Math 461 B/C, Spring 2009
Midterm Exam 3 Solutions and Comments

1. Let X have normal distribution with mean 1 and variance 4.

   (a) Find $P(|X| \le 1)$.

10 pts

Solution.

$$
\begin{aligned}
P(|X| \le 1) &= P(-1 \le X \le 1) \\
&= P\left(\frac{1-1}{2} \le \frac{X-1}{2} \le \frac{-1-1}{2}\right) \\
&= \Phi(0) - \Phi(-1) = \Phi(0) - (1 - \Phi(1)) \\
&= 0.5 - (1 - 0.84) = \boxed{0.34}
\end{aligned}
$$

   (b) Write down the p.d.f. $f_X(x)$ of X. (The formula should be an explicit, elementary function of x; in particular, it should not involve Φ.)

10 pts

Solution. The general formula for the density of a normal distribution with parameters μ and σ is $f(x) = (1/\sqrt{2\pi}\sigma)e^{-(x-\mu)^2/(2\sigma^2)}$. Here $\mu = 1$, $\sigma = \sqrt{4} = 2$, so

$$
f_X(x) = \boxed{\frac{1}{2\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-1}{2}\right)^2\right\}, \quad -\infty < x < \infty}
$$

   (c) Let $Y = e^X$. Find the p.d.f. $f_Y(y)$ of Y. (Again, the formula should be an explicit, elementary function of y.)

10 pts

Solution. [This is a standard change-of-variables exercise, of much the same type as Problem 39, Chapter 5, from HW 7.]

We start by computing the c.d.f. of Y: The range of Y is $(0, \infty)$ (since $Y = e^X$ and X has range $(-\infty, \infty)$), and for y in this range

$$
(*) \qquad F_Y(y) = P(Y \le y) = P(e^X \le y) = P(X \le \ln y) = \Phi\left(\frac{\ln y - 1}{2}\right).
$$

Differentiating, we get the density:

$$
\begin{aligned}
f_Y(y) &= \frac{d}{dy}\Phi\left(\frac{\ln y - 1}{2}\right) = \frac{1}{2y}\Phi'\left(\frac{\ln y - 1}{2}\right) \\
&= \frac{1}{2y\sqrt{2\pi}} \exp\left(-\frac{1}{2 \cdot 2^2}(\ln y - 1)^2\right) \\
&= \boxed{\frac{1}{2y\sqrt{2\pi}} e^{-(1/8)(\ln y - 1)^2} \quad (0 < y < \infty)}
\end{aligned}
$$

Comments: The key to the solution is the relation $(*)$, which relates the **c.d.f.** of Y, $F_Y(y)$, to the **c.d.f.** of X, and to the Φ-function. As emphasized in class and in hw solutions, the correct approach in change-of-variables problems for densities is via the corresponding c.d.f.'s. To earn credit on this problem, this key step has to be present. Simply substituting $x = \ln y$ into the formula for the density of X would be completely wrong, and although it leads to an answer that looks superficially similar to the correct answer, the approach is totally flawed.

2. Let X and Y be independent random variables, each exponentially distributed with mean 1/10.

(a) Find $P(X \geq 5Y)$.

10 pts **Solution.** [This is a variation of the "light-bulb race" problem from class, which asked for the probability $P(X \geq Y)$, with X and Y having different exponential distributions.]
Since an exponential distribution with parameter λ has mean $\mu = 1/\lambda$, the parameter λ in the distribution of X and Y must be $\lambda = 1/(1/10) = 10$, and the joint c.d.f. is

$$f(x,y) = f_X(x)f_Y(y) = 10e^{-10x}10e^{-10y}, \quad 0 < x < \infty, 0 < y < \infty.$$

Hence,

$$\begin{aligned}
P(X \geq 5Y) &= \int_{y=0}^{\infty}\int_{x=5y}^{\infty} 10e^{-10x}10e^{-10y}dydx \\
&= \int_{y=0}^{\infty} e^{-5\cdot 10y}10e^{-10y}dy \\
&= \int_{y=0}^{\infty} 10e^{-60y}dy = \boxed{\frac{1}{6}}
\end{aligned}$$

Remark: The parameter λ in the exponential distribution is **not** equal to the mean μ, but the reciprocal: Thus, $\lambda = 1/\mu = 1/(1/10) = 10$, and **not** $\lambda = 1/10$. While the final answer, 1/6, ends up being the same with the latter choice of λ, the argument is not correct if one works with $\lambda = 1/10$.

