TEST #2
STA 5326 **Name:**________________________________
October 30, 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 **12** pages.
- There are a total of **100** points.

Problem 1. An urn contains R red balls and G green balls. A player randomly selects **10** balls, doing this one by one and with**OUT** replacement. (Assume $R+G \geq 10$.) The player wins a dollar every time he selects three red balls in a row. To be more precise, a player receives a dollar after the i-th draw, if he selected red balls on draws $i-2$, $i-1$, and i. Note that under these rules, a player drawing 4 red balls in a row receives a total of 2 dollars; a player drawing 5 red balls in a row receives a total of 3 dollars, etc. Let X be the player's total winnings.

This is similar to exercise C1 and the lecture example on pages 22 to 27 of notes6.pdf.

(a) (9%) Find EX.

The solution exactly parallels that of C1.
Let $Z_i = I(\textit{red balls on draws } i, i+1, i+2)$ *then* $X = \sum_{i=1}^{8} Z_i$ *so that*

$$\begin{aligned} EX &= \sum_{i=1}^{8} EZ_i \\ &= \sum_{i=1}^{8} P(\textit{red balls on draws } i, i+1, i+2) \\ &= 8P(\textit{red balls on draws } 1,2,3) \\ &= 8\frac{R}{R+G} \cdot \frac{R-1}{R+G-1} \cdot \frac{R-2}{R+G-2} \end{aligned}$$

The assertion $P(\textit{red balls on draws } i, i+1, i+2) = P(\textit{red balls on draws } 1,2,3)$ *is similar to what is proved on pages 22 to 24 of notes6.pdf, and students may use this withOUT giving a proof.*

[**Problem 1 continued**]

(b) (9%) Find EX^2.

Again the solution parallels that of C1, but with the change that

$$\begin{aligned}
&P(\textit{red balls on draws } i, i+1, i+2) = P(\textit{red balls on draws } 1,2,3) \\
&\quad = \frac{R}{R+G}\cdot\frac{R-1}{R+G-1}\cdot\frac{R-2}{R+G-2} \equiv p_3 \\
&P(\textit{red on } i, i+1, i+2, i+3) = P(\textit{red on } 1,2,3,4) \\
&\quad = \frac{R}{R+G}\cdot\frac{R-1}{R+G-1}\cdot\frac{R-2}{R+G-2}\cdot\frac{R-3}{R+G-3} \equiv p_4 \\
&P(\textit{red on } i, i+1, i+2, i+3, i+4) = P(\textit{red on } 1,2,3,4,5) \\
&\quad = \frac{R}{R+G}\cdot\frac{R-1}{R+G-1}\cdot\frac{R-2}{R+G-2}\cdot\frac{R-3}{R+G-3}\cdot\frac{R-4}{R+G-4} \equiv p_5 \\
&P(\textit{red on } i, i+1, i+2, j, j+1, j+2) = P(\textit{red on } 1,2,3,4,5,6) \\
&\quad = \frac{R}{R+G}\cdot\frac{R-1}{R+G-1}\cdot\frac{R-2}{R+G-2}\cdot\frac{R-3}{R+G-3}\cdot\frac{R-4}{R+G-4}\cdot\frac{R-5}{R+G-5} \equiv p_6 \\
&\quad \textit{for } |i-j| \geq 3
\end{aligned}$$

(introducing p_3, p_4, p_5, p_6 *for convenience below) which students may use without proof.*

Using the above and counting the number of terms of the various types exactly as in C1 leads to the answer:

$$EX^2 = 8p_3 + (7\cdot 2)p_4 + (6\cdot 2)p_5 + (64-8-14-12)p_6$$

Problem 2.

(a) (9%) Prove the "two-way" rule for expectation (also known as the "law of the unconscious statistician") which says that $Eg(X) = EY$ where $Y = g(X)$. Assume that $g(x)$ is a strictly increasing differentiable function and that X has a density.

This is exercise 2.21.

