Name: ______________________

Signature: ______________________

Math 5651. Lecture 001 (V. Reiner) Midterm Exam III
Tuesday, November 16, 2010

This is a 115 minute exam. No books, notes, calculators, cell phones or other electronic devices are allowed. There are a total of 100 points. To get full credit for a problem you must show the details of your work. Answers unsupported by an argument will get little credit. Do all of your calculations on this test paper.

Problem	Score
1.	______
2.	______
3.	______
4.	______
5.	______
Total:	______

Problem 1. *(20 points total)* Let X be a continuous random variable with probability density function $f(x) = 2x$ for $0 \le x \le 1$, and $f(x) = 0$ for all other x.

a. (8 points) Compute the expectation EX and variance $\mathrm{Var}(X)$.

b. (6 points) Let $Y = -3X + 7$. Compute the expectation EY and variance $\mathrm{Var}(Y)$. You must justify your answer.

c. (6 points) Compute the correlation $\rho(X, Y)$. You must justify your answer.

Problem 2. *(20 points total)*

a. (5 points) Let X be a discrete random variable taking on two possible values $+1, -1$, with probability function $f(+1) = \frac{3}{4}, f(-1) = \frac{1}{4}$

Compute the moment generating function $\psi_X(t)$ as a function of t, with no summations in your final answer.

b. (10 points) Consider a particle that begins at 0 on the real number line, and in each second, either takes a step one unit to the right with probability $\frac{3}{4}$, or a step one unit to the left with probability $\frac{1}{4}$.

Let Y be the random variable which is the x-coordinate of this particle after n seconds.

Compute the moment generating function $\psi_Y(t)$ as a function of t, again with no summations in your final answer.

c. (5 points) Compute the second moment $\mu_2(Y) = E(Y^2)$, as a function of n, simplified as much as possible.

Problem 3. *(20 points)* Let X, Y be *independent* random variables for which $EX, EY, \mathrm{Var}(X), \mathrm{Var}(Y)$ all exist.

Assuming $EX = EY$, prove that $E((X-Y)^2) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Problem 4. *(20 points total; 5 points each)* Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, with expectation $EX_i = 10$ and variance $\mathrm{Var}(X_i) = 3$.

a. (5 points) Compute the variance $\mathrm{Var}(3X_1 - 7X_2 + 12)$.

b. (5 points) Let $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ denote the sample mean of $X_1, \ldots, X_n$. What is its expectation $E(\overline{X}_n)$?

c. (5 points) What is its variance $\mathrm{Var}(\overline{X}_n)$?

d. (5 points) How large must n be, at a minimum, before Chebyshev's inequality implies that $\Pr(9 \leq \overline{X}_n \leq 11) > 0.9$?

Problem 5. *(20 points total)* Let (X, Y) be random variables defined by the following process. First X is chosen with uniform distribution on $[0, 3]$, and then *after* the value x for X has been chosen, Y is chosen uniformly on the interval $[0, x]$.

a. (7 points) Write down the marginal probability density functions $f_1(x)$ and $f_2(y)$ for each value of x and y.

b. (6 points) Compute the conditional expectation $E(Y|X = x)$ as a function of x.

c. (7 points) If one knows that $X = \frac{1}{2}$, what prediction M for the value of Y will minimize the expectation $E(|Y - M|)$ of the absolute value of the error $Y - M$? Explain.

Brief solutions

1.(a)

$$EX = \int_{x=0}^{x=1} x \cdot 2x dx = \frac{2}{3}.$$

$$\mathrm{Var}(x) = E(X^2) - (EX)^2 = \int_{x=0}^{x=1} x^2 \cdot 2x dx - \frac{4}{9} = \frac{1}{18}.$$

(b)

$$E(Y) = E(-3X+7) = -3EX + 7 = 5.$$

$$\mathrm{Var}(Y) = \mathrm{Var}(-3X+7) = (-3)^2\mathrm{Var}(X) = \frac{1}{2}.$$

(c)

$$\rho(X,Y) = \rho(X,-3X+7) = \frac{\mathrm{Cov}(X,-3X+7)}{\sigma(X)\sigma(-3X+7)} = \frac{-3\mathrm{Cov}(X,-X)}{\sigma(X)\cdot 3\sigma(X)} = -1$$

2.(a)

$$\Psi_X(t) = E(e^{tX}) = \frac{3}{4}e^{t(+1)} + \frac{1}{4}e^{t(-1)} = \frac{1}{4}\left(3e^t + e^{-t}\right).$$