(b) Let Z be the maximum (i.e., larger) of X and Y. Find the density, $f_Z(z)$, of Z.

10 pts **Solution.** [This is a variation on Problem 49(b), Chapter 6 from HW 9, and an illustration of the max/min trick that came up in class on several occasions.]
As usual, we first compute the c.d.f. $F_Z(z)$. We have

$$\begin{aligned}
F_Z(z) &= P(Z \leq z) = P(\max(X,Y) \leq z) = P(X \leq z, Y \leq z) \text{ (by the maximum trick)} \\
&= P(X \leq z)P(Y \leq z) \text{ (by the independence of } X \text{ and } Y\text{)} \\
&= (1 - e^{-10z})^2, \quad 0 \leq z < \infty.
\end{aligned}$$

Now take the derivative to get the p.d.f.:

$$f_Z(z) = F_Z'(z) = 2(1 - e^{-10z})10e^{-10z} = \boxed{20\left(e^{-10z} - e^{-20z}\right), \quad 0 \leq z < \infty}$$

(c) Let $S = X + Y$. Find the density, $f_S(s)$, of S.

10 pts **Solution.** [This is just like HW Problem 27 in Chapter 6, but simpler, since the case distinction necessary in that hw problem is not needed here.]
As in the hw problem, we use the convolution formula for the density of a sum of two independent random variables: We have, for $s \geq 0$,

$$\begin{aligned}
f_S(s) &= \int_{x=-\infty}^{\infty} f_X(x)f_Y(s-x)dx \\
&= \int_{x=0}^{s} 10e^{-10x}10e^{-10(s-x)}dx \\
&= 100e^{-10s}\int_{x=0}^{s} dx \\
&= \boxed{100se^{-10s}, \quad 0 \leq s < \infty}
\end{aligned}$$

3. Let X have uniform distribution on the interval $(0, 1)$. Given $X = x$, let Y have uniform distribution on the interval $(0, x)$.

   (a) Find the joint density of X and Y. Be sure to specify the range.

10 pts

Solution. [This is a problem worked out in class.]
The given assumptions on X and Y are:
(1) X has uniform distribution on $[0, 1]$, and
(2) given $X = x$, Y has uniform distribution on $(0, x)$.
This translates into

(1) $$f_X(x) = 1, \quad 0 \le x \le 1,$$

(2) $$f_{Y|X}(y|x) = \frac{1}{x}, \quad 0 \le y \le x.$$

Hence the joint density is

$$f(x, y) = f_{Y|X}(y|x) f_X(x) = \frac{1}{x} \cdot 1 = \boxed{\frac{1}{x}, \quad 0 \le x \le 1, 0 \le y \le x}$$

   (b) Find the marginal density $f_Y(y)$ of Y. Be sure to specify the range.

10 pts

Solution.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx = \int_{x=y}^{1} \frac{1}{x}dx = \boxed{-\ln y, \quad 0 < y < 1},$$

Comments: As pointed out in class and in the solutions to several hw problems (e.g., Problem 8, Chapter 6, from HW 8), in computing marginal densities it is absolutely crucial to keep track of ranges of densities and to use these ranges in determining integration limits. In the above situation, integrating from $x = 0$ to $x = 1$ (instead of $x = y$ to $x = 1$) would be a fatal mistake, and would result in a nonsensical answer, namely, the integral $\int_0^1 \frac{1}{x}dx$, which does not exist.

   (c) Find $E(XY)$.

10 pts

Solution.

$$E(XY) = \iint xyf(x, y)dydx = \int_{x=0}^{1} \int_{y=0}^{x} xy\frac{1}{x}dydx$$
$$= \int_{x=0}^{1} \frac{x^2}{2}dx = \boxed{\frac{1}{6}}$$

4. The following questions are independent of each other.

   (a) Suppose that X and Y are independent random variables with $\mathrm{Var}(X) = 1$, $\mathrm{Var}(Y) = 2$. Find $\mathrm{Var}(1 - 2X + 3Y)$.

8 pts

Solution. [This is an exercise in using the properties of a variance.]

$$\mathrm{Var}(1 - 2X + 3Y) = 0 + (-2)^2 \mathrm{Var}(X) + 3^2 \mathrm{Var}(Y) = 4 \cdot 1 + 9 \cdot 2 = \boxed{22}.$$

(b) Suppose X and Y are random variables such that $\text{Var}(X+Y) = 15$ and $\text{Var}(X-Y) = 5$. Find $\text{Cov}(X, Y)$.

