CS 155: Homework 1

Douglas Lanman
dlanman@brown.edu
14 September 2006

Problem 1

Prove Lemma 1.2 in [1] (i.e., the union bound).

For any finite or countably infinite sequence of events $E_1, E_2, \ldots$, the union bound is given by

$$\Pr\left(\bigcup_{i\geq 1} E_i\right) \leq \sum_{i\geq 1} \Pr(E_i).$$

Recall that Lemma 1.1 from [1] states that

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2),$$

for any two events E_1 and E_2. To derive the union bound, we can iteratively apply Lemma 1.1.

$$\begin{aligned}
\Pr(E_1 \cup E_2 \cup E_3) &= \Pr(E_1 \cup E_2) + \Pr(E_3) - \Pr((E_1 \cup E_2) \cap E_3) \\
&= \Pr(E_1) + \Pr(E_2) + \Pr(E_3) - \Pr(E_1 \cap E_2) - \Pr((E_1 \cup E_2) \cap E_3) \\
&= \sum_{i=1}^{3} \Pr(E_i) - \sum_{i=2}^{3} \Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right)
\end{aligned}$$

By induction, we find the following result

$$\Pr\left(\bigcup_{i\geq 1} E_i\right) = \sum_{i\geq 1} \Pr(E_i) - \sum_{i\geq 2} \Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right). \tag{1}$$

To complete our proof, we simply need to demonstrate that the second term in Equation 1 is non-negative.

$$\sum_{i\geq 2} \Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right) \geq 0$$

Note that, for any $i \geq 2$, $\Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right)$ represents the probability that both event E_i occurs and at least one event from $\{E_1, \ldots, E_{i-1}\}$ occurs. If and only if E_i and $\bigcup_{j=1}^{i-1} E_j$ are mutually disjoint will this probability be equal to zero. In this case, the summation equals zero and the union bound holds with equality (since the second term in Equation 1 is zero). Alternatively, if the probability $\Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right) > 0$, then E_i and $\{E_1, \ldots, E_{i-1}\}$ cannot be disjoint. As a result, the union bound will still hold (but without strick equality) since the second term in Equation 1 will be non-zero.

To complete our proof, we must show that, for any set of events $\{E_i\}$, the union bound defines a valid probability function such that $0 \leq \Pr\left(\bigcup_{i\geq 1} E_i\right) \leq 1$. As we have already shown, the maximum value of the union bound corresponds to the case in which the events are mutually disjoint. Since the events are drawn from a sample space Ω and $\Pr(\Omega) = 1$, we must have $\Pr\left(\bigcup_{i\geq 1} E_i\right) \leq 1$. Similarly, to minimize the union bound, we seek to maximize the second term in Equation 1. As defined, we must have $\Pr\left(\left(\bigcup_{j=1}^{i-1} E_j\right) \cap E_i\right) \leq \Pr(E_i)$ such that $\Pr\left(\bigcup_{i\geq 1} E_i\right) \leq \Pr(E_1)$. Since E_1 is also drawn from the sample space Ω we must have $0 \leq \Pr\left(\bigcup_{i\geq 1} E_i\right) \leq 1$. (QED)

Problem 2

Suppose that a fair coin is flipped n times. For $k > 0$, find an upper bound on the probability that there is a sequence of $\log_2 n + k$ consecutive heads.

Let's begin by defining the probability space and events we will analyze. First, let H_i be the event that the i^{th} coin comes up heads. Similarly, let S_i denote the event that $\log_2 n + k$ consecutive coin flips are heads, starting with the i^{th} flip. We can derive an upper bound on the probability p that there is a sequence of $\log_2 n + k$ consecutive heads using the "union bound" we derived in Problem 1. Note that only a single run of $\log_2 n + k$ heads within n flips is required for success. As a result, we can apply the union bound to the sequence of events S_i to obtain

$$p = \Pr\left(\bigcup_{i \in I} S_i\right) \leq \sum_{i \in I} \Pr(S_i),$$

where $I = \{1, 2, \ldots, n - \log_2 n - k + 1\}$. Note that the limits of the summation have been selected to prevent indexing events which do not exist (e.g., S_0 or $S_{n-\log_2 n - k + 2}$).

