MthStat 465, Spring 2005, Homework Number 2

1. A complete set of events in a sample space is called a *sigma algebra*. What three properties must a set of events have in order to be a sigma algebra?

   A set of subsets of a sample space S is a sigma algebra if

   (a) It contains S;
   (b) It is closed under complementation;
   (c) It is closed under countably infinite unions.

2. Give a full model (Sample space, sigma algebra, probability measure) to describe the following: A fair coin is tossed twice. Explicitly list all the events. What is the probability that there are more heads than tails?

   Each outcome can be represented by an ordered pair, where the first element of the ordered pair is the result of the first toss, and the second in the result of the second toss. Let H denote Head and T denote T. The sample space S is

$$S = \{(H,H),(H,T),(T,H),(T,T)\}.$$

   The set of events should be all subsets of S:

$$\left\{\begin{array}{cccc} S & \emptyset & \{(H,H),(H,T)\} & \{(T,T),(T,H)\} \\ \{(H,H),(T,H)\} & \{(T,T),(H,T)\} & \{(H,H),(T,T)\} & \{(H,T),(T,H)\} \\ \{(H,H)\} & \{(H,T),(T,H),(T,T)\} & \{(H,T)\} & \{(H,H),(T,H),(T,T)\} \\ \{(T,H)\} & \{(H,T),(H,H),(T,T)\} & \{(T,T)\} & \{(H,H),(T,H),(H,T)\} \end{array}\right\}$$

   The probability of any event is the number of its elements divided by 4.

   Since the event "more heads than tails" is $\{(H,H)\}$, the probability of more heads than tails is 1/4.

3. (From *Basic Probability Theory* by R. Ash) A public opinion poll circa 1850 consisted of the following three questions:

   (a) Are you a registered Whig?
   (b) Do you approve of President Fillmore's performance in office?
   (c) Do you favor the Electoral College system?

   A group of 1000 people were polled. Each person answered each question "yes" or "no". It was found that

   (a) 520 people were against the Electoral College system.
   (b) 325 people answered yes to exactly 2 of the questions, and no to a third.
   (c) 100 people answered yes to all three questions.
   (d) 125 registered Whigs approved of President Fillmore's performance.

   How many of those who favored the Electoral College system disapproved of Fillmore's performance and were not registered Whigs? You might want to use a Venn diagram to answer this question.

   This problem is most simply answered by using a Venn Diagram. Let W be the set of Whigs, let E be the supporters of the electoral college system, and let F be the set of people who approve of Fillmore. The diagram divides the sample space into the following 8 non-overlapping parts:

   (a) $W \cap E \cap F$, containing a people;
   (b) $W \cap E \cap F^c$, containing b people;
   (c) $W \cap E^c \cap F$, containing c people;
   (d) $W^c \cap E \cap F$, containing d people;
   (e) $W^c \cap E^c \cap F$, containing e people;
   (f) $W^c \cap E \cap F^c$, containing f people;

(g) $W \cap E^c \cap F^c$, containing g people;
(h) $W^c \cap E^c \cap F^c$, containing h people.
In terms of the variables a through h the given information is
(a) 1000 people polled:

$$a + b + c + d + e + f + g + h = 1000.$$

(b) 520 people were against the Electoral College system:

$$h + g + c + e = 520.$$

(c) 325 people answered yes to exactly 2 of the questions, and no to a third:

$$b + c + d = 325.$$

(d) 100 people answered yes to all three questions:

$$a = 100.$$

(e) 125 registered Whigs approved of President Fillmore's performance:

$$a + c = 125.$$

The solution of the problem is the value of f. From the first two equations we know that

$$a + b + d + f = 480$$

From the last two equations we know that $c = 25$. Using this in the third equation we find that $b + d = 300$. Therefore

$$480 = a + b + d + f = 100 + 300 + f$$

so $f = 80$.

4. Show that for any sets A and B that $(A \cap B)^c = A^c \cup B^c$. Illustrate your argument with a Venn diagram. Use this to conclude that the intersection of two events is also an event.

   Suppose that $x \in (A \cap B)^c$. Then x is not in both A and B, so either $x \in A^c$ or $x \in B^c$. This means $x \in A^c \cup B^c$. On the other hand, if $x \in A^c \cup B^c$ then either x is not in A or x is not in B. Either way, x is not in the intersection of A and B, so $x \in (A \cap B)^c$.

   I can't draw you the picture here.

   Suppose then that A and B are events. Then A^c and B^c are events. This means that $A^c \cup B^c$ is an event too. Since $(A \cap B)^c = (A^c \cup B^c)$ we know that $(A \cap B)^c$ is an event, and it complement, $A \cap B$ is an event too.

