University of Illinois **Fall 2012**

ECE 313: Problem Set 13

Law of large numbers, central limit theorem, and joint Gaussian distribution

Due:	Wednesday, December 12, at 4 p.m.
Reading:	*ECE 313 Notes* Sections 4.10 & 4.11.
Reminder:	The final exam will be held 17, December 8-11 a.m. See http://courses.engr.illinois.edu/ece313/Exams.html for details.

1. [**Achieving potential in a class**]
(The following is roughly based on the ECE 313 grading scheme, ignoring homework scores and the effect of partial credit.) Consider a class in which grades are based entirely on midterm and final exams. In all, the exams have 100 separate parts, each worth 5 points. A student scoring at least 85%, or 425 points in total, is guaranteed an A score. Throughout this problem, consider a particular student, who, based on performance in other courses and amount of effort put into the class, is estimated to have a 90% chance to complete any particular part correctly. Problem parts not completed correctly receive zero credit.

   (a) Assume that the scores on different parts are independent. Based on the LLN, about what total score for the semester are we likely to see?

   (b) Under the same assumptions in part (a), using the CLT (without the continuity correction, to be definite) calculate the approximate probability the student scores enough points for a guaranteed A score.

   (c) Consider the following variation of the assumptions in part (a). The problem parts for exams during the semester are grouped into problems of four parts each, and if X_i and X_j are the scores received on two different parts of the same problem, then the correlation coefficient is $\rho_{X_i,X_j} = 0.8$. The scores for different problems are independent. The total score for problem one, for example, could be written as $Y_1 = X_1 + X_2 + X_3 + X_4$, where X_i is the score for the i^{th} part of the problem. Find the mean and variance of Y_1.

   (d) Continuing with the assumptions of part (c), the student takes a total of 25 problems in the exams during the semester (with four parts per exam). Using the CLT (without the continuity correction, to be definite) calculate the approximate probability the student scores enough points for a guaranteed A score.

2. [**The CLT and the Poisson distribution**]
Let X be a Poisson random variable with parameter $\lambda = 10$, and let $\widetilde{X}$ be a Gaussian random variable with the same mean and variance as X.

   (a) Explain why X can be written as the sum of ten independent, identically distributed random variables. What is the distribution of those random variables? (Hint: Think about Poisson processes. By the CLT, we thus expect that the CDFs of X and $\widetilde{X}$ are approximately equal.)

   (b) Compute $p_X(12)$, the pmf of X evaluated at 12.

   (c) Compute $P\{11.5 \le \widetilde{X} \le 12.5\}$, which, by the CLT with the continuity correction, we expect to be fairly close to the answer of part (b).

   (d) Compute $f_{\widetilde{X}}(12)$, the pdf of $\widetilde{X}$ evaluated at 12. (We expect the answer to be fairly close to the answer of part (b).)

3. **[Gaussian approximation for confidence intervals]**
Recall that if X has the binomial distribution with parameters n and p, the Chebychev inequality implies that if $\widehat{p} = \frac{X}{n}$, then p is contained in the (random) confidence interval $\left[\widehat{p} - \frac{a}{2\sqrt{n}}, \widehat{p} + \frac{a}{2\sqrt{n}}\right]$ with probability at least $1 - \frac{1}{a^2}$. A less conservative, commonly used approach is to note that by the CLT,

$$P\{|X - np| \geq a\sigma\} \approx 2Q(a).$$

   (a) Calculate the value of n sufficiently large that, by the Chebychev inequality, the random interval $[\widehat{p} - 0.1, \widehat{p} + 0.1]$ contains the true value p with probability at least 99%.

   (b) Calculate the value of n sufficiently large that, according to the CLT, the approximate probability the same random interval contains the true value p is at least 99%. Explain your reasoning.

4. **[Transforming joint Gaussians to independent random variables]**
Suppose X and Y are jointly Gaussian such that X is $N(0, 9)$, Y is $N(0, 4)$, and the correlation coefficient is denoted by ρ. The solutions to the questions below may depend on ρ and may fail to exist for some values of ρ.

   (a) For what value(s) of a is X independent of $X + aY$?

   (b) For what value(s) of b is $X + Y$ independent of $X - bY$?

   (c) For what value(s) of c is $X + cY$ independent of $X - cY$?

   (d) For what value(s) of d is $X + dY$ independent of $(X - dY)^3$?

5. **[Estimation of jointly Gaussian random variables]**
Suppose X and Y are jointly Gaussian random variables such that X is $N(4, 16)$, Y is $N(5, 25)$. and $\rho = 0.4$. Let $Z = X + 4Y - 1$.

   (a) Find $E[Z]$ and $\mathrm{Var}(Z)$.

   (b) Calculate the numerical value of $P\{Z \geq 40\}$.

   (c) Find the unconstrained estimator $g^*(Z)$ of Y based on Z with the minimum MSE, and find the resulting MSE.

6. **[Estimating the third power of a jointly Gaussian random variable]**

   (a) Suppose $Z = \mu + W$ where W is a $N(0, \sigma^2)$ random variable. Express $E[Z^3]$ in terms of μ and σ^2. (Note that Z is a $N(\mu, \sigma^2)$ random variable, and you are finding an expression for the moment of such a random variable.)

   (b) Suppose X and Y are jointly Gaussian such that X and Y are each standard normal, and the correlation coefficient between X and Y is ρ. Describe the marginal distribution of Y given $X = u$.

   (c) Under the same assumptions as (b), express $E[Y^3|X = u]$ in terms of ρ and u.