MATH 151, FINAL EXAM
Winter Quarter, 21 March, 2014

Time: 3 hours, 8:30-11:30

Instructions:

(1) Write your name in blue-book provided and sign that you agree to abide by the honor code.

(2) The exam consists of 6 questions, each worth 15 points. The breakdown of points appears at the beginning of each question. Read all questions carefully and answer in the blue book.

(3) All work must be shown for full credit.

(4) Permitted materials: You may use homework, class notes, calculator and use or cite the text book.
Question 1 (5+5+5)

There are two urns, one urn containing 3 black balls and 6 white balls, while the other urn contains 100 white balls. An urn is selected uniformly at random and then a ball is drawn uniformly at random from the chosen urn.

(a) What is the probability that the drawn ball is black?

ANS: Let $U_1$ denote the event of choosing the urn having only white balls and $B_1$ the event that black ball is drawn. Clearly, $P(B_1|U_1) = 0$ and since the urn is chosen uniformly we know that $P(U_1) = P(U_1^c) = 1/2$, while the uniform drawing of ball from the urn means that $P(B_1|U_1^c) = 3/(3 + 6) = 1/3$. Consequently,

$$P(B_1) = P(B_1|U_1)P(U_1) + P(B_1|U_1^c)P(U_1^c) = 0 \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$  

(b) If a black ball was drawn then it is thrown away, whereas if a white ball was drawn it is returned to its original urn. After that, a second ball is drawn as above (that is, first selecting an urn uniformly at random, then a ball uniformly from it). What is the probability that the second ball is black?

ANS: Let $B_2$ and $U_2$, denote the events of choosing in second drawing a black ball, or the urn having only white balls, respectively. In the event of $B_1^c$ (white ball in first drawing), we repeat the same experiment as in part (a), hence $P(B_2|B_1^c) = \frac{1}{6}$. If event $B_1$ occured, we again have $P(U_2) = P(U_2^c) = 1/2$ and $P(B_2|U_2) = 0$. However, we discarded the (first) chosen black ball, so have now only 2 black balls in the urn from which the first ball came, hence $P(B_2|U_2^c, B_1) = 2/(2 + 6) = 1/4$, yielding similarly to part (a) that $P(B_2|B_1) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$. Combining all of the above we arrive at,

$$P(B_2) = P(B_2|B_1^c)P(B_1^c) + P(B_2|B_1)P(B_1) = \frac{1}{6} \times \frac{5}{6} + \frac{1}{8} \times \frac{1}{6} = \frac{40 + 6}{36 \times 8} = \frac{23}{144}.$$  

(c) Suppose that the first drawn ball was white and returned to its original urn. What is the probability that another ball drawn from that same urn, will be black?

ANS: In solving part (a) you saw that the probability $5/6$ of event $B_1^c$ (first drawn ball was white), is composed of $P(B_1^c, U_1) = 1/2$ while $P(B_1^c, U_1^c) = 1/3$. So, by the definition of the conditional probability $P(U_1|B_1^c) = (1/2)/(5/6) = 3/5$. This is precisely the probability of using the urn $U_1$ in the experiment described in part (c), so repeating the analysis of part (a) we conclude that the event $B$ of getting black ball when using same urn twice, has the following probability conditioned on $B_1^c$,

$$P(B|B_1^c) = P(B|U_1, B_1^c)P(U_1|B_1^c) + P(B|U_1^c, B_1^c)P(U_1^c|B_1^c) = 0 \times \frac{3}{5} + \frac{1}{3} \times \frac{2}{5} = \frac{2}{15}.$$  

2
Question 2 (3+4+4+4)

Suppose continuous random variables \( X \) and \( Y \) have the joint probability density function 
\[
f_{X,Y}(x,y) = c(y-x) \quad \text{for} \quad 0 < x < y < 1.
\]

(a) Find the value of \( c \).

**ANS:** Recall that \( \int \int f_{X,Y}(x,y) \, dx \, dy = 1 \), which means here that
\[
1 = c \int_0^1 dx \int_x^1 (y-x) \, dy = c \int_0^1 dx \int_0^{1-x} t \, dt = \frac{c}{2} \int_0^1 (1-x)^2 \, dx = \frac{c}{6} u^3 \bigg|_0^1 = \frac{c}{6}
\]
(using change of variables, first \( t = y-x \), then \( u = 1-x \)). In conclusion, \( c = 6 \).