8 pts

Solution. [This is an exercise in using the properties of a covariance.]
We have

$$\begin{aligned}
\text{Var}(X+Y) &= \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X, Y), \\
\text{Var}(X-Y) &= \text{Var}(X) + \text{Var}(-Y) + 2\,\text{Cov}(X, -Y) \\
&= \text{Var}(X) + \text{Var}(Y) - 2\,\text{Cov}(X, Y),
\end{aligned}$$

(since $\text{Var}(-Y) = \text{Var}(Y)$ and $\text{Cov}(X, -Y) = -\text{Cov}(X, Y)$). Subtracting these two equations and substituting the given values for $\text{Var}(X+Y)$ and $\text{Var}(X-Y)$, we get $15 - 5 = 4\,\text{Cov}(X, Y)$, so $\text{Cov}(X, Y) = \boxed{2.5}$.

(c) Suppose that X and Y are independent random variables with moment-generating functions

$$M_X(t) = (1-2t)^{-2}, \quad M_Y(t) = (1-2t)^{-3}.$$

Find $\text{Var}(X+Y)$.

9 pts

Solution. [This is a simplified form of Problem 9 from HW 10.]
Since X and Y are independent, $X+Y$ has mgf

$$M_{X+Y}(t) = M_X(t) M_Y(t) = (1-2t)^{-5}.$$

Hence,

$$\begin{aligned}
M'_{X+Y}(t) &= 5 \cdot 2(1-2t)^{-6}, \quad M'_{X+Y}(0) = 10, \\
M''_{X+Y}(t) &= 10 \cdot 6 \cdot 2(1-2t)^{-7}, \quad M''_{X+Y}(0) = 120, \\
\text{Var}(X+Y) &= M''_{X+Y}(0) - M'_{X+Y}(0)^2 = 120 - 10^2 = \boxed{20}
\end{aligned}$$

Comments: (1) This problem was meant to be done in the same way as Problem 9 from HW 10, namely, using the multiplication property of mgf's. With this method the calculations required are minimal. One could also compute separately $\text{Var}(X)$ and $\text{Var}(Y)$ via the corresponding mgf's, and adding up the two variances, but this requires much more calculation.
(2) It is crucial that the derivatives of the moment-generating functions be evaluated at $t = 0$ to get the moments. Without plugging in $t = 0$ one would get a function of t as answer, which makes no sense since expectation and variance are numerical quantities and thus should not involve any variables.

(d) (True/false questions) For each of the following statements, indicate **clearly** whether the statement is true or false by placing a mark **(T)** (True) or **(F)** (False) next to the question; **if you are unsure, leave the answer blank**.

Grading: 3 points for a correct answer, 0 points for an answer left blank, and -2 points for each incorrect answer. Thus, the maximal score is 15 points, 4 correct and 1 incorrect answer result in a total score of $4 \cdot 3 + 1 \cdot (-2) = 10$ points, while 4 correct answers and one answer left blank results in a total of 12 points.

A. If X and Y are independent, then $\text{Cov}(X, Y) = 0$.
B. If $E(XY) = E(X)E(Y)$, then X and Y are independent.
C. If X and Y are independent, then $M_{XY}(t) = M_X(t) M_Y(t)$.
D. If X and Y are independent continuous r.v.'s, then $f_{X|Y}(x|y) = f_X(x)$.

E. If $X_1, X_2, \ldots$ are mutually independent r.v.'s, and $S_n = \sum_{i=1}^n X_i$, then $S_1, S_2, \ldots$, are also mutually independent.

15 pts

Solution.

A. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.
$\boxed{\text{True}}$ (One of the properties of independent r.v.'s. (cf. p. 2 of the variance/covariance/mgf handout).)

B. If $E(XY) = E(X)E(Y)$, then X and Y are independent.
$\boxed{\text{False}}$ (The condition $E(XY) = E(X)E(Y)$ is equivalent to $\mathrm{Cov}(X, Y) = 0$, but the latter does not imply that X and Y are independent (cf. the note on p. 2 of the variance/covariance/mgf handout).)

C. If X and Y are independent, then $M_{XY}(t) = M_X(t)M_Y(t)$.
$\boxed{\text{False}}$ (Independence implies that the mgf of the *sum* (not the product) is the product of the individual mgf' (cf. the note on p. 2 of the variance/covariance/mgf handout).)

