EE 178/278A Probabilistic Systems Analysis

Spring 2014 Tse/Hussami Lecture 12

Midterm Exam Comments

(1) **Balls and bins problem:** The main point in this problem is to understand the logical relationship between events, rather than computing probabilities right away. For example parts b and d are about determining whether events are disjoint.

There are multiple ways of defining the sample space for the experiment. One way is ball centric, and the other bin centric. We start with the ball centric approach, where an outcome describes the location of each ball. An outcome is described by $b_1, b_2, \cdots, b_m$ where b_i describes the bin location of ball i. Therefore, $\Omega = \{(b_1, \cdots, b_m) : b_i = 1, \cdots, n\}$. By symmetry of the model, the probability of any outcome is the same. There are n possible locations for each ball, so the total number of outcomes is equal to $|\Omega| = n^m$ and $\mathbf{P}(\omega) = n^{-m}$. This is the easy way of solving this problem.

However many of you chose the bin centric way. Now instead we describe the number of balls there are in each bin. If we count how many balls there are in each bin, we lose the information about the ball labels. For this problem, the events we are interested in are only about the number of balls in each bin. Therefore, the description of the number of balls in each bin is a valid sample space description for the purpose of our questions, if we are able to give a probability assignment to each outcome. However, each outcome here might have a different probability.

(2) **Joint probability mass function:** Let X have a pmf $p_X(a) = \mathbf{P}(X = a)$ for $a \in \mathbf{A}$, and Y a pmf $p_Y(b) = \mathbf{P}(Y = b), b \in \mathbf{B}$. Is it sufficient to have these two individual mass functions? No. They do not give us information about the relationship between the two random variables X and Y. They could be independent, positively correlated or negatively correlated. Therefore, the relationship should be described by a 2 dimensional function which is the joint pmt:

$$p_{X,Y}(a,b) = \mathbf{P}(X = a, Y = b), a \in \mathbf{A} \text{ and } b \in B.$$

This function corresponds to the tables we drew in the law of large numbers section of lecture 10. The sum of all the entries in part a is equal to 1. The ideal explanation involves going back to the sample space. The key was to notice the events $\{X = a \cap Y = b\}$ are disjoint and cover the whole sample space. Therefore, the events $\{X = a \cap Y = b\}$ form a partition of the sample space. The second part asks you to compute

$$\sum_{a \in \mathbf{A}} p_{X,Y}(a,b).$$

This is equivalent to summing over one row of the matrix, which takes the union of the events in this row. Therefore this event is equivalent to $\{Y = b\}$. So we can conclude that

$$\mathbf{P}(Y = b) = \sum_{a \in \mathbf{A}} p_{X,Y}(a,b).$$

This is called the marginal probability mass function. When two events $\{X = a\}$ and $\{\mathbf{Y} = b\}$ are independent,

$$p_{X,Y}(a,b) = p_X(a)p_Y(b).$$

Two random variables X and Y are independent if

$$\mathbf{P}(X = a \text{ and } Y = b) = \mathbf{P}(X = a)\mathbf{P}(Y = b)\ \forall a \in \mathbf{A}, b \in \mathbf{B}.$$

Any answer that used this strong assumption was not correct, because you were not given that the variables are independent. Finally, to compute the pmf of Z, we think about partitioning the events $\{Z = c\} = \{f(X,Y) = c\}$ into smaller events. For example, consider the event $Z = 1$, and suppose that X and Y are binary random variables and that $Z = f(X,Y) = X + Y$. Then the event $\{Z = 1\}$ can be split into the events $\{X = 1, Y = 0\}$ and $\{X = 0, Y = 1\}$ which are disjoint. Therefore, adding up their probabilities gives the probability $\mathbf{P}(Z = X + Y = 1)$.

Generalizing the above example, the event $\{Z = c\}$ can be split by looking at all the possible (a,b) pairs such that $f(a,b) = c$. Now, $\mathbf{P}(Z = c) = \sum_{(a,b):f(a,b)=c} p_{X,Y}(a,b)$.

(4) **Conditional expectation:** This is a generalization of question number 1 of homework 5. In part a you had to show:

$$\mathbb{E}[X] = \sum_b \mathbb{E}[X|Y = b]\mathbf{P}(Y = b).$$

The expectation of X is given by the sum over terms involving a random variable Y. For each value that Y takes on, we look at the probability of that event happening and multiply it by the conditional expectation of the value of X for that specific value of Y. This is analogous to the law of total probability which says that: $\mathbf{P}(X = a) = \sum_b \mathbf{P}(X = a|Y = b)\mathbf{P}(Y = b)$. Part b shows an application of this rule. You are asked to compute the expected time that it takes James Bond to exit the cell. How do we define Y here so that we can apply the equation in part a? We let Y be the method of escape in the first trial. It is important to define random variables and events clearly. There are 3 possibilities $Y = 0$ if he escapes via the door on the first try, $Y = 1$ if he escapes via the duct on the first try, and $Y = 2$ if he escapes via the pipe on the first try. X is the total duration until exit. How do we compute $\mathbb{E}[Y]$? We compute $\mathbb{E}[X|Y = 0] = 0$, $\mathbb{E}[X|Y = 1] = 2 + \mathbb{E}[X]$ and $\mathbb{E}[X|Y = 2] = 5 + \mathbb{E}[X]$ and then apply the result of part a.

(5) **Genome sampling:** It is important to clarify whether you are defining a random variable or an event. There was a common mistake in question (e), where you were asked for the upper bound on the probability that at least one position on the genome is not covered. Many of you used the Chebyshev inequality. If X_i is defined as being the indicator random variable which is equal to 1 if position i is not covered by a read and 0 otherwise. The Chebyshev inequality can be used to bound $\mathbf{P}(X \geq 1)$ as long as you can compute the variance of $X = \sum_i X_i$. However, the variance of X cannot be easily computed as the sum of the variance, because as you showed in a previous part of this question, the X_i variables are not independent. The correct approach was to use the union bound in this question.