(b) Find \( P(Y > 2X) \).

**ANS:** For subset \( A = \{(x,y) : y > 2x\} \) of \( \mathbb{R}^2 \), we use the formula \( P((X,Y) \in A) = \int_A f_{X,Y}(x,y) \, dx \, dy \) (and note that if \( x > 0.5 \) then \( 2x > 1 \)), to deduce that
\[
P(Y > 2X) = \int_0^{0.5} dx \int_{2x}^1 c(y-x) \, dy = c \int_0^{0.5} dx \int_x^{1-x} t \, dt
\]
\[
= \int_0^{0.5} [3(1-x)^2 - 3x^2] \, dx = u^3 \bigg|_{0.5}^1 - u^3 \bigg|_0^{0.5} = 1 - 2 \times 0.5^3 = \frac{3}{4}
\]
(using same change of variables as in part (a)).

(c) Find the marginal density \( f_Y(y) \) of \( Y \).

**ANS:** By definition, the marginal density of \( Y \) is non-zero only when \( 0 < y < 1 \), in which case it is given by
\[
f_Y(y) = \int_0^y c(y-x) \, dx = 3 \int_0^y (2t) \, dt = 3y^2.
\]

(d) Find \( E[Y] \).

**ANS:** By part (c) and the definition of expectation of continuous random variable \( Y \),
\[
E[Y] = \int_0^1 y f_Y(y) \, dy = \int_0^1 3y^3 \, dy = \frac{3}{4} u^4 \bigg|_0^1 = \frac{3}{4}
\]
(or use \( E[Y] = \int \int y f_{X,Y}(x,y) \, dx \, dy \) to bypass part (c), if having difficulty with the latter).

Question 3 (5+5+5)

A fair die is rolled 720 times.

(a) Use normal approximation to find the probability that the number of times in which the die lands on face 6, is strictly between 100 and 140.
ANS: The number $S_6$ of times in which the die lands on face 6 follows the Binomial($n, p$) distribution with $n = 720$ and $p = 1/6$. As such it has mean $E(S_6) = np = 120$ and standard-deviation $SD(S_6) = \sqrt{np(1-p)} = 10$. With 140 and 100 being $120 \pm 2 \times 10$, i.e. two standard-deviations from the mean, the approximation for the probability $P(100.5 < S_6 < 139.5)$ in question, by a standard normal variable $Z$, is thus $P(|Z| \leq 1.95)$. That is, $\Phi(1.95) - \Phi(-1.95) = 2 \times 0.9744 - 1 = 0.9488$ (see Table 5.1, page 190 of text book).

(b) Suppose this die landed on face 6 in precisely 95 of its 720 rolls. What is then the probability that it landed on face 5 in at most 140 among these 720 rolls?

ANS: Excluding the $S_6$ rolls in which the die landed on face 6, there are $n' = n - S_6$ independent rolls in each of which the outcome is uniformly chosen among $\{1, 2, 3, 4, 5\}$. Therefore, conditional on the value of $S_6$, the number of times $S_5$ in which the die landed on face 5, is again Binomial, now of parameters $n' = n - S_6 = 720 - 95 = 625$ and $p' = 1/5$. This Binomial variable has mean $E(S_5|S_6 = 95) = n'p' = 625/5 = 125$ and $SD(S_5|S_6 = 95) = \sqrt{n'p'(1-p')} = 10$. With $140 = 125 + 1.5 \times 10$ being 1.5 standard-deviations from the mean, the probability $P(S_5 < 140.5|S_6 = 95)$ in question, has the normal approximation $P(Z < 1.55) = \Phi(1.55) = 0.9394$ (or 0.9265, for those who interpreted "at most 140" as meaning $S_5$ being strictly smaller than 140).

(c) What is the least number of times you need to roll such fair die in order for the proportion of times it lands on face 6 be in the interval $[1 - 0.04/6, 1 + 0.04/6]$ with probability at least 0.95?