5. Suppose that our sample space, S, is $S := \{0, 1, 2, 3, \cdots, 10\}$. The set of all subsets of a complete sigma algebra on S. Define

$$p_k := \frac{10!}{k!(10-k)!} \left(\frac{3}{7}\right)^k \left(\frac{4}{7}\right)^{10-k}$$

and assign probabilities by the rule

$$\Pr(F) = \sum_{k \in F} p_k.$$

What is the exact value of $\Pr(\{0, 1, 2, 3, 4\})$? Is Pr a legitimate assignment of probabilities? Generalize this example to other sets of the form $S = \{0, 1, 2, \ldots, N\}$ where N is a positive integer.

$$
\begin{aligned}
\Pr(\{0,1,2,3,4\}) &= p_0 + p_1 + p_2 + p_3 + p_4 \\
&= \left(\frac{4}{7}\right)^{10} + 10\left(\frac{4}{7}\right)^{9}\left(\frac{3}{7}\right) + 45\left(\frac{4}{7}\right)^{8}\left(\frac{3}{7}\right)^{2} \\
&\quad +120\left(\frac{4}{7}\right)^{7}\left(\frac{3}{7}\right)^{3} + 210\left(\frac{4}{7}\right)^{6}\left(\frac{3}{7}\right)^{4}
\end{aligned}
$$

This is a legitimate assignment of probabilities because

(a) Each event is assigned a non-negative probability;

(b) The probability of the sample space is 1, since

$$
\begin{aligned}
\Pr(S) &= \sum_{k=0}^{10} \frac{10!}{k!(10-k)!}\left(\frac{3}{7}\right)^{k}\left(\frac{4}{7}\right)^{10-k} \\
&= \left(\frac{3}{7}+\frac{4}{7}\right)^{10} \\
&= 1
\end{aligned}
$$

by using the binomial theorem;

(c) The formula forces the probability of the union of disjoint sets to be the sum of their probabilities.

We can generalize to $S = \{0, 1, \dots, N\}$ and

$$p_k = \frac{N!}{k!(N-k)!}x^k(1-x)^{N-k}$$

for any $x \in [0,1]$ and any positive integer N.

6. What is meant by geometric probability. Illustrate your answer with examples. Include the following problem. Consider all quadratic polynomials of the form $p(x) = x^2 + bx + a$ where $-2 \le b \le 2$ and $-2 \le a \le 2$. If a $p(x)$ is chosen at random from this set of polynomials, what is the probability that is has two distinct real roots?

By geometric probability we mean the assignment of probability by the ratio of the size of an event to the size of the sample space in case the sample space is a geometric object, such as a disk, a sphere, the interior of a square, and so forth.

For example, choosing a number at random in an interval $[a, b]$ is an example of geometric probability. Another example is throwing darts at a dart board, if we assume that we have equal probability of hitting regions of the board that have the same area.

In the polynomial example, the polynomial $p(x) = x^2 + bx + c$ has two real roots if and only if $b^2 - 4a > 0$. (Use the quadratic formula!!) So we need the ratio of the area of the region $b^2/4 > a$, $-2 \le b \le 2$, and $-2 \le a \le 2$ to the region $-2 \le b \le 2$, and $-2 \le a \le 2$. The latter has area 16, while the former (using calculus) has area

$$8 + \int_{-2}^{2} \frac{b^2}{4}\,db = \frac{28}{3}.$$

The probability in question is then $7/12$.

7. The symbol ${}_nC_r$ is often used to denote the number of ways to choose r things out of n things. Why is ${}_nC_0 = 1$ and ${}_nC_1 = n$ when n is a positive integer? What is the meaning of ${}_0C_0$ and why? Why is ${}_nC_k =_n C_{n-k}$ when n and k are integers and $0 \leq k \leq n$? Why should ${}_nC_k = 0$ if $k > n > 0$?

   When we speak of choosing some number of items from some other number of items we are speaking of the number of subsets of a given size there are in a certain set. Since there is only one set with no items in it, namely the empty set, we must have ${}_nC_0 = 1$ if n is a positive integer. ${}_nC_1 = n$ because each element of the set can be used to form a one element subset, giving us n one element sets. Since the empty set has only one subset, we have ${}_0C_0 = 1$.

   ${}_nC_k =_n C_{n-k}$ when n and k are integers and $0 \leq k \leq n$ because there is a one-to-one correspondence between each subset and its complement.

   Finally ${}_nC_k = 0$ if $k > n > 0$ because the number of element of a subset cannot exceed the number of elements of the original set.

8. Derive the formula

$$ {}_nC_k = \frac{n!}{k!(n-k)!} $$

   for integers n and k satisfying $0 \leq k \leq n$.

