TEST #3
STA 5326 **Name:**__________________________________
December 4, 2014

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

Directions

- This exam is **closed book** and **closed notes**. (You will have access to a copy of the "Table of Common Distributions" given in the back of the text.)
- 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.)
- The different problems are not related. The different parts of a problem are sometimes unrelated. If you cannot solve part of a problem, you should still go on to look at the later parts.
- Show and explain your work (including your calculations) for all problems unless you are explicitly told otherwise. **No credit is given without work.** But don't get carried away! Show enough work so that what you have done is clearly understandable.
- The grader should be able to see how you got from one step to the next. If you needed scratch paper to work something out, make sure to transfer your work to the exam.
- If you need more space for a problem than what is given, use the back of the same page and write **Work on Back** to indicate you have done so.
- On problems where work is shown, **circle your answer**. (On these problems, you should give only one answer!)
- All the work on the exam should be your own. No "cooperation" is allowed.
- Arithmetic does **not** have to be done completely **unless requested otherwise**. Answers can be left as fractions or products. You do not have to evaluate binomial coefficients, factorials or large powers. Answers can be left as summations (unless there is a simple closed form such as when summing a geometric or exponential series).
- You need only pens, pencils, erasers and a calculator. (You will be supplied with scratch paper.)
- 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 **11** pages.
- There are a total of **100** points.

Problem 1. (10%) Prove that $EY = E[E(Y \mid X)]$.

(Assume that (X, Y) has a joint density $f_{X,Y}(x, y)$ and that EY exists.)

See notes 11, page 2.

If students exactly reproduce the argument in the notes, give them full credit. But you should look carefully: in going from line 2 to line 3 of the proof, two things happen; y is brought inside the inner integral (over x), and then the order of integration is changed.

Problem 2. (10%) Suppose that

$$f_X(x) = \begin{cases} 6x(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

and that for $0 < x < 1$,

$$f_{Y|X}(y|x) = \begin{cases} \dfrac{2y}{x^2} & \text{for } 0 < y < x \\ 0 & \text{otherwise}\,. \end{cases}$$

Find $\mathrm{Var}(Y|X = x)$ for $0 < x < 1$.

See notes9, page 19. The calculation is similar to (but simpler than) the calculation of $Var(Y|X = x)$ on notes9, page 22 (which uses $E(Y|X = x)$ on notes9, pages 17). Note that $f_X(x)$ is never used.

Problem 3. (12%) Prove that if the joint cumulative distribution function (cdf) satisfies

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)\,,$$

then for any pair of intervals (a,b) and (c,d),

$$P(a \leq X \leq b, c \leq Y \leq d) = P(a \leq X \leq b)P(c \leq Y \leq d)\,.$$

This is exercise 4.9.

The solution in the manual is not quite right; it handles the lower endpoints of the intervals incorrectly. But if students reproduce this solution, give them full credit.

Problem 4. (12%) Let X = the number of trials to get the first head and Y = the number of trials to get the **first tail** in repeated tosses of a fair coin. Are X and Y independent? (Answer "Yes" or "No" and give a detailed justification of your answer.)

This is similar to exercise 4.11. The answer is "No".
The simplest justification is to note that $P(X = 1, Y = 1) = 0$ *and* $P(X = 1) = P(Y = 1) = 1/2$ *so that* $P(X = 1, Y = 1) \neq P(X = 1)P(Y = 1)$ *which violates independence.*

Another way to justify "No" is to note that the support of the distribution of (X, Y) *is* **not** *a product set. The support consists of the possibilities marked with a plus sign below:*

	1	2	3	4	5	...
1	0	+	+	+	+	...
2	+	0	0	0	0	
3	+	0	0	0	0	
4	+	0	0	0	0	
5	+	0	0	0	0	
⋮	⋮					

This follows by noting that either X *or* Y *must be 1 (but not both), and whichever is not 1 can take any larger value.*

For students who give this type of solution, it is not sufficient to just say "the support is not a product set". To get full credit, students must actually specify the support or give a specific example demonstrating the support is not a product set (e.g. they could note that $(1, 1)$ *is not in the support of* (X, Y)*, but that 1 is in the support of both* X *and* Y*, which violates the product set condition).*

Problem 5. (12%) Let (X, Y) be a bivariate random vector with joint pdf $f(x, y)$ which is positive on the entire plane R^2. Let

$$U = a\,e^X + b \quad \text{and} \quad V = c\,e^Y + d$$

where a, b, c, and d are fixed constants with $a > 0$ and $c > 0$. Find the joint pdf of (U, V). (Make sure to specify the support of the joint pdf.)

