Math 10A | Fall, 2016

Professor Ken Ribet

Homework due Monday, November 21, 2016

I have selected a few problems from old final exams and homework assignments. This is a fairly light assignment, with a five-day turnaround time.

Before stating the first problem, we will discuss a preliminary problem and its solution. Here's the preliminary problem:

Use the Central Limit Theorem to estimate the probability of getting heads more than 60 times when a fair coin is tossed 100 times.

Here's the idea for the solution:

We consider the random variables X_1, X_2, $\ldots$ X_{100} that concern the various coin tosses: the first, the second, the third and so on. These are independent and identically distributed. They take the value 1 on `H` and the value 0 on `T`, so they have expected value $\frac{1}{2}$. If X is one of the variables, then $\mathrm{Var}[X] = E[X^2] - E[X]^2 = \frac{1}{4}$. (Note that $X^2 = X$ because $1^1 = 1$ and $0^2 = 0$.) Hence X has standard deviation $\sigma = \frac{1}{2}$. Let $\overline{X} = \dfrac{X_1 + X_2 + \cdots + X_{100}}{100}$ be the average of the X_i. Then, as we all saw in class on Tuesday, $\overline{X}$ has mean $\frac{1}{2}$ and standard deviation $\dfrac{\sigma}{\sqrt{100}} = \dfrac{1}{20}$. It follows (as we all discussed on Tuesday) that $Y = 20(X - \frac{1}{2}) = 20X - 10$ has mean 0 and standard deviation 1. To have more than 60 heads is to have $X > 0.6$, i.e., to have $Y > 2$.

Now the Central Limit Theorem states that Y is "like" a standard normal variable. This is a qualitative statement; a slightly more helpful comment is that the analogue of Y approaches a standard normal variable as the number of coin tosses approaches ∞. Of course, 100 is not ∞, but the conceit of the problem is that 100 is big enough to be very much like ∞. Therefore the probability that Y is bigger than 2 is very close to the probability that a standard normal random variable is bigger than 2. Note that 1 is the standard deviation of the standard normal variable, so we are asking for the probability to be more than two standard deviations out from the mean (which is 0). Now in the standard set-up (table on p. 566 of the text), 50% of the scores are between 0 and ∞ and 47.72% of the scores are between 0 and 2. Hence only 2.28% of the scores are bigger than 2. So, all told, our answer to the question is that the Central Limit Theorem estimates the probability of getting more than 60 heads in 100 tosses as 0.0228.

Actually, folks, this is an over-estimate. I calculated the true probability using `Sage` and got 0.0176 as my number. The true probability, by the way, is $\frac{1}{2^{100}} \sum_{i=61}^{100} \binom{100}{i}$, where $\binom{100}{i}$ is the binomial coefficient "100 choose i." Warning: Math 10B talks about these coefficients quite a lot.

1. Use the Central Limit Theorem to decide which of the following two scenarios is the more likely one: getting more than 70 heads in 100 tosses of a fair coin or getting more than 600 heads in 1000 tosses of a fair coin. (Hint: you are trying to see which of two probabilities is the bigger one; you don't have to calculate either of them numerically.)

2. Show that $f(x) = \frac{b}{2} e^{-b|x|}$ is a probability density function when b is a positive real number.

3. Find all real numbers C such that $C \frac{e^x}{1 + e^{2x}}$ is a PDF of a random variable. For each such C, find the corresponding CDF.

4. Suppose that X is a random variable whose PDF is $f(x)$ and that a is a real number. What is the PDF for the random variable $X - a$?

5. Suppose that X is a random variable whose PDF is $f(x)$ and that a is a nonzero real number. What is the PDF for the random variable aX? (We discussed this in class on Tuesday.)

6. Compute the expected value of a random variable whose PDF is 0 for $x < 1$ and $3x^{-4}$ for $x \geq 1$.

The last two problems concern the maximum likelihood method. James will discuss this method in class on Thursday.

7. Suppose that $x_1, x_2, \ldots, x_n$ form a random sample from a distribution with PDF $f(x)$ as in Probem 2. (The parameter b is unknown.) Find the maximum likelihood estimate for b.

8. Bob owns an aquarium with an unknown number of fish. Say there are n fish. We can see that one of the fish is red and all the others are black. Bob behaves in the following bizarre way (remember: this is math homework). First he picks out a fish at random and observes that it is black. He returns this fish to the aquarium and adds eight more black fish (so now there are $n + 8$ fish). Finally, Bob picks out a fish at random and observes that it is red. The question is to use the maximum likelihood method to obtain an estimate for n.