Exam II

STAT 430
FALL 2017

Instructions

- The exam is a take home exam. It is due November 18th on blackboard by 5:00 pm.
- If you have any questions about, or need clarification on the meaning of an item on this exam, please ask your instructor during my office hours or email.
- No other form of external help is permitted attempting to receive help or provide help to others will be considered cheating.
- **Do not cheat on this exam.** Academic integrity demands an honest and fair testing environment. Cheating will not be tolerated and will result in an immediate score of 0 on the exam and an incident report will be submitted to the dean's office.

Name: ______________________________

Student ID: ______________________________

1. (5 points) Suppose that U is a uniform random variable on the interval $[0, 1]$. Also suppose that F is any invertable cumulative density function. Find the cdf of $X = F^{-1}(U)$.

2. (5 points) Let X be any random variable with mean μ and variance σ^2. What is the mean and variance of the random variable $Z = \frac{X-\mu}{\sigma}$?

3. (5 points) For any random variables X and Y show that $Var(X-Y) = Var(X)+Var(Y)-2Cov(X, Y)$.

4. (5 points) For any random variables X and Y, find that $Var(X \cdot Y)$ in terms of $E(X), E(Y), Var(X), Var(Y)$, and $Cov(X, Y)$.

5. Suppose that $X_1, X_2, ..., X_n$ are independent random variables with the same expected value μ and the same variance σ^2.
   (a) (5 points) What is the expected value and variance of $\bar{X} = \frac{1}{n}(X_1 + X_2 + ... + X_n)$?
   (b) (5 points) What is the expected value and variance of $T = X_1 + 2X_2 + ... + nX_n$?

6. Let X_1 and X_2 are independent, identically distributed random variables with the same probability density function
$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$
(i.e., X_1 and X_2 are gamma random variables).
   (a) (5 points) Find the probability density function for $Y_1 = 1/X_1$ (be sure to include the range).
   (b) (5 points) Find the probability density function for $Y_2 = X_1 + X_2$ (be sure to include the range).
   (c) (5 points) What is the expected value of $Z = Y_1 \cdot Y_2$?

7. **Stick Breaking** Consider the following scenario: I hold a wooden stick of length L in my left hand and randomly select a point at which to break the stick. I name the portion still in my left hand Stick A and place it on a table. I then pick up what was left of the stick and, holding it in my left hand, break it again by selecting another point at random. I name the piece in my left hand Stick B and place it on the table with Stick A.
   (a) (5 points) What is the probability that Stick A is longer than Stick B?
   (b) (5 points) What is the probability that Stick B is longer than the length of the stick left over?
   (c) (5 points) What is the probability that the sum of lengths of Stick A and Stick B is less than $\frac{1}{2}L$?
   (d) (5 points) Which length has the highest variance? The length of Stick A or the length of Stick B?

8. Let X be a $N(\mu, \sigma^2)$.
   (a) (5 points) Derive the moment generating function for X.
   (b) (5 points) Using the moment generating function for X, find the moment generating function for $Y = X - \mu$.
   (c) (5 points) Show that $E(Y^k) = E\left((X-\mu)^k\right)$ is 0 if k is odd.
   (d) (5 points) Find an expression for $E(Y^k) = E\left((X-\mu)^k\right)$ if k is even in terms of , μ, and k.

9. (30 points) **Gibbs Sampling** One of the sampling techniques commonly used in Bayesian Analysis is the Gibbs sampler. Essentially, the gibbs sampler is a method of generating multiple observations $x_1, x_2, \ldots, x_n$ from random variables $X_1, X_2, \ldots, X_n$ with joint distribution $f(x_1, x_2, \ldots, x_n)$. The alogorithm works like this:

   1. With the ith observation $x_1, x_2, \ldots, x_n$, generate the $i+1$th observation $x_1^*, x_2^*, \ldots, x_n^*$ as follows:
      i. Generate x_1^* from the conditional distribution $f(x|X_2 = x_2, X_3 = x_3, \ldots, X_n = x_n)$
      ii. Generate x_2^* from the conditional distribution $f(x|X_1 = x_1^*, X_3 = x_3, \ldots, X_n = x_n)$
      iii. Generate x_3^* from the conditional distribution $f(x|X_1 = x_1^*, X_2 = x_2^*, \ldots, X_n = x_n)$
      ...
      n. Generate x_n^* from the conditional distribution $f(x|X_1 = x_1^*, X_2 = x_2^*, \ldots, X_{n-1} = x_n^*)$

   In this way, each iteration generates values $x_1, \ldots, x_n$ following the joint distribution of $X_1, \ldots, X_n$. As an added bonus, if we look only at the values generated for x_1, the values will follow the marginal for X.

   Suppose that $\theta \sim Beta(3, 1)$, $X|\theta, Y = y \sim \text{binom}(y, \theta)$, and $Y|\theta \sim \text{binom}(10, \theta)$.

   Write a function in R that generates samples from the joint distribution of X, Y and θ using a Gibbs Sampling algorithm. Use it to get 10,000 samples distribution after letting the sampler run 10,000 times to avoid any problems from your initial starting point. Submit your function, your initial starting value, plots of the 10,000 values for each of the variables. Provide summary statistics including the mean and variance of your samples.