This is similar to exercise 4.22.

The support of (X, Y) is $\mathcal{A} = R^2$, the entire plane.
Since the pair (e^X, e^Y) can be any two positive values, the support of (U, V) is $\mathcal{B} = \{(u, v) : u > b, v > d\} = (b, \infty) \times (d, \infty)$.
The inverse transformation is:

$$X = \log((U - b)/a) = \log(U - b) - \log(a)\,, \quad Y = \log((V - d)/c) = \log(V - d) - \log(c)$$

The Jacobian of the inverse transformation is:

$$J = \begin{vmatrix} \frac{1}{u-b} & 0 \\ 0 & \frac{1}{v-d} \end{vmatrix} = \frac{1}{(u-b)(v-d)} > 0 \quad \textit{for } (u, v) \in \mathcal{B}.$$

Therefore

$$f_{U,V}(u, v) = \frac{f\left(\log\left(\frac{u-b}{a}\right), \log\left(\frac{v-d}{c}\right)\right)}{(u-b)(v-d)} \quad \textit{for } u > b, v > d.$$

Problem 6. Suppose that X has density $f_X(x) = 2x$ for $0 < x < 1$ and $Y|X$ has a binomial distribution with n trials and success probability X.

This is a modification of exercise 4.31. The solution is essentially the same.

(a) (5%) Find EY.

$EY = EE(Y|X) = E(nX) = nE(X) = n \int_0^1 x \cdot 2x \, dx = \frac{2n}{3}$

[**Problem 6 continued**]

(b) (7%) Find $\mathrm{Var}(Y)$.

$$
\begin{aligned}
Var(Y) &= Var(E(Y|X)) + E\,Var(Y|X) \\
&= Var(nX) + E(nX(1-X)) \\
&= n^2\,Var(X) + nE(X(1-X)) \\
&= n^2\left\{EX^2 - (EX)^2\right\} + n\int_0^1 x(1-x)\cdot 2x\,dx \\
&= n^2\left\{\int_0^1 x^2\cdot 2x\,dx + \left(\frac{2}{3}\right)^2\right\} + 2n\int_0^1 x^2(1-x)\,dx \\
&= n^2\left\{\frac{1}{2} - \frac{4}{9}\right\} + 2n\left(\frac{1}{3} - \frac{1}{4}\right) \\
&= \frac{n^2}{18} + \frac{n}{6}
\end{aligned}
$$

$Var(X)$ may also be found by noting that $X \sim Beta(2,1)$ and using the formula for the variance of the Beta distribution given in the appendix.
The calculation of $E(X(1-X))$ may also be done using the Beta integral.

If students use the Beta integral in their calculations, they might leave Gamma functions in their answer. That is OK.

[**Problem 6 continued**]

(c) (6%) Find the marginal mass function (pmf) of Y.

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_{Y|X}(y|x) = 2x\, I_{(0,1)}(x) \binom{n}{y} x^y (1-x)^{n-y} I_{\{0,1,\ldots,n\}}(y) \\ &= 2\binom{n}{y} x^{y+1}(1-x)^{n-y} I_{(0,1)}(x) I_{\{0,1,\ldots,n\}}(y) \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x)\, dx = 2\binom{n}{y} \int_0^1 x^{y+1}(1-x)^{n-y}\, dx = 2\binom{n}{y} B(y+2, n-y+1) \\ &= 2\binom{n}{y} \frac{\Gamma(y+2)\Gamma(n-y+1)}{\Gamma(n+3)} = \frac{2 \cdot n!(y+1)!(n-y)!}{y!(n-y)!(n+2)!} \\ &= \frac{2(y+1)}{(n+2)(n+1)} \quad \textit{for } y = 0, 1, \ldots, n\,. \end{aligned}$$

The support $0, 1, \ldots, n$ *must be specified in some way to get full credit.*

Students may leave Gamma functions in their answer and still receive full credit. They don't have to simplify it down to the final form given above.