   The case where $k = 0$ follows from the discussion of the previous problem and the fact that $0!$ is defined to be 1:

$$ {}_nC_0 = 1 = \frac{n!}{0!(n)!} $$

   Now, supposing that both n and k are positive integers, we reason as follows. Suppose I wanted to count the number of sequences of length k I can make by choosing from a set of N symbols, without replacement. Call this number ${}_nP_k$. On the one hand, if I choose the symbols one at a time, we see that

$$ {}_nP_k = n \times (n-1) \times \cdots \times (n-(k-1)) = \frac{n!}{(n-k)!}. $$

   On the other hand, I could first choose a set of k symbols to use, which can be done ${}_nC_k$ ways, and put those k symbols in order, which can be done $k!$ ways, so we see from the multiplication formula that

$$ {}_nP_k =_n C_k \times k!. $$

   Therefore

$$ \frac{n!}{(n-k)!} =_n P_k =_n C_k \times k!, $$

   which, when solved for ${}_nC_k$, gives the desired formula.

   Give two different arguments demonstrating that ${}_nC_k +_n C_{k+1} =_{n+1} C_{k+1}$.

On the one hand we can argue algebraically:

$$\begin{aligned}
{}_nC_k + {}_nC_{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} \\
&= \frac{n!}{k!(n-k-1)!}\left(\frac{1}{n-k} + \frac{1}{k+1}\right) \\
&= \frac{n!}{k!(n-k-1)!}\left(\frac{n+1}{(n-k)(k+1)}\right) \\
&= \frac{(n+1)!}{(k+1)!(n-k)!} \\
&= {}_{n+1}C_{k+1}
\end{aligned}$$

Better yet, we can argue straight from the definition. If we want to select subsets from $\{1, 2, \ldots, N+1\}$ there are two types of subsets: those that contain $N+1$ and those that don't. If we want to count subsets of size $k+1$ that don't contain $N+1$ we are selecting them from $\{1, 2, \ldots, N\}$ and this can be done ${}_NC_{k+1}$ ways. If we want them to contain $N+1$ we must select subsets of size k from $\{1, 2, \ldots, N\}$ which can be done ${}_NC_k$ ways, and then add in $N+1$. Since the total number of subsets of size $k+1$ from $\{1, 2, \ldots, N+1\}$ is ${}_{N+1}C_{k+1}$ we must have

$${}_{N+1}C_{k+1} = {}_NC_{k+1} + {}_NC_k.$$

What is the relation to Pascal's triangle?

Pascal's triangle is an algorithm for generating these combinatorial coefficients recursively. It says that if we arrange them in this triangular pattern:

$$\begin{array}{ccccccccccccc}
 & & & & & & {}_0C_0 & & & & & & \\
 & & & & & {}_1C_0 & & {}_1C_1 & & & & & \\
 & & & & {}_2C_0 & & {}_2C_1 & & {}_2C_2 & & & & \\
 & & & {}_3C_0 & & {}_3C_1 & & {}_3C_2 & & {}_3C_3 & & & \\
 & & {}_4C_0 & & {}_4C_1 & & {}_4C_2 & & {}_4C_3 & & {}_4C_4 & & \\
 & {}_5C_0 & & {}_5C_1 & & {}_5C_2 & & {}_5C_3 & & {}_5C_4 & & {}_5C_5 & \\
{}_6C_0 & & {}_6C_1 & & {}_6C_2 & & {}_6C_3 & & {}_6C_4 & & {}_6C_5 & & {}_6C_6
\end{array}$$

and so on, and then pick an element in any row, it is equal to the sum of the two elements in the row above and just to the right and left. This is precisely what our formula says.

9. Why is it better to define the independence of A and B by requiring $\Pr(A \cap B) = \Pr(A)\Pr(B)$ than $\Pr(A|B) = \Pr(A)$?

Independence is a symmetric relation between A and B, so the definition should reflect this symmetry if possible. The relation $\Pr(A \cap B) = \Pr(A)\Pr(B)$ is symmetric in A and B, while $\Pr(A|B) = \Pr(A)$ is not. There is no reference to the magnitude of probabilities in independence, yet $\Pr(A|B) = \Pr(A)$ requires $\Pr(B) > 0$. Finally, $\Pr(A \cap B) = \Pr(A)\Pr(B)$ implies $\Pr(A|B) = \Pr(A)$ if $\Pr(B) > 0$, so the former is a generalization of the latter, and, therefore, is more applicable.

Suppose that $\Pr(A) = 1/3$, $\Pr(B) = 1/5$ and A and B are independent. What is $\Pr(A \cup B)$?