ANS: The proportion of times the die lands on face 6 is $n^{-1}S_6$. By part (a) and linearity with respect to scaling of the mean and standard deviation, this random variable has mean $1/6$ and standard-deviation $\sigma = \sqrt{5/n}/6$. Under the normal approximation, we ask for an interval of $\pm x$ from the mean (with $x = 0.04/6$), to have probability $\Phi(x/\sigma) - \Phi(-x/\sigma) \geq 0.95$. Hence $\Phi(x/\sigma) \geq 0.975$, implying by Table 5.1 of text that $x \geq 1.96\sigma$. Solving for the given values of $x = 0.04/6$ and $\sigma = \sqrt{5/n}/6$, yields $n \geq 5 \times 49^2 = 12,005$.

Question 4 (3+4+4+4)

Suppose random variables $X_i$, $i = 1, 2, \ldots, n$ have mean 0, variance 1 and are (pairwise) uncorrelated (but not necessarily independent). That is, $\text{Cov}(X_i, X_j) = 0$ for any $i \neq j$. Let $S_n = \sum_{i=1}^{n} X_i$ and find the values of:

(a) $E[S_n]$.

ANS: By linearity of the mean (see formula (7.2.2), page 282 of text), we have that $E[S_n] = \sum_{i=1}^{n} E(X_i) = 0$.  

4
(b) \( \text{Var}[S_n] \).

**ANS:** By bi-linearity of the variance (see formula (7.4.1), page 306 of text), we see that

\[
\text{Var}[S_n] = \sum_{i=1}^{n} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) = n + 0 = n.
\]

(c) The limit as \( n \to \infty \) of \( P(|n^{-1}S_n| \geq \epsilon) \) for any fixed \( \epsilon > 0 \).

**ANS:** Applying Chebychev’s inequality for \( S_n \) of mean \( \mu_n = 0 \) and variance \( \sigma_n^2 = n \), we deduce that

\[
P(|n^{-1}S_n| \geq \epsilon) = P(|S_n - \mu_n| \geq n\epsilon) \leq \frac{\sigma_n^2}{(n\epsilon)^2} = \frac{1}{n\epsilon^2}
\]

which for any fixed \( \epsilon > 0 \) goes to zero as \( n \to \infty \).

(d) The correlation \( \text{Cor}(X_1 - X_2, X_2 + X_3) \).

**ANS:** The key is the bi-linearity of the covariance (parts (iii) and (iv) of Proposition 7.4.2 of the text). Thanks to it we have that

\[
\text{Cov}(X_1 - X_2, X_2 + X_3) = \text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_3) - \text{Cov}(X_2, X_2) - \text{Cov}(X_2, X_3)
\]

\[
= 0 + 0 - 1 - 0 = -1,
\]

by our assumption that \( X_i \) are uncorrelated and of variance 1. Similarly,

\[
\text{Var}(X_1 - X_2) = \text{Cov}(X_1, X_1) - 2\text{Cov}(X_1, X_2) + \text{Cov}(X_2, X_2) = 1 - 2 \times 0 + 1 = 2,
\]

and likewise, \( \text{Var}(X_2 + X_3) = 2 \). In conclusion, \( \text{Cor}(X_1 - X_2, X_2 + X_3) = -1/(\sqrt{2} \sqrt{2}) = -1/2 \) (see formula in page 310 of text; we also went over it in class).

**Question 5 (3+4+4+4)**

Suppose \( X \) and \( Y \) are independent, continuous, non-negative random variables of densities \( f_X(x) \) and \( f_Y(y) \).

(a) What is the joint density \( f_{X,Y}(x,y) \) of \( (X,Y) \)?

**ANS:** By independence we have that \( f_{X,Y}(x,y) = f_X(x)f_Y(y) \).

(b) Show that the density \( f_Z(z) \) of \( Z = X/Y \) is given by \( f_Z(z) = \int_0^\infty y f_Y(y) f_X(yz) dy \).

**ANS:** We start with the joint density formula of Section 6.7 of text, to get \( f_{Z,Y} \) in terms of \( f_{X,Y} \) of part (a). Here \( g_1(x,y) = x/y \) and \( g_2(x,y) = y \), resulting with matrix of partial derivatives having the entries \( \partial g_1/\partial x = 1/y \), \( \partial g_1/\partial y = -x/y^2 \), \( \partial g_2/\partial x = 0 \) and \( \partial g_2/\partial y = 1 \), hence determinant \( J = 1/y \). Since \( X = ZY \), this leads to \( f_{Z,Y}(z,y) = f_Y(y)f_X(yz)y \). We then integrate over \( y \geq 0 \) to get the stated marginal density \( f_Z(z) \).
(c) Find the distribution function $F_Z(z)$ of $Z = X/Y$ in case $X$ and $Y$ are independent exponential random variables of parameter 1.