(b) Note that $Y = X_1 + X_2 + \cdots + X_n$ where X_i are i.i.d. random variables having the same distribution as X above. Hence

$$\Psi_Y(t) = \Psi_X(t)^n = \frac{(3e^t + e^{-t})^n}{4^n}$$

(c) Since $\mu_2(Y) = \Psi''_Y(t=0)$ we need to compute

$$\Psi'(t) = \frac{1}{4^n}n\left(3e^t + e^{-t}\right)^{n-1}\left(3e^t - e^{-t}\right)$$

and

$$\Psi''(t) = \frac{n}{4^n}\left[(n-1)\left(3e^t + e^{-t}\right)^{n-2}\left(3e^t - e^{-t}\right)^2 + \left(3e^t + e^{-t}\right)^{n-1}\left(3e^t + e^{-t}\right)\right]$$

and finally

$$\mu_2(Y) = \Psi''_Y(t=0) = \frac{n}{4^n}\left[(n-1)4^{n-2}\cdot 4 + 4^n\right] = \frac{n(n+3)}{4}$$

3. One has

$$\begin{aligned} E((X-Y)^2) &= E(X^2 - 2XY + Y^2) \\ &= E(X^2) - 2E(XY) + E(Y^2) && \text{by linearity of expectation} \\ &= E(X^2) - 2EX \cdot EY + E(Y^2) && \text{due to independence of } X, Y \\ &= E(X^2) - (EX)^2 + E(Y^2) - (EY)^2 && \text{since } EX = EY \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

4.(a)

$\text{Var}(3X_1 - 7X_2 + 12) = \text{Var}(3X_1) + \text{Var}(7X_2) = 9\text{Var}(X_1) + 49\text{Var}(X_2) = 174$

where the first equality used the independence of X_1, X_2.

(b)

$$E(\overline{X}_n) = E(\frac{1}{n}(X_1 + \cdots + X_n) = \frac{1}{n}(E(X_1) + \cdots + E(X_n)).$$

(c)

$$\begin{aligned} \text{Var}(\overline{X}_n) &= \text{Var}(\frac{1}{n}(X_1 + \cdots + X_n) \\ &= \frac{1}{n^2}(\text{Var}(X_1) + \cdots + \text{Var}(X_n)) \\ &= \frac{1}{n^2}(3n) = \frac{3}{n}. \end{aligned}$$

(d) We want

$$\Pr(|\overline{X}_n - 10| \le 1) > 0.9$$

or equivalently

$$\Pr(|\overline{X}_n - 10| > 1) < 0.1$$

Chebyshev's inequality says

$$\Pr(|\overline{X}_n - E\overline{X}_n| > 1) \le \frac{\text{Var}\overline{X}_n}{1^2} = \frac{3/n}{1^2} = \frac{3}{n}$$

So we need $\frac{3}{n} < 0.1$, that is, $n > 30$.

5.(a) The problem gives us the marginal distribution

$$f_1(x) = \begin{cases} \frac{1}{3} & \text{if } x \in [0,3] \\ 0 & \text{otherwise} \end{cases}$$

and the conditional distribution

$$g_2(y|x) = \begin{cases} \frac{1}{x} & \text{if } y \in [0,x] \\ 0 & \text{otherwise.} \end{cases}$$

Hence we can compute the joint distrbution

$$f(x,y) = g_2(y|x)f_1(x) = \begin{cases} \frac{1}{3x} & \text{if } 0 \leq y \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

and then the other marginal distribution

$$f_2(y) = \int_{x=-\infty}^{x=+\infty} f(x,y)dx = \begin{cases} \int_{x=y}^{x=3} \frac{1}{3x} dx = \frac{1}{3}\left(\log(3) - \log(y)\right) & \text{if } 0 \leq y \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$E(Y|X=x) = \int_{y=-\infty}^{y=+\infty} y g_2(y|x)dy = \int_{y=0}^{y=x} \frac{y}{x} dy = \frac{x}{2}.$$

(c) This is asking for the median value m of Y conditioning on the value $X = \frac{1}{2}$, that is, the solution m to this equation:

$$\begin{aligned} \frac{1}{2} &= \int_{y=-\infty}^{y=m} g_2\left(y|x = \frac{1}{2}\right) dy \\ &= \int_{y=0}^{y=m} 2dy \\ &= 2m \end{aligned}$$

and hence $m = \frac{1}{4}$.