Probability Generally

Probability questions have:

random process/phenomena which generates observed outcome – defines a probability distribution

sample space all possible outcomes from the process. We denote it by $\Omega$

event subset of possible outcomes which you are interested in knowing the probability of

Examples

Example 1: Randomly pick 2 nucleotides from given DNA sequences. What is the probability of picking two of the same nucleotides?

My sample space is all possible 2 letter combinations:

$$\Omega = \{ \text{AA, CA, GA, TA, AG, CG, GG, TG, AT, CT, GT, TT, AC, CC, GC, TC}\}$$

The event $E = \{\text{AA, TT, GG, CC}\}$.

And the random process ... we need more information to precisely define the probability distribution. 

- ■

Example 2 (Infinite sample space) Toss a coin until heads comes up. What is the probability heads turns up after an even number of tosses?

\[\Omega = \{1, 2, 3, \ldots\}\]

The question defines an event of interest $A = \{2, 4, 6, \ldots\}$

Making a tree diagram, we see that $P(1) = 1/2, P(2) = 1/4, P(3) = 1/8, \ldots$

The distribution function in this case is: $m(i) = 1/2^i$

—by symmetry you would think $P(A) = 1/2$, and by symmetry you would be wrong!—-

- ■

Set Facts, Venn Diagrams

Union $A \cup B \rightarrow A$ or $B$

Intersection $A \cap B$ or $AB \rightarrow A$ and $B$

Difference $A \setminus B$ or $A \setminus B \rightarrow A$ but not $B$

Complement $A^c$ or $\tilde{A}$ or $\bar{A}$ $\rightarrow$ not $A$

A nice website to play around with Venn Diagrams and Probability:
http://stat-www.berkeley.edu/users/stark/Java/Venn.htm

Rules for set manipulation:
(a) Distributive Rule: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(b) De Morgan’s Laws: $(A \cup B)^c = A^c \cap B^c$; $(A \cap B)^c = A^c \cup B^c$.

Events are said to be mutually exclusive if $A \cap B = \emptyset$ (they can’t happen at the same time)

Note that this is a property of the sets, not of the probability.
Example 3 (Mutually Exclusive Sets): Suppose we are interested in the set of outcomes of 4 rolls of the dice where the first die to be a 1 or 6. If I falls first, clearly it rules out the possibility of 6 being first, so the two sets of outcomes, \{first die =1\} and \{first die =6\}, are mutually exclusive.

What happens if we want to examine the sets: \{throw a 1 first\} and throw one 6 in the four throws? Then the sets \{first die =1\} and \{one of 4 die =6\} are no longer mutually exclusive.

Properties of Probability Distributions

Basic Axioms of Probability:

(a) \(0 \leq P(A) \leq 1\) for any set \(A\).

(b) \(P(\Omega) = 1\), where \(\Omega\) is the sample space.

(c) If \(E_1, E_2, \ldots, E_n\) is a sequence of mutually exclusive events, that is for all \(i \neq j\) different: \(E_i \cap E_j = \phi\) we have the finite additivity property: \(P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)\)

Example 1 (continued) So, from the previous example, if we have three events, \(A=\{\text{ends in G}\}, B=\{\text{ends in T}\}, C=\{2\ \text{repeated nucleotides}\}\), then

\[P(A \cup B) = P(A) + P(B)\]

But

\[P(A \cup C) \neq P(A) + P(C)\]

Consequences of Axioms:

Property 1: \(P(A \cup B) = P(A) + P(B) - P(A \cap B)\)

Property 2: \(P(\Omega) = P(E \cup E^c) = P(E) + P(E^c)\)

Property 3: If \(E \subseteq F\), \(P(E) \leq P(F)\).

Property 4: \(P(E \cup F) = P(E) + P(F) - P(E \cap F)\)

Special Case: If \(E\) and \(F\) are disjoint, (mutually exclusive), then: \(P(E \cup F) = P(E) + P(F)\)

Property 5: If \(F \subseteq E\) \(P(E \cap F^c) = P(E) - P(F)\)

Definition: We define a partition of the sample space \(\Omega\) to be a sequence of pairwise disjoint sets \(A_1, A_2, \ldots, A_n\), whose union is \(\Omega\).

Property 6: If \(A_1, A_2, \ldots, A_n\) forms a partition of \(\Omega\) then: \(P(E) = \sum_{i=1}^{n} P(E \cap A_i)\)

Property 7: \(P(A) = P(A \cap B) + P(A \cap B^c)\)

Probability that Stanford will beat Cal?

You may believe that it is much more likely that Stanford will win and you say the odds are “2 to 1” that Stanford wins – that it is twice as likely that Stanford wins than not. This implies you believe the probability of Stanford winning is \(2/3\). Why? r to 1 odds that \(E\) occurs? This means you believe it is \(r\) times more likely that \(E\) occurs than not.

\(P(E) = rP(E^c)\)

So that \(P(E) = \frac{r}{r+1}\).

