Homework #04 • MATH495/STAT490 • Continuous Conditioning

- submit your write-up before 12 noon on **Thursday** 09 October.
- page limits will be enforced.
- highlight major results.
- **to aid the grader, please begin each lettered problem on a new page**.

A) Two Random Variables (4 pages + plot, 10 pts) Based on problem #16 from Chapter 3 of Ross. If you can explain why (in a sentence), you may set $\mu_x = \mu_y = 0$ as they are just a distraction. You will find the following integral identity to be useful for part (a)

$$\int_{-\infty}^{+\infty} e^{-(az^2+bz+c)}\, dz = \sqrt{\frac{\pi}{a}}\; e^{b^2/4a}\; e^{-c}\ .$$

Add a part (c), which is to verify the covariance formula. You might then wish for two other integrals. Fortunately, if you take first and second derivatives of the above with respect to b, your wish is my command! Please put some thought into how you can organize and explain your steps & not present pages of just algebra. You may, of course, do the details elsewhere but the key logic must be made clear..

Finally, run the matlab script *w05bivar.m* and print out the scatterplots. Annotate this page to illustrate the above theoretical results.

B) How Many Till The Decrease? (3 pages, 10 pts) Problem #81 from Chapter 3 of Ross. The first parts are subtle in that you can get the correct answer through faulty mathematics. Your explanation will be crucial here. An additional part

(a′) Find a value of x such that the value of $f(x)$ is obvious. Use this value to check your result from (a).

will help you to see things more clearly.

Bonus #1: Add a new menu choice to *w03distr.m* to verify your result.

Bonus #2: What differential equation determines $E[N^2]$? Solve this equation and hence, obtain the variance.

C) Pre-election Polls (3 pages, 10 pts) Take a quick scan of the two online news articles on the class webpage – note the occurances of the phrases *accurate* and *margin of error*. The point of this problem is to design a simulation which demonstrates the tagline, *results are considered accurate within X percentage points, 95 times out of 100.*

Say, in a population of voters, 64% will vote 1 and the rest will vote 0. When a single random voter (x_j) is polled, the result to a polltaker is a Bernoulli trial with probabilities determined by the yet unknown fractions of voters. Of course, after the election it will be clear that the Bernoulli trials will have had $p = 0.64$ for $x_j = 1$ and $q = 1 - p = 0.36$ for $x_j = 0$. Recall that a Bernoulli trial has mean $\mu = p$ and variance $\sigma^2 = pq$. A sample poll of n voters is the random variable

$$\tilde{p} = \frac{x_1 + \ldots + x_n}{n}$$

which according to the central limit theorem has mean μ and variance σ^2/n.

How large does n have to be so that with 95% probability the random variable $\tilde{p}$ is within 3 percentage points of the true value p? (Hint: over a wide range of Bernoulli trials $\sigma^2 = pq \approx 1/4$.) Demonstrate the validity of your result by modifying the script *w05vote.m* (which generates a million happy voting citizens) to perform 100 sample polls.

Comment on the correctness of the Evi Mustel quote in the *margin of error* article.