**ANS:** Setting $\mu = 1 + z$, we have from the identity stated in part (b) and the formula for exponential density, that

$$f_Z(z) = \int_0^\infty ye^{-y}e^{-zy}dy = \int_0^\infty ye^{-\mu y}dy = \mu^{-2} \int_0^\infty te^{-t}dt = \mu^{-2}$$

( usando change of variable $t = \mu y$ and recalling that exponential(1) variables have mean 1).

Having $f_Z(z) = (1 + z)^{-2}$ for $z \geq 0$, we integrate to get the distribution function

$$F_Z(z) = \int_0^z (1 + u)^{-2}du = 1 - \frac{1}{1 + z} = \frac{z}{1 + z}.$$

(d) Let $q_n(x) = P(\max_{1 \leq i \leq n} X_i \leq \log n + x)$ for independent exponential variables $X_i, i = 1, 2, \ldots$, of parameter 1. Compute $q_n(x)$ for each fixed $x$ and positive integer $n$, then find the limit of $q_n(x)$ when $n \to \infty$.

**ANS:** Since $X_i$ are independent and identically distributed,

$$q_n(x) = \prod_{i=1}^n P(X_i \leq \log n + x) = F_X(\log n + x)^n.$$

Setting $F_X(x) = 1 - e^{-x}$ we get that $q_n(x) = (1 - n^{-1}e^{-x})^n$ whose limit as $n \to \infty$ is $\exp(-e^{-x})$.

**Question 6 (5+5+5)**

Jack and Jill eat lunch together. Jack arrives at 12:00+T, where $T$ has an *exponential distribution* with a *mean value* of 2 minutes. Every whole minute, starting at 11:57AM, Jill’s watch makes one ”beep” sound. Jill independently notices each ”beep” with a chance $p = 0.2$ and arrives exactly 3 minutes after she first noticed a ”beep”.

(a) What is the probability of both Jack and Jill arriving before 12:03PM?

**ANS:** Jill arrives at time $11:59+N$ where $N = 1, 2, \ldots$ has the geometric distribution of success probability $p$. Noting that $T$ is an exponential variable of parameter $\lambda = 1/E(T) = 1/2$ independent of $N$, we see that the probability of interest is

$$P(N \leq 3, T < 3) = P(N \leq 3)P(T < 3) = (p + (1-p)p + (1-p)^2p)(1 - e^{-3\lambda})$$

$$= 0.488(1 - e^{-1.5}) = 0.379.$$

(b) What is the probability that Jack arrives before Jill?
ANS: Using the notations of part (a), Jack arrives before Jill if and only if \( N > T + 1 \), the probability of which is by independence of \( T \) and \( N \), given by

\[
P(N > T + 1) = 1 - P(N \leq T + 1) = 1 - \sum_{k=0}^{\infty} P(N = k + 1)P(T \geq k)
\]

\[
= 1 - p \sum_{k=0}^{\infty} (1 - p)^k e^{-\lambda k} = 1 - \frac{p}{1 - (1 - p)e^{-\lambda}} = 0.61
\]

(for \( p = 0.2 \) and \( \lambda = 1/2 \)).

(c) Their bill is settled by flipping fair coin \( n = 5 \) times during lunch, with Jill paying the bill if there is a string of \( k = 3 \) consecutive heads among these \( n \) coin flips, and otherwise Jack pays it. Who is more likely to pay the bill, Jack or Jill?

ANS: The probability that Jill pays the bill is \( P(L_n \geq k) \) which is given in page 143 of the text book. For \( n = 5 \) and \( k = 3 \) the only relevant terms in this formula are \( r = 1, 2, 3 = n - k + 1 \). Further, since \( n - 2k = 5 - 6 < 0 \), it suffices to take only \( r = 1 \). For fair coin, i.e. \( p = 1/2 \), it yields \( P(L_5 \geq 3) = (n - k + 2)2^{-(k+1)} = 1/4 \). In conclusion, Jack is more likely to pay the lunch bill.