D. If X and Y are independent continuous r.v.'s, then $f_{X|Y}(x|y) = f_X(x)$.
$\boxed{\text{True}}$ By the definition of the conditional density and the independence of X and Y,

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x).$$

E. If $X_1, X_2, \ldots$ are mutually independent r.v.'s, and $S_n = \sum_{i=1}^n X_i$, then $S_1, S_2, \ldots$, are also mutually independent.
$\boxed{\text{False}}$ (In Problem 5 in HW 10, we calculated that $\mathrm{Cov}(S_n, S_k)$ equals $k\sigma^2$ if $k \le n$, and in particular is not 0. Hence S_n and S_k cannot be independent.)

5. Suppose two points are chosen independently and uniformly from the interval $[0, 1]$. Let D denote the distance (in absolute value) between these two points.

(a) Find $P(D \le 0.1)$.

10 pts

Solution. [This problem was worked in class. Similar problems appeared in the hw assignments; e.g., Problem 13, Chapter 6, in HW 8.]
We have $D = |X - Y|$, where X and Y denote the two points, and we need to compute $P(|X - Y| \le 0.1)$. Since X and Y are independent and uniformly distributed on $[0, 1]$, the joint distribution of X and Y is uniform over the unit square, we can compute probabilities involving X and Y via areas. With R denoting the region on which $|x - y| \le 0.1$ (see sketch), we get

$$P(D \le 0.1) = P(|X - Y| \le 0.1) = \frac{\text{Area}(R)}{\text{Area(Unit Square)}}$$
$$= \frac{1 - (0.9^2}{1} = \boxed{0.19}$$

Comments: The above geometric approach, via areas, is the one used in the class problem and the homework problems referred to above. One could, in principle, compute $P(|X - Y| \le 0.1)$ via double integrals, but this approach is much more complicated and highly error prone. In particular, to get the correct integration limits one needs to split up the integral into three separate double integrals, corresponding to the ranges $0 \le x \le 0.1$, $0.1 \le x \le 0.9$, and $0.9 \le x \le 1$. (See picture.)

(b) Find $E(D^2)$.

10 pts

Solution. Write $D^2 = (X - Y)^2$, expand the square and use the properties of an expectation:

$$\begin{aligned} E(D^2) &= E((X-Y)^2) = E(X^2) - 2E(X)E(Y) + E(Y^2) \\ &= \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} = \boxed{\frac{1}{6}} \end{aligned}$$

(since $E(X^2) = \int_0^1 x^2 \cdot 1 dx = 1/3$, and $E(X) = \int_0^1 x dx = 1/2$ for a uniform distribution on $[0, 1]$).

(c) (Extracredit) The following is a 2-dimensional analog of the previous question: For this part, suppose two random points are chosen independently and uniformly from the *unit square* $[0, 1] \times [0, 1]$, and let D denotes the distance (in the usual geometric sense) between the two points. Find $E(D^2)$.

Note: To earn the extra credit requires a rigorous mathematical solution, using an appropriate probabilistic framework, with properly defined random variables, correct notation, etc. No credit will be given for answers arrived at by informal reasoning, intuition, guessing, etc. No questions, hints , pointers, etc. on this problem (otherwise, it wouldn't deserve the "extracredit" designation). Use back of page if you need more space.

10 pts

Solution. Let (X_1, Y_1) and (X_2, Y_2) denote the coordinates of the two points. By assumption, the X_i's, Y_i's are mutually independent, and each is uniformly distributed on the interval $[0, 1]$. The square of the distance between these two points is

$$D^2 = (X_2 - X_1)^2 + (Y_2 - Y_1)^2.$$

Hence,

$$\begin{aligned} E(D^2) &= E((X_2 - X_1)^2 + (Y_2 - Y_1)^2) \\ &= E((X_2 - X_1)^2) + E((Y_2 - Y_1)^2). \end{aligned}$$

Squaring out and using the properties of the expectation and the independence assumptions, we get

$$E((X_2 - X_1)^2) = E(X_2^2) - 2E(X_1)E(X_2) + E(X_1^2) = \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} = \frac{1}{6}$$

since $E(X_i) = \int_0^1 x dx = \frac{1}{2}$ and $E(X_i^2) = \int_0^1 x^2 dx = \frac{1}{3}$. By symmetry, $E((Y_2 - Y_1)^2) = \frac{1}{6}$, so

$$E(D^2) = \frac{1}{6} + \frac{1}{6} = \boxed{\frac{1}{3}}$$

Maximal Total Score (excluding extracredit points): $\boxed{\text{150 points}}$