We have

$$\begin{aligned}
\Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\
&= \Pr(A) + \Pr(B) - \Pr(A)\Pr(B) \\
&= \frac{1}{3} + \frac{1}{5} - \frac{1}{15} \\
&= \frac{7}{15}
\end{aligned}$$

10. (From *Basic Probability Theory* by R. Ash) In a certain village 20% of the population has disease D. A test is administered which has the property that if the person has D, the test will be positive 90% of the time, and if the person does not have D, the test will still be positive 30% of the time. All those whose test is positive are given a drug which invariably cures the disease, but produces a characteristic rash 25% of the time. Given that a person picked at random has the rash, what is the probability that that person had D to begin with?

    Let D be the event that a person has the disease, let T be the event that the test is positive, and let R be the event that a rash is present. We want to find $\Pr(D|R)$. Since we know that $\Pr(D|R) = \Pr(D \cap R)/\Pr(R)$ we will compute $\Pr(D \cap R)$ and $\Pr(R)$ separately. Since you cannot get the rash if the test is negative (because no drug is administered), $R \cap T^c = \emptyset$.

    First,

$$\begin{aligned}
\Pr(D \cap R) &= \Pr(D \cap R \cap T) + \Pr(D \cap R \cap T^c) \\
&= \Pr(D \cap R \cap T) + \Pr(D \cap \emptyset) \\
&= \Pr(D \cap R \cap T) + \Pr(\emptyset) \\
&= \Pr(D \cap R \cap T) \\
&= \Pr(R|D \cap T)\Pr(D \cap T) \\
&= \frac{1}{4}\Pr(T|D)\Pr(D) \\
&= \frac{1}{4} \times \frac{9}{10} \times \frac{1}{5} \\
&= \frac{9}{200}
\end{aligned}$$

On the other hand, $\Pr(R) = \Pr(R \cap D) + \Pr(R \cap D^c)$, and

$$\begin{aligned}
\Pr(D^c \cap R) &= \Pr(D^c \cap R \cap T) + \Pr(D^c \cap R \cap T^c) \\
&= \Pr(D^c \cap R \cap T) + \Pr(D^c \cap \emptyset) \\
&= \Pr(D^c \cap R \cap T) + \Pr(\emptyset) \\
&= \Pr(D^c \cap R \cap T) \\
&= \Pr(R|D^c \cap T)\Pr(D^c \cap T) \\
&= \frac{1}{4}\Pr(T|D^c)\Pr(D^c) \\
&= \frac{1}{4} \times \frac{3}{10} \times \frac{4}{5} \\
&= \frac{12}{200}
\end{aligned}$$

so $\Pr(R) = 21/200$ and $\Pr(D|R) = (9/200)/(21/200) = 3/7$. This is the same as $\Pr(D|T)$ because the incidence of rash does not depend on whether or not you actually have the rash. If you write out all the formulas without simplifying the fractions you can see the factors of 1/4 divide out.

11. Suppose that $S = \{1, 2, 3\}$, the sigma algebra is $\{\{1\}, \{2, 3\}, \emptyset, S\}$, and $\Pr(\{1\}) = 1/6$ and $\Pr(\{2, 3\}) = 5/6$. Define $X : S \to (-\infty, \infty)$ by $X(1) = 2$, $X(2) = 3$ and $X(3) = 4$. Explain why X is not a random variable.

For X to be a random variable we have to have $\{s : a < X(s) \leq b\}$ be an event for every choice of a and b. Since $\{s : 3 < X(s) \leq 5\} = \{3\}$ which is not an event, X cannot be a random variable.

How would the sigma algebra have to be modified so that X becomes a random variable?

X would be a random variable if every subset of S were an event.

Instead, what value should be assigned to $X(2)$ to make X a random variable?

If $X(2) = X(3)$ then X will be a random variable. An example of this sort was given in the on-line lecture notes.

Let $R : S \to (-\infty, \infty)$ be given by $R(1) = 2$, $R(2) = R(3) = 4$. List out $\{R = 2\}$, $\{R < 4\}$ and $\{R \leq 4\}$, and give the probability of each.

(a) $\{R = 2\} = \{s : R(s) = 2\} = \{1\}$ and $\Pr(\{1\}) = 1/6$.

(b) $\{R < 4\} = \{s : R(s) < 4\} = \{1\}$ and $\Pr(\{1\}) = 1/6$.

(c) $\{R \leq 4\} = \{s : R(s) \leq 4\} = \{1, 2, 3\}$ and $\Pr(\{1, 2, 3\}) = 1$.

Make a graph of $F_R(t) := \Pr(\{R \leq t\})$ for $-\infty < t < \infty$.

I can't insert a graph, but the rule for F_R is

$$F_R(t) = \begin{cases} 0 & \text{if} \quad t < 2 \\ 1/6 & \text{if} \quad 2 \leq t < 4 \\ 1 & \text{if} \quad 4 \leq t \end{cases}$$