Yet more remarks on Probability A, sheet 1 (11/02/2012)

Minhyong Kim

6. In this problem, we randomly pick a $Q \in [0,1]$ using a uniform distribution, and then start tossing a coin that has

$$P(\text{head}) = Q.$$

The results are rather interesting, but there was just one computation I wished to show that can be useful to the other parts. This is the conditional expectation

$$E(X_i|X_j),$$

for $i \neq j$, where X_i is the function that is 1 if the i-th toss is a head and 0 otherwise. When making statements like that, it is useful to clarify for yourself precisely what the probability space is: In the first approximation, it is the set S_n of all sequences

$$S_n = \{s = (T, H, H, T, H, T, \ldots, H, T)\}$$

of length n made up of H's and T's. The value $X_i(s)$ of the random variable is 1 if s has H in the i-th place, and zero if s has T in the $i-$th place. That is, X_i is the indicator function of the subset of S_n consisting of all sequences that have H in the i-th place. Clearly, $|S_n| = 2^n$. What is the probability measure on S_n? Here is where we have to become more precise, because the probability of any given sequence depends on Q, and we realise that the probability space is actually

$$[0,1] \times S_n = \{(Q, s)\}$$

consisting of pairs, a number in $[0,1]$ and a sequence. The probability measure of $dI \times \{s\}$, where dI is a small interval of length dQ around the point Q, is

$$Q^{h(s)}(1-Q)^{n-h(s)}dQ,$$

where we denote by $h(s)$ the number of heads in the sequence s. In particular, if $Q = 1/2$, then all sequences are equally likely, but not when $Q \neq 1/2$. Because of this dependence on Q, when calculating, for example, the probability that you will get 2 heads, you must integrate a conditional probability

$$P(h = 2) = \int_0^1 P(h = 2|Q)dQ = \int_0^1 \binom{n}{2} Q^2(1-Q)^{n-2}dQ.$$

I'll leave it to you to calculate this integral. One I will do for you is

$$E(X_i) = P(X_i = 1) = \int_0^1 P(X_i = 1|Q)dQ = \int_0^1 QdQ = 1/2,$$

the same as if we had a fair coin.
Now, for the conditional expectation $E(X_i|X_j)$, for $i \neq j$. We just calculate separately,

$$E(X_i|X_j = 0) = P(X_i = 1|X_j = 0) = \frac{P(X_i = 1, X_j = 0)}{P(X_j = 0)}$$

$$= \int_0^1 P(X_i = 1, X_j = 0|Q)dQ/(1/2) = \int_0^1 Q(1-Q)dQ/(1/2) = [(1/2) - (1/3)]/(1/2) = 1/3.$$

and

$$E(X_i|X_j = 1) = P(X_i = 1, X_j = 1)/P(X_j = 1) = \int_0^1 Q^2 dQ/(1/2) = 2/3.$$

The two equations can be conveniently expressed as

$$E(X_i|X_j) = (X_j + 1)/3.$$

If you think about it the conclusion is a bit surprising. It says the probability of getting a head in the i-th place depends on what happens in the j-th place, in fact, quite heavily. Why is that? The reason is that *we do not know* Q, when we look at the results. Hence, if we see a head in place j, it increases the probability that Q is large, and hence, that the i-th place will also be head. The calculation we did merely quantifies this commonsensical observation.

Perhaps this problem has some practical application. Say someone tosses a coin and gets a head. You are invited to bet on the outcome of the second toss. What should you guess? Well, there is always the possibility that the coin is not entirely perfect, that there is a slight bias towards either head or tail. Perhaps there is no perfect coin, in fact. So then, the first head can be interpreted as a bit of evidence that the bias is towards head. So you should bet on head for the second toss as well. For a more extreme example, if there were a sequence of ten consecutive heads, then it's probably a pretty good idea to bet on head for the eleventh toss.
By the way, in practice, $E(X_i|X_j = 1)$ is hardly likely to be as big as in our idealised problem. This is because we were assuming Q is chosen from $[0, 1]$ with a uniform distribution. An actual coin is likely to be quite close to fair, unless the person is deliberately trying to cheat you. What would be a reasonable model for the distribution of Q? It is important to remark that Q is a well-defined quantity in practice. For any given coin we can actually toss the coin 10 million times, say, and calculate

$$Q \approx \text{number of heads}/10^7.$$

We are then asking what the distribution of Q might be as we go through many actual coins. Do you have some idea to model this distribution?

There are many places in literature where related scenarios pop up. I recommend the prologue of Roger Penrose's book 'Shadows of the Mind' as well as the beginning pages of Tom Stoppard's play 'Rosenkrantz and Guildenstern are dead.'