So you think the probability of winning is \(r/\ r + 1\)

If we generalise this to an event \(E\) whose odds are \(r\) to \(s\), written \((r : s)\), then \(P(E) = \frac{r/s}{r/s+1} = \frac{r}{r+s}\)

If \(P(E) = p\) then the odds in favor of \(E\) are: \(r : s\) where \(r/s = p/(1-p)\).  

Conditional Probability – Appendix 10.4

Suppose that you have to guess the suit of a playing card drawn from a 52 pack at random, you probability of guessing correctly is 1/4.

Now suppose I tell you that it is red. You then have a 1/2 chance, because there are less choices, you can use the additional information to restrict the space of possible outcomes.
Definition and Properties

Definition of the conditional probability of $A$ given $B$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Averaging rule:

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

General Averaging Rule:
Let $F_i, i = 1 \ldots n$ be a partition of $S$, by which I mean that the events $F_i$ are all mutually exclusive and that their union is $S$, then:

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

Bayes Rule
To find the opposite conditional probability than the ones given:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

General Bayes Rule:
Let $F_i, i = 1 \ldots n$ be a partition of $S$, then $\{EF_i\}_{i=1..n}$ is a partition of $E$ and:

$$P(F_i|E) = \frac{P(E|F_i)P(F_i)}{\sum_{j=1}^{n} P(E|F_j)P(F_j)}$$

Conditional Probability IS a probability
Conditional probability obeys the three axioms necessary to having a probability, provided of course it is defined, ie the event that we are conditionning on has probability non zero.

Independence

Intuitive definition:
Knowing that another event $B$ occurs does not affect the probability of the event $A$ of occurring, this means that $A$ and $B$ are independent.

Definition of Independence as a Formula:

$$P(A \text{ given } B) = P(A|B) = P(A)$$

$$P(A \text{ and } B) = P(A \cap B) = P(A) \times P(B) =$$

This also means $A$ and $B$ are independent if

$$P(A|B^c) = P(A), \quad P(B) < 1$$

Example Conditional probabilities can both be bigger or smaller:
In considering colorblindness, suppose I consider the binary random variables associated to color blindness and gender (associate 0 if male, 1 if female), these are called indicator variables, we can tabulate the probabilities of all 4 possible pairs of outcomes as:

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorblind</td>
<td>14</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Not Colorblind</td>
<td>186</td>
<td>199</td>
<td>385</td>
</tr>
<tr>
<td>Total</td>
<td>1/2</td>
<td>1/2</td>
<td>1</td>
</tr>
</tbody>
</table>
So that from this table of joint distribution we read:

\[
\begin{align*}
P(\text{colorblind}) & = P(C) = \frac{15}{400} \\
P(\text{man}) & = P(M) = \frac{1}{2} \\
P(\text{colorblind and male}) & = P(C \text{ and } M) = \frac{14}{400} \\
P(\text{colorblind given male}) & = \frac{P(C \text{ and } M)}{P(M)} = \frac{14}{\frac{1}{2}} = \frac{14}{200} = \frac{7}{100} \\
P(\text{colorblind}) & = P(C) = \frac{15}{400} \\
P(\text{colorblind given woman}) & = \frac{P(C \text{ and } W)}{P(W)} = \frac{\frac{1}{400}}{\frac{1}{2}} = \frac{1}{200}
\end{align*}
\]

We see that for women: \( P(C|W) \leq P(C) \)
And for men: \( P(C|M) \geq P(C) \)

Thus, when \( A \) and \( B \) are not independent:
Sometimes we have \( P(A \text{ given } B) \leq P(A) \),
Sometimes we have \( P(A \text{ given } B) \geq P(A) \),

When two events are independent the probability of them both occurring is just the product of their probabilities.

When random mating occurs in a population, if \( P(\text{receive } X \text{ with colorblind}) = .07 \), then \( P(\text{receive } 2X) = 0.07 \times 0.07 = 0.0049 \). The multiplication rule applies because of the independence of the X chromosomes (mother and father).

Examples:
We draw two cards one at a time from a shuffled deck of 52 cards.

A The first card is a 7 ♣.
B The second card is a queen ♥.

\[ P(B|A) \neq P(B) \]

These events are not independent.

Beware the multiplication rule is ONLY available if and only if the events ARE independent.

Problems To Try:
1. De Méré’s problem is whether or not it is more likely to get at least one double six in 24 throws of a pair of dice or to get at least one six in 4 throws of a die?
2. A rare disease is caused by a mutation of a gene. Individuals heterozygous for this mutation, show symptoms at various stages (heterozygous means one of your two copies you received from each of your parents has the mutation and one of the 2 copies does not). 1/3 of heterozygotes show symptoms by age 35. We know Joe’s mother was heterozygous for the disease. Joe’s is 35 and has shown no symptoms. What is the probability Joe is heterzygous – i.e. the probability he got his mother’s mutated copy of the gene, and not her normal copy? (the disease is rare, so you can assume Joe’s father did not have the disease)
3. A False Positive Puzzle
   (a) In random testing, you test positive for a disease
   (b) In 5% of cases, this test shows positive even when the subject does not have the disease
   (c) In the population at large, one person in 1,000 has the disease.

What is the chance that you have the disease?