University of Illinois Spring 2008

ECE 313: Problem Set 5
Discrete Random Variables; Estimation

This Problem Set contains seven problems

Due:	Wednesday February 20 at the beginning of class.
Reading:	**Chapter 4; then Chapter 3.** Powerpoint slides: Lectures 8-11
Noncredit Exercises	DO NOT turn these in. **Chapter 4:** 34, 35, 38, 39, 40-43, 48, 51-59; Theoretical Exercises 16-18; Self-Test Problems 9, 13, 15, 16.

1. Let $\mathcal{X}$ denote a binomial random variable with parameters (N, p).

   (a) Show that $\mathcal{Y} = N - \mathcal{X}$ is a binomial random variable with parameters $(N, 1-p)$.

   (b) What is $P\{\mathcal{X} \text{ is even}\}$? Hint: Use the binomial theorem (Ross, page 8) to write an expression for $(x+y)^n + (-x+y)^n$ and then set $x = p$, $y = 1-p$.

2. Let $\mathcal{Y}$ denote a Poisson random variable with parameter λ.

   (a) Show that $P\{\mathcal{Y} \text{ is even}\} = \exp(-\lambda)\cosh(\lambda)$.

   (b) In Problem 1 above, you proved (I hope!) that the probability that a binomial random variable $\mathcal{X}$ with parameters (N, p) is *even* is $[1 + (1-2p)^N]/2$. Now, for large N and small p, the binomial probability $P\{\mathcal{X} = k\}$ is well approximated by the Poisson probability $\exp(-Np)(Np)^k/k!$. Under the same conditions, show that $[1 + (1-2p)^N]/2 \approx \exp(-Np)\cosh(Np)$ and thus your answer of part (a) is consistent with the previous result.

   (c) Now suppose that the value of λ is unknown. The experiment is performed and it is observed that $\mathcal{Y} = k$. What is the *likelihood* of this observation? What is the *maximum likelihood* estimate $\hat{\lambda}$ of λ? That is, what choice of positive number $\hat{\lambda}$ maximizes the likelihood of the observation $\mathcal{Y} = k$?

3. Suppose that 105 passengers hold reservations for a 100-passenger flight from Chicago to Champaign. The number of passengers who show up at the gate can be modeled as a binomial random variable $\mathcal{X}$ with parameters $(105, 0.9)$.

   (a) On average, how many passengers show up at the gate?

   (b) If $\mathcal{X} \leq 100$, everyone who shows up gets to board the flight. Find $P\{\mathcal{X} \leq 100\}$.

   (c) Explain why the number of *no-shows* can be modeled as a binomial random variable $\mathcal{Y}$ with parameters $(105, 0.1)$.

   (d) Notice that the probability that everyone who shows up gets to go can also be expressed as $P\{\mathcal{Y} \geq 5\}$. Use the *Poisson approximation* to compute $P\{\mathcal{Y} \geq 5\}$ and compare your answer to the "more exact" answer that you found in part (b).

4. ["I am from Iowa; I only work in outer space ..."] Each box of Cornies, the breakfast of silver medalists, contains either a picture of Homer Simpson or a picture of Bart Simpson with probabilities $\frac{2}{3}$ and $\frac{1}{3}$ respectively. The contents of each box may be considered to be independent of the contents of other boxes. Little Jimmy T. Kirk of Cedar Rapids, Iowa, asks his mother to buy boxes of Cornies until he has accumulated at least one picture of both Homer and Bart.

   (a) What is the minimum number of boxes that Mrs Kirk must purchase?

   (b) Let $\mathcal{X}$ denote the number of boxes that Mrs Kirk buys till Jimmy has his heart's desire. What is the pmf of $\mathcal{X}$? What is the expected value of $\mathcal{X}$?

(c) The following year, pictures of Harold and Kumar replace those of Homer and Bart in boxes of Cornies. Being even more spoiled than before, Jimmy wants to have at least *two* pictures of each. Repeat parts (a) and (b) for these conditions.

5. A long message is divided into L packets of N bits each (including headers, addresses, timestamps, data bits, CRC bits, tail, flags etc.) and transmitted over a channel with bit error probability p. If the CRC detects that a packet is received in error, the packet transmission is repeated. *But,* if a packet has been transmitted 5 times and still has not been received correctly on the fifth try, then it is deemed to be lost and is not transmitted again.

   (a) What is the probability that the CRC indicates no error in a received packet?

   (b) What is the probability that a packet is transmitted successfully (i.e. is not deemed to be lost)?

   (c) Let $\mathcal{X}_i$ denote the number of times that the i-th packet is transmiitted. What is the pmf of $\mathcal{X}_i$? What is $\mathsf{E}[\mathcal{X}_i]$?

   (d) What is the probability that none of the L packets are lost?

6. There are N multiple-choice questions (with 5 possible answers each) on a certain exam. A student knows the answers to K questions and answers them correctly. On the remaining $N-K$ questions, the student guesses randomly among the 5 choices. The examiner knows N, and can observe the values of $\mathcal{C}$, the number of *c*orrect answers, and $\mathcal{W} = N - \mathcal{C}$, the number of *w*rong answers on the answer sheet. Note that $\mathcal{C}$ can have values K, $K+1$, $\ldots$, N. What the examiner is really interested in, though, is *estimating* the value of K.

   (a) Explain why it is reasonable to model $\mathcal{W}$ as a binomial random variable with parameters $(N-K, 0.8)$. What assumptions are you making?

   (b) Suppose that n answers are incorrect, that is, $\mathcal{W} = n$ and $\mathcal{C} = N - n$. What is the *likelihood* of this observation? Hint: your answer will depend on N, n and the unknown parameter K that the examiner is interested in estimating.

   (c) Having observed that $\mathcal{W} = n$, the examiner is sure that K cannot exceed $N-n$, i.e., K can have value $0, 1, 2, \ldots, N-n$ only. Use the method of Proposition 6.1 of Chapter 4 in Ross to show that the likelihood you found in part (b) is maximized at $\hat{K} = \lfloor N - 1.25n + 1 \rfloor$.

   (d) Since $\mathcal{C} = N - n$, a *guessing penalty* is applied by subtracting $\lfloor 0.25n \rfloor$ from $\mathcal{C}$ to get an estimate of K. For $N = 100$ and $K = 90$, compare the *examiner's estimate* $\tilde{K} = N - n - \lfloor 0.25n \rfloor$ and the maximum likelihood estimate $\hat{K}$ for each possible value that n can take on, *viz.* $n = 0, 1, \ldots 10$. Notice that lucky guesses cause the examiner to overestimate K while the unlucky student who blows all ten problems has to suffer the further indignity of having the score reduced to something smaller than K.

   (e) [Noncredit exercise] If think that the result of part (d) is grossly unfair, write a letter to the Educational Testing Service complaining about the guessing penalty.

7. An urn contains 10 red balls and an unknown number x of blue balls. The experiment consists of drawing one ball at random from the urn and noting its color. Consider 100 independent trials of this experiment. Thus, the ball drawn is replaced, and the urn shaken well before the next ball is drawn. It is observed that 25 of the drawings resulted in a red ball and 75 in a blue ball.

   (a) What is the maximum likelihood estimate $\hat{x}$ of the number of blue balls x in the urn?

   (b) Let $\hat{p} = \dfrac{10}{10+\hat{x}}$. With what level of confidence can we say that $p = \dfrac{10}{10+x}$ lies in the interval $[\hat{p} - 0.1, \hat{p} + 0.1]$?

   (c) Find a confidence interval with confidence level 0.96 for p .