At this point, we need to determine $\Pr(S_i)$. For any given sequence starting at flip i, each coin toss will be independent of the others (i.e., $\{H_i\}$ are mutually independent). As a result, we can express the desired probability

$$\Pr(S_1) = \Pr\left(\bigcap_{i=1}^{log_2 n + k} H_i\right) = \prod_{i=1}^{log_2 n + k} \Pr(H_i) = \left(\frac{1}{2}\right)^{\log_2 n + k} = \frac{1}{2^k n},$$

since $Pr(H_i) = 1/2$. Similarly, we must also have $\Pr(S_i) = 1/(2^k n)$ for all complete sequences starting at i. Substituting into the previous equation, we obtain an upper bound for p.

$$p \leq \sum_{i=1}^{n - log_2 n - k + 1} \frac{1}{2^k n} = \frac{n - log_2 n - k + 1}{2^k n} \leq 2^{-k}$$

Note that the numerator satisfies $n - \log_2 n - k + 1 \leq n$, since $\log_2 n + k > 0$ for $n > 0$ and $k > 0$. In conclusion, we have derived an upper bound on the probability p that there is at least one sequence of $\log_2 n + k$ consecutive heads in n coin flips given by

$$\boxed{p \leq 2^{-k}}.$$

Problem 3

The following problem is known as the Monty Hall problem, after the host of the game show "Let's Make a Deal". There are three curtains. Behind one curtain is a new car, and behind the other two are goats. The contestant chooses the curtain that she thinks the car is behind. Monty then opens one of the other curtains to show a goat. (If Monty has more than one choice, assume he chooses uniformly at random). The contestant can then stay with the original curtain or switch to the other unopened curtain. Should the contestant switch or not, or does it make no difference?

This is a classic problem in probability and the "counter-intuitive" result can best be seen by applying Bayes' Law. To begin our analysis let's enumerable the sample space. Let O_i correspond to the event where Monty opens door i. In addition, let C_i be the event that the car is behind door i. Without loss of generality we can assume that the contestant always initially chooses the first door and that Monty chooses the second (since we could always permute the door labels to achieve this condition). Subject to this condition, the sample space Ω can be enumerated simply by the position of the car as $\Omega = \{C_1, C_2, C_3\}$.

Recall from [1] that Bayes' Law is given by

$$\Pr(E_j \,|\, B) = \frac{\Pr(E_j \cap B)}{\Pr(B)} = \frac{\Pr(B \,|\, E_j)\Pr(E_j)}{\sum_{i=1}^{n}\Pr(B \,|\, E_i)\Pr(E_i)},$$

where $E_1, E_2, \ldots, E_n$ are mutually disjoint sets such that $\bigcup_{i=1}^{n} E_i = E$. In the context of conditional probabilities, we would like to determine the following quantities (where we have applied Bayes' Law since $\{C_i\}$ are mutually disjoint).

$$\Pr(C_1 \,|\, O_2) = \frac{\Pr(O_2 \,|\, C_1)\Pr(C_1)}{\sum_{i=1}^{3}\Pr(O_2 \,|\, C_i)\Pr(C_i)}, \;\; \Pr(C_3 \,|\, O_2) = \frac{\Pr(O_2 \,|\, C_3)\Pr(C_3)}{\sum_{i=1}^{3}\Pr(O_2 \,|\, C_i)\Pr(C_i)}$$

Note that $\Pr(C_1 \,|\, O_2)$ represents the probability that the car is behind the original door, whereas the $\Pr(C_3 \,|\, O_2)$ corresponds to the probability that the car is behind the remaining door. To solve this problem, we would like to determine which probability is larger (or prove that they are equal).

