Advanced Statistics

Homework 1

Due: Oct. 24,2012

Jin-Lung Lin

1. Define and explain the following terms
   (a) sample space
   (b) random variable
   (c) conditional probability
   (d) distribution function
   (e) independence
   (f) Bayes's Rule

2. A study of the post treatment behavior of a large number of drug abusers suggest that the likelihood of conviction within a 2-year period after treatment may depend upon the offenders education. The proportions of the total number of cases falling in four education-conviction categories in the following table.

	Status Within Two Years After Treatment		
Education	Convicted	Not Convicted	Totals
10 years or more	.10	.30	.40
9 years or less	.27	.33	.60
Totals	.37	.63	1.00

Suppose that a single offender is selected from the treatment program. Define the events:
A: The offenders has 10 or more years of education.
B: The offenders is convicted within 2 years after completion of treatment.
Find: a. P(A), b. P(B), c. P(AB), d. P(A∪B), e. P($\bar{A}$), f. P($\overline{A \cup B}$), g. P($\overline{AB}$), h. P(A|B),i. P(B|A)

3. (2.99) A diagnostic test for a disease is said to be 90$\%$ accurate in that if person has the disease, the test will detect it with probability .9. Also, if a person does not have the disease,

the test will report that he or she does not have it with probability .9. Only 1$\%$ of the population has the disease in question. If a person is chosen at random from the population and the diagnostic test indicates that she has the disease, what is the conditional probability that she does, in fact, have the disease? Are you surprised by the answer? Would you call this diagnostic test reliable?

4. (3.59) The probability of a customer arrival at a grocery service counter in any 1 second is equal to .1. Assume that customer arrive in a random stream and hence that an arrival in any 1 second is independent of all others.
   a. Find the probability that the first arrival will occur during the third 1-second interval.
   b. Find the probability that the first arrival will not occur until at least the third 1-second interval. c. Apply the result in (b) for the case r=7, p=.5 to determine the values of y for which $p(y) > p(y-1)$.

5. (3.98) Customers arrive at a checkout counter in a department store according to a Poisson distribution at an average of seven per hour. During a given hour, what are the probabilities that
   a. no more than three customers arrive?
   b. at least two customers arrive?
   c. exactly five customers arrive?

6. (3.133) The exercise demonstrates that, in general, the results provided by Tchebysheff's theorem cannot be improved upon. Let Y be a random variable such that

$$p(-1) = 1/18, p(0) = 16/18, \text{ and } p(1) = 1/18$$

   a. Show that E(Y)=0,and V(Y)=1/9.
   b. Use the probability distribution of Y to calculate $P(|Y-\mu| \geq 3\sigma)$. Compare this exact probability with the upper bound provided by Tchebysheff's theorem to see that bound provided by Tchebysheff's theorem is actually attained when k=3.
   c. In (b) we guaranteed E(Y)=0 by placing all probability mass on the values -1, 0, and 1, with $p(-1) = p(1)$. The variance was controlled by the probabilities assigned to $p(-1)$ and $p(1)$. Using this same basic idea, construct a probability distribution for a random variable X that will yield $P(|X-\mu_X| \geq 2\sigma_X)$=1/4.
   d. If any $k > 1$ is specified, how can a random variable W be constructed so that $P(|W-\mu_W| \geq k\sigma_W) = 1/k^2$?