[**Problem 2 continued**]

(b) (8%) Consider a random right triangle with vertices $(0, 0)$, $(b, 0)$, and (b, Y) where $b > 0$ is a constant and Y is random. This triangle is constructed so that the angle at the origin (call it X) is a random variable with density

$$f_X(x) = \cos(x) \quad \text{for } 0 < x < \pi/2.$$

Here Y is the height of the random triangle determined by the angle X. Use the "two-way" rule to find EY.

This is somewhat similar to Exercise 2.12.
Solution: $Y = b \tan X$ *so that* $EY = E(b \tan X) = \int_0^{\pi/2} (b \tan x) \cos(x)\, dx = b \int_0^{\pi/2} \sin x\, dx = b\,(-\cos x)\big|_0^{\pi/2} = b\,.$

Problem 3. (16%) Al and Bob each roll a fair six-sided die 100,000 times and count the number of 1's they observe. Use a normal approximation to find the probability they both roll exactly the same number of 1's.

Show the argument and do NOT just quote a result from lecture. For full credit, your answer must be a **numerical value** and you must obtain your answer **withOUT using normal tables or your calculator's Φ function**!

This is similar to exercise C5 and the lecture discussion on pages 24-28 of notes7.pdf. The final answer may be obtained by plugging $n =$ 100,000 and $p = 1/6$ into the formula $1/\sqrt{4\pi np(1-p)}$ on page 27 of notes7.pdf. But students should lose a lot of credit if that is all they do; they are supposed to give the argument leading to this too. Note: Tables will not give the answer for this problem since the z-score (ε in the lecture notes) is too small.

[**Problem 3 continued**]

[Additional work space if needed]

Problem 4. Let X have a discrete uniform distribution on N_0 to N_1 that puts equal probability on each of the integers $N_0, N_0 + 1, \ldots, N_1 - 1, N_1$.

This is exercise 3.1. See also Problem #4 on Test #2 from 2012.

Note: Students are allowed to use the appendix which gives the answer for the special case of the discrete uniform on 1 *to* n. *So the easiest solution is just to reduce the general case* N_0 *to* N_1 *to the special case* 1 *to* n *as in the solution of Problem #4 on Test #2 from 2012.*

(a) (8%) Find EX

[**Problem 4 continued**]

(b) (8%) Find $\mathrm{Var}(X)$.

Problem 5. (16%) Suppose the random variables Y and $X_1, X_2, X_3, \ldots$ satisfy

$$Y \sim \text{Exponential}(1) \quad \text{and}$$
$$X_n \sim \text{Geometric}\left(\frac{1}{n}\right) \quad \text{for all } n.$$

Use moment generating functions (mgf's) to prove that

$$\frac{X_n}{n} \xrightarrow{d} Y \quad (\text{as } n \to \infty).$$

See pages 19-22 of notes5.pdf.

[**Problem 5 continued**]

[Additional work space if needed]

Problem 6.

(a) (4%) Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with cumulative distribution function (cdf) F. State a formula for the **cdf** of $Y = \max_{1 \le i \le n} X_i$. (No work is required.)

See page 7 of notes7.pdf. The answer is $F_Y(t) = (F(t))^n$.

(b) (3%) Now take $n = 3$ and assume X_1, X_2, X_3 are i.i.d. Exponential(β) random variables. Apply the result of part (a) to find the **cdf** of $Y = \max_{1 \le i \le 3} X_i$. (No work is required.)

See page 7 of notes7.pdf. The answer is $F_Y(t) = (1 - e^{-t/\beta})^3$ *for* $t \ge 0$.

(c) (10%) Use the result of part (b) to find EY. **Show your work.** (Do NOT use the heuristic argument based on the memoryless property. If you wish, you may use a certain result from homework without proof.)

See page 13 of notes7.pdf for a messy and sketchy solution.
The calculation can be done in either of two ways:

(1): $EY = \int_0^\infty (1 - F_Y(t))\, dt$ *(This is the result of a homework exercise which the students may use without proof.) This calculation is sketched on page 13 of notes7.pdf.*

(2): Differentiate the cdf to find the pdf $f_Y(y)$ *and then calculate* $\int y f_Y(y)\, dy$.

For either approach, it is probably easiest to expand the power and integrate term by term. But there might be another way.