Problem 7. (12%) Let X_i, $i = 1, 2$, be independent Gamma(α_i,1) random variables. Find the density (pdf) of $Z = X_1/X_2$.

Students can use formula (2) on page 16 of notes10.pdf to answer this question, if they happen to remember it. (They don't have to re-derive it.) If they don't remember it, they can re-derive it following the argument on page 17 of notes10.pdf. This problem is similar to exercise 4.19(b), but requires a different bivariate transformation. The calculations are similar to those in 4.19(b).

For those who remember formula (2) on page 16 of notes10.pdf, you simply plug into the formula using $Z = Y/X$ with $Y = X_1$ and $X = X_2$ leading to:

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} |x| f_{X,Y}(x, zx)\, dx = \int_{-\infty}^{\infty} |x| f_{X_2,X_1}(x, zx)\, dx \\
&= \int_0^{\infty} x \cdot \frac{x^{\alpha_2-1}e^{-x}}{\Gamma(\alpha_2)} \frac{(zx)^{\alpha_1-1}e^{-zx}}{\Gamma(\alpha_1)}\, dx \\
&= \frac{z^{\alpha_1-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^{\infty} x^{\alpha_1+\alpha_2-1} e^{-(1+z)x}\, dx \\
&\quad (\textit{Let } u = (1+z)x.) \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}} \int_0^{\infty} \frac{u^{\alpha_1+\alpha_2-1}e^{-u}}{\Gamma(\alpha_1+\alpha_2)}\, du \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \cdot \frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}}, \quad z > 0.
\end{aligned}
$$

We know in advance that the support must be $(0, \infty)$ since $Z = X_1/X_2$ can be any positive value.

Problem 8. Suppose (X, Y) have the joint density (pdf)

$$f(x, y) = \frac{\sqrt{3}}{2\pi} \exp\left\{-\left(x^2 - xy + y^2\right)\right\}$$

This is a bivariate normal density with $\mu_X = \mu_Y = 0$, $\sigma_X^2 = \sigma_Y^2 = 2/3$, and $\rho = 1/2$.

(a) (10%) Calculate $f_{Y|X}(y|x)$, the conditional density of Y given $X = x$.

This is a special case of exercise 4.45(b).

We know that $f_{Y|X}(y|x)$ is a density function in y for any fixed value of x. Think of $f_{Y|X}(y|x)$ as a function of y with x being a fixed value. Any quantity not depending on y is a constant. We may ignore multiplicative constants in our work since, in the end, the conditional density must integrate to 1, and this condition will determine the appropriate normalizing constant.

$$\begin{aligned}
f_{Y|X}(y|x) &= \frac{f(x,y)}{f_X(x)} \propto f(x,y) \propto \exp\left\{-\left(x^2 - xy + y^2\right)\right\} \\
&\propto \exp\left\{-\left(y^2 - xy\right)\right\} \\
&= \exp\left\{-(y - x/2)^2 + x^2/4\right\} \\
&\propto \exp\left\{-(y - x/2)^2\right\} \\
&\propto \frac{1}{\sqrt{2\pi(1/2)}} \exp\left\{-\frac{(y - x/2)^2}{2(1/2)}\right\}
\end{aligned}$$

The above is the correct answer to part (a).

For part (b), we note the following:

$= N(\mu = x/2, \sigma^2 = 1/2)$ *density*

The answer to part (a) can be written in many algebraically equivalent ways, all of which are correct:

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi(1/2)}} \exp\left\{-\frac{(y - x/2)^2}{2(1/2)}\right\} = \frac{1}{\sqrt{\pi}} \exp\left\{-(y - x/2)^2\right\} = \frac{1}{\sqrt{\pi}} \exp\left\{-y^2 + xy - x^2/4\right\}.$$

(b) (4%) What is the conditional distribution of Y given $X = x$? Name the distribution whose density you found above, and give the values of any parameters. (No work is required.)

The answer is Normal(mean $= x/2$, variance $= 1/2$).

Give two points for stating "Normal", and one point each for the correct values of μ and σ^2. Students should get zero points for stating "bivariate normal", which is very wrong.