Note: The answers are recorded after all the questions. KLW

Stat 280 Mid-term test Oct 28, 1999

Instructions: The test is open book. Attempt all questions. You have 75 minutes for this test - the marks assigned add to 75.

1. (8 marks) Explain why (no calculations please):
(The original formula was hand-written - I try to indicate here what it was)
(n-1)C(i-1,j,n-i-j) +(n-1)C(i,j-1,n-i-j)+(n-1)C(i,j,n-i-j-1)
=nC(i,j,n-i-j)
where nC(a,b,c) means the number of ways n things can be allocated to subgroups of size a, b, and c.

2. (10 marks).

A charity sells 500 raffle tickets. There are two prizes to be awarded; the same ticket is allowed to win both prizes. Two tickets are drawn at random, with replacement, to identify the winners. What is the probability that a customer who has purchased 50 tickets wins at least one prize?

3. (12 marks)

A diagnostic test is used to by a clinic to determine whether or not a person has a certain disease. If the test is positive, the person is assumed to have the disease, while if it is negative, the person is assumed not to have the disease. Suppose 10% of people tested have the disease, 90% of people that have the disease test positive, and 95% of those who are disease-free test negative. What is the probability that the clinic will make a mistaken diagnosis?

4. (10 marks)

Suggest a distribution that would model the variability in the number of students in this class that attend the Tuesday lecture of STAT 280. Assume that each Tuesday's attendance is a random value of an IID random variable. There are 45 students registered in this course.
Explain your choice of model and parameters. Base your choice on one of the named models mentioned in the course text.

5. (15 marks)

Thousands of identical (in appearance) firecrackers are stored in a warehouse. The probability that, when ignited, they make an audible "bang" is a random variable X. X has a Beta distribution with parameters a=1 and b=3. When they do make an audible bang, the loudness is recorded as Y, which has a Gamma distribution with r=3 and lambda = 50.
a) If 100 of these firecrackers are selected at random, what would the histogram of loudness values look like?
b) How would you determine to three decimal places the exact standard deviation of the distribution of loudness for a single randomly selected firecracker?

6. (20 marks)

a) Why do the Central Limit Theorem and the Poisson Process together suggest that a Gamma distribution with large r will be approximately Normally distributed?
b) What normal distribution would you use to approximate, as well as possible, a Beta (a=3,b=5) distribution?
c) In your MINITAB simulation of the Gambler's ruin experiment, predict what the estimated probability would be for p=.45 if the target for the gambler were $20, and the gambler's initial stake was $5. (Assume each game is for $1, as usual).
d) How does one generate values that look like random outcomes of a random variable whose CDF is F(), even if F is not one of the distributions for which your software generates values automatically?

Answers:

1. The right side is the number of allocations of n things to three subgroups of sizes i,j,n-(i+j) respectively

The left side is the sum of three terms, each term is the number of allocations of n-1 things to three particular sized subgroups. The number of allocations of n things can be generated from these three suballocations by considering one particular object of the n objects. It may be allocated to group 1, 2, or 3. If it allocated to group 1, then there are a number of allocations of the n-1 remaining objects is the first term on the left side of the identity. If it is allocated to the second group or third group, then the number of allocations is as given by the second and third terms. Since in the allocations of the n, this object must be somewhere, and since an allocation with it in one group cannot be the same as the allocation with it in another group, the identity follows.

2. Let W1 be the event that there is a winner in the 50 tickets on the first draw
Let W2 ...
P(W1 union W2) = 1- P(W1 intersection W2) = 1-P(W1)P(W2) = 1-(50/450)^2=.19

3. D: has disease T: positive on test
Given P(D)=.1 P(T|D)=.9 P(T'|D')=.95

We want P(mistaken diagnosis) = P(D and T') + P(D' and T)

But P(D and T') = P(T'|D)P(D)=(1-.9) x .1 = .01
and P(D' and T) = P(T|D')P(D')=(1-.95) x (1-.1) = .045

So P(mistake)= .010 + .045 = .055

4. Beta a=25 b=5 produces the right shape. Note mean is 25/30 = .83 so about 83% of the class is present on average and the SD about this mean is sqrt[25*5/(900*31)]=.07 so the percentage could vary from 69% to 97% which seems about right. To get the model in the right scale, one just multiplies the Beta(25,5) by 45. The reason Beta is indicated is that there are natural limits at 0 and 45 for this distribution.
The fact that the upper limit is attainable suggests that a model like normal would not be as good since it has infinite support.

5. The proportion of firecrackers that make an audible sound is Beta(1,3) or about (p +- SD of p) .25 +- .19 (using formulas for mean and SD of Beta)

so we can at least infer than a majority of the firecrackers will emit 0 sound. The rest will have a sound level described by Gamma(3, 50) which has mean .06 and SD about .034. So the histogram should have more than half its area in the cell which includes 0 and a right-skewed distribution for the rest with mean and SD about .06 and .034 resp.

6. a) Gamma(r,lambda) is sum of iid expo(lambda), the waiting time for the rth event in a Poisson Process. By CLT, as r increases, this distribution approaches a normal distribution.

b) The normal distribution with mean 3/8 and variance 15/(64 x 9) = 5/192 =.026

c) Let a=.55/.45

prob of ruin = 1- (1-a^5)/(1-a^20)=1-(1-2.73)/(1-55.3)=1-1.73/54.3 = 1-.032 = .968

d) use the inverse function F^(-1)(ui) to turn the random ui into values of a random variable having F as its cdf.