To evaluate these expressions, we must first determine the simple event and conditional probabilities (which follow directly from the problem statement).

$$\Pr(C_1) = \Pr(C_2) = \Pr(C_3) = \frac{1}{3}$$
$$\Pr(O_2 \,|\, C_1) = \frac{1}{2},\ \Pr(O_2 \,|\, C_2) = 0,\ \Pr(O_2 \,|\, C_3) = 1$$

Subsituting these expressions, we can obtain estimates of the desired probabilities.

$$\Pr(C_1 \,|\, O_2) = \frac{\Pr(O_2 \,|\, C_1)\Pr(C_1)}{\sum_{i=1}^{3}\Pr(O_2 \,|\, C_i)\Pr(C_i)} = \frac{(1/2)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} = \frac{1}{3}$$

$$\Pr(C_3 \,|\, O_2) = \frac{\Pr(O_2 \,|\, C_3)\Pr(C_3)}{\sum_{i=1}^{3}\Pr(O_2 \,|\, C_i)\Pr(C_i)} = \frac{(1)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} = \frac{2}{3}$$

In conclusion, we find that **the contestant should always switch** since $\Pr(C_3 \,|\, O_2) > \Pr(C_1 \,|\, O_2)$. From Bayes' Law we find that the contestant will win the car with probability 2/3 if he switches, whereas he will only win with probability 1/3 by staying with the original door.

Problem 4

A medical company touts its new test for a certain genetic disorder. The false negative rate is small: if you have the disorder, the probability that the test returns a positive result is 0.999. The false positive rate is also small: if you do not have the disorder, the probability that the test returns a positive result is only 0.005. Assume that 2% of the population has the disorder. If a person chosen uniformly from the population is tested and the result comes back positive, what is the probability that the person has the disorder?

This problem can be also solved using Bayes' Law. Recall from the previous problem that Bayes' Law is given by

$$\Pr(E_j \,|\, B) = \frac{\Pr(E_j \cap B)}{\Pr(B)} = \frac{\Pr(B \,|\, E_j)\Pr(E_j)}{\sum_{i=1}^{n} \Pr(B \,|\, E_i)\Pr(E_i)},$$

where $E_1, E_2, \ldots, E_n$ are mutually disjoint sets such that $\bigcup_{i=1}^{n} E_i = E$. Before we blindly apply this theorem, we must identify the probability space and the events we will analyze. Let us define D as the event that the patient has the disorder. Similarly, let T be the event that the test is positive. As a result, the sample space is composed of four outcomes $\{(D,T),\, (D,\overline{T}),\, (\overline{D},T),\, (\overline{D},\overline{T}\})$. From the problem statement, these simple events have the following probabilities.

$$\Pr(D) = 0.02,\ \Pr(\overline{D}) = 0.98$$
$$\Pr(T \,|\, D) = 0.999,\ \Pr(T \,|\, \overline{D}) = 0.005$$

This question asks us to determine the probability that a person has the disorder given that they test positive. In terms of conditional probabilities, we want to determine $\Pr(D \,|\, T)$. Since the simple events $\{D, \overline{D}\}$ are mutually disjoint we can apply Bayes' Law as follows.

$$\Pr(D \,|\, T) = \frac{\Pr(T \,|\, D)\Pr(D)}{\Pr(T \,|\, D)\Pr(D) + \Pr(T \,|\, \overline{D})\Pr(\overline{D})}$$

Substituting for the known probabilities we have

$$\Pr(D \,|\, T) = \frac{(0.999)(0.02)}{(0.999)(0.02) + (0.005)(0.98)} = \frac{999}{1244} \approx 0.803.$$

In conclusion, we find that the probability that the person has the disorder to be approximately **80.3%**. Intuitively, we should have expected such a result since the incidence of the disorder in the general population is so low, however the Bayesian approach has quantified this expectation.

References

[1] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, 2005.