MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Fall 2006 15.085/6.436

Recitation 4

1 Inclusion-Exclusion Formula

We know that

$$P(A_1 \bigcup A_2) = P(A_1) + P(A_2) - P(A_1 \bigcap A_2)$$

Can we generalize this formula to n events $A_1, A_2, \ldots, A_n$?

Theorem:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_j P(A_j) - \sum_{j<k} P(A_j \bigcap A_k) + \sum_{j<k<l} P(A_j \bigcap A_k \bigcap A_l) - \cdots + (-1)^{n+1} P(\bigcap_{j=1}^{n} A_j)$$

Before proving this, we derive the following identity. Since

$$(x+y)^n = \sum_{i=1}^{n} \binom{n}{i} x^i y^{n-i}$$

it follows that

$$0 = (-1+1)^n = \sum_{i=1}^{n} (-1)^i \binom{n}{i} \tag{1}$$

Proof: Let I_k be the indicator function of the event A_k, and let I be the indicator function of the event $\bigcup_{i=1}^{n} A_i$. We need to show that

$$I = \sum_j I_j - \sum_{j<k} I_j I_l + \sum_{j<k<l} I_j I_k I_l - \cdots + (-1)^n \prod_j I_j \tag{2}$$

and then the theorem will follow by taking expectation of both sides. We will show that both sides of the above equation evaluate to the same thing for all events ω.

Let ω be an element of the sample space, and let Z be the number of sets A_i such that $\omega \in A_i$. If $Z = 0$, then both sides of Eq. (2) evaluate to 0. Suppose now that $Z > 0$. Then the left hand side of Eq. (2) is 1 while the right hand side is

$$\sum_{i=1}^{Z} (-1)^{i+1} S_i(\omega)$$

where $S_i(\omega)$ is the number of nonzero terms in the sum

$$\sum_{j_1<j_2<\cdot<j_i} I_{j_1} I_{j_2} \cdots I_{j_i}$$

In other words, $S_i(\omega)$ is the number of different groups of i events ω belongs to. But since ω belongs to a total of Z of the events $A_1, \ldots, A_n$, it follows that

$$S_i(\omega) = \binom{Z}{i}$$

so that the right-hand side of Eq. (2 is

$$\sum_{i=1}^{Z} (-1)^{i+1} \binom{Z}{i}$$

or

$$1 - \sum_{i=0}^{Z} (-1)^{i} \binom{Z}{i}$$

which by Eq. (1) is equal to 1.

2 Joint Lives

Of the $2n$ people in a given collection of n couples, exactly m die. Assuming that the m have been picked at random, find the mean number of surviving couples. *Hint:* Use indicator functions.

Solution

Let N be the random number of surviving couples. For the couple indexed by $i = 1, \ldots, n$, let A_i (resp. B_i) the event that the first (resp. second) partner of couple i survives.

$N = \sum_{i=1}^{n} \mathbf{1}_{A_i} \mathbf{1}_{B_i}$ Hence, $E[N] = \sum_{i=1}^{n} E[\mathbf{1}_{A_i} \mathbf{1}_{B_i}] = nE[\mathbf{1}_{A_1} \mathbf{1}_{B_1}]$.

On the other hand, observe that $E[\mathbf{1}_{A_1} \mathbf{1}_{B_1}] = \mathbb{P}(\text{a couple survives}) = \frac{2n-m}{2n} \frac{2n-1-m}{2n-1}$ since the first partner survives with probability $\frac{2n-m}{2n}$ and the second survives with probability $\frac{2n-1-m}{2n-1}$ given that the first person survives.

Hence, $E[N] = n \frac{2n-m}{2n} \frac{2n-1-m}{2n-1}$.

3 Uniform random variables and infinite coin tosses

See appendix B of lecture 5.

6.436J / 15.085J Fundamentals of Probability
Fall 2008