Statistics 220/301 **Banerjee**
Spring 2001

Quiz 3

May 31

Your section: _____ Print your name: ______________________________

Sign your name: ______________________________

This is an open book exam. However, you are not allowed to pass any material (such as books, notes, or calculators) to each other. The quiz consists of 3 problems. Each problem carries 10 points but the maximum you can score is 25. You have 35 minutes. Good luck!

For the purposes of the quiz, the continuity correction can be ignored.

Problem 1. Two persons do the following experiment a 100 times. Each throws a die independently of the other. If the numbers on the two dice match they record the difference; otherwise they record a 1. Formulate an appropriate box model for the experiment. Hence, find the chance that in 100 repetitions of the experiment the two dice match more than 17 times. (10 points)

Solution: The chance expt. has two outcomes, 0 and 1. A 0 is recorded if the numbers on the two dice match; the chance of the two dice matching is $6/36 = 1/6$ (since there are 6 outcomes favorable to this, $(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)$ out of 36 possible outcomes). So the chance of recording a 0 is $1/6$ and the chance of recording a 1 is $5/6$. Thus the appropriate box is one that has 6 tickets, with 1 ticket marked 0 and 5 tickets marked 1.

The sum of the numbers on the tickets after a 100 draws is precisely the number of times that the two dice do not match. We need to find the chance that the two dice do not match less than $83\,(100 - 17)$ times, or in other words the chance that the sum of the numbers on the tickets after a 100 draws is less than 83. The average of the box is $5/6$ and the S.D. of the box is 0.373 ; so the expected value of the sum is $100 \times 5/6 = 83.3$ (appx) and the S.E. for the sum is $\sqrt{100} \times .373 = 3.73$. The probability histogram for the sum of the draws can be assumed to follow the normal curve and so the required chance is the area under the normal curve to the left of $(83 - 83.3)/3.73 = -.08$ (appx). This can be found by referring to the normal table and is approximately 47 %.

Problem 2. I am tossing a very unfair coin, loaded so heavily in favour of heads that the chances are 3 in a million that it will come up tails on any single toss. I now toss the coin 200 times. Find the chance of at least 1 tails in 200 tosses. What does this tell us about the probability histogram for the number of tails in 200 tosses of this coin ? In particular, would it be well-approximated by the normal curve ? Answer briefly. (10 points)

Solution: The chance of at least one tails in 200 tosses = 1 - chance of no tails in 200 tosses.

Now the chance of no tails in 200 tosses is precisely the chance of 200 heads in 200 tosses $= (1 - .000003)^{200} = 0.999$ (appx). Thus the chance of at least 1 tails in 200 tosses is $1 - .999 = .001$. (appx)

The number of tails in 200 tosses can take any value between 0 and 200. The above computation shows that the chance that the total number of tails is more than or equal to 1 is just .001 and the chance that it is 0 is .999. Thus the probability histogram for the total number of tails is extremely lopsided (assymetric) with almost all the area in the block over 0. Hence it cannot be well-approximated by the normal curve.

This is the situation where the underlying box is very lopsided (a million tickets with 3 marked 1 (coresponding to tails) and the rest marked 0) and hence the normal approximation takes a long time to work.

Problem 3. Two hundred draws are made at random with replacement from the box $[-3, -1, 0, 1, 2]$. The sum of the non-positive numbers will be around (a) give or take (b) or so. Put in appropriate numbers for (a) and (b). Hence find the chance that in two hundred draws the sum of the positive numbers is between -190 and -130. (10 points)

Solution: Since we are worried about the sum of the non-positive numbers, whenever we get a positive number out of the box we take it to be 0 for the purpose of forming a sum. Thus the sum of the non-positive numbers is like the sum of the numbers on the tickets in 200 draws from the box $[-3, -1, 0, 0, 0]$. The average of the box is -0.8 and the SD of the box is 1.17 (appx). Thus the expected value for the sum after 200 draws is $200 \times 0.8 = 160$ and the S.E is $1.17 \times \sqrt{200} = 16.5$. Thus (a) is -160 and (b) is 16.5.

This part is straightforward normal approximation to the probability histogram. We convert -190 and -130 to standard units using (a) and (b). In std. units -130 is $(-130 - (-160))/16.5 = 30/16.5 = 1.8$ (appx) and -190 is -1.8 (appx). Thus we need to find the area under the normal curve between -1.8 and 1.8 and this is approximately 92.8 %.