I_{B^c} + I_{A^c}\, I_{B^c} - I_A\, I_B\, I_{C^c})\,.$$

Write your answer as some simple expression involving one or more of the probabilities $P(A)$, $P(B)$, $P(C)$, $P(A \cap B)$, $P(A \cap C)$, $P(B \cap C)$, $P(A \cap B \cap C)$.

This problem uses the same sort of manipulations as exercise C0. See pages 9–12 of notes4.pdf. The correct answer is $P(A \cap B \cap C)$. The answer $P(A)P(B)P(C)$ is **wrong** *since we are told that A, B, C are arbitrary events; they are* **not** *known to be independent.*

Problem 6. (12%)

(a) Suppose $X \sim \text{Gamma}(\alpha = 100, \beta = 2)$. Find a good approximation to:

$$P(200 < X < 201)$$

Part (a) is worth 8 points and part (b) worth only 4 points.

A normal approximation produces an accurate approximation in this case. We know the gamma distribution is approximately normal for large values of α. The Gamma distribution is continuous, so we do NOT use the continuity correction.

It is a serious error to use the continuity correction here, so students should lose several points if they do this.

(b) Suppose $X \sim \text{Gamma}(\alpha = 1, \beta = 200)$. Find a good approximation to:

$$P(200 < X < 201)$$

When $\alpha = 1$, the Gamma distribution is an exponential distribution, which is NOT approximately normal regardless of the value of β, which is a scale parameter and does not affect the shape of the distribution. The normal approximation should NOT be used here. Anyone using the normal approximation in this part should receive no credit.

We actually know a simple formula for the cdf of the exponential distribution given by $F(x) = 1 - e^{-x/\beta}$, $x > 0$. The exact answer is then given by $F(201) - F(200)$. Anyone giving this answer should receive full credit (even though it is not an approximation).

For a continuous distribution with a continuous density f, the probability of a small interval of length dx containing the point x is approximately $f(x)\,dx$. The interval from 200 to 201 is quite small relative to $\beta = 200$, so the simple approximation $f(x)\,dx$ (setting $x = 200$ and $dx = 1$ and $f(x) = (1/\beta)e^{-x/\beta}$) works very well here. Anyone who correctly carries out this approach should receive full credit.

Problem 7. (12%) Find the moment generating function for a random variable with density

$$f(x) = \begin{cases} -\dfrac{2x}{\beta^2} & \text{for } -\beta < x < 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\beta > 0$.

This is very similar to exercise 2.30(b). The answer is:

$$\frac{-2}{t^2\beta^2}(\beta t e^{-\beta t} + e^{-tb} - 1)$$

for $t \neq 0$*, and* $M(0) = 1$.

No work is required for the problems on this page.

Problem 8. (4%) Let $X_1, X_2, X_3, \ldots$ be iid random variables having a Cauchy distribution with pdf

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, the sample mean of the first n random variables. Circle the single choice from the list below which correctly completes the follow statement:
As $n \to \infty$, the sequence of sample means $\bar{X}_n$, $n = 1, 2, 3, \ldots$ converges __________________.

See page 6 of notes4.pdf. Since the mean of the Cauchy does not exist and the variance is not finite, the LLN and CLT fail for the Cauchy distribution. This rules out most of the answers below. But it was mentioned on page 6 of notes4.pdf that $\bar{X}_n$ has the same distribution as X_1 for all n. Thus $\bar{X}_n$ converges in distribution to X_1, which is one of the choices below.

a) in distribution to Exp(1)

b)$\star$ in distribution to X_1

c) in distribution to $N(0,1)$

d) in distribution to $N(0,1/n)$

e) in distribution to $N(0,n)$

f) to 0 with probability 1

g) to 1 with probability 1

h) to 1/n with probability 1

i) to ∞ with probability 1

j) to n with probability 1

k) to $N(0,1)$ with probability 1

Problem 9. (4%) Which ONE of the plots given below could be the plot of a moment generating function? Circle the single correct choice from the list below.

a) **b**) **c**) **d**) **e**) **f**)$\star$

See page 6 of notes5.pdf. An mgf M must satisfy $M(0) = 1$ and $M(t) \geq 0$ for all t, and be convex and infinitely differentiable. The only plot that appears to satisfy all of these is plot F.