Lecture 5

Conditional Probability

Review Exercises

1. Suppose you roll a fair 6-sided die 12 times. What is the probability of getting each value exactly twice?

2. You keep rolling a fair 100-sided die until it is strictly larger than 60. Let \( A_j \) denote the event that you roll a total of \( j \) times.

   (a) What is \( P(A_j) \)?

   (b) What is \( \sum_{j=1}^{\infty} P(A_j) \)?

   (c) Let \( N_j \) denote the event that you roll a total of \( j \) times, and the last roll is a 90. Compute \( P(N_j|A_j) \).

   (d) What is probability the last roll is 90? [Use LOTP and the previous parts.]

   (e) Suppose you roll a fair 100-sided die once. Let \( C \) be the event of rolling strictly larger than 60, and let \( D \) be the event of rolling a 90. What is \( P(D|C) \)?

Solutions

1. \[
\frac{\binom{12}{2,2,2,2,2,2}}{6^{12}} = \frac{12!/(2!)^6}{6^{12}}.
\]

   This can also be written as

   \[
   \frac{\binom{12}{2} \binom{10}{2} \binom{8}{2} \binom{6}{2} \binom{4}{2} \binom{2}{2}}{6^{12}}.
   \]

2. Here we can define \( S \) as follows:

   \[
   S = \{(d_1, d_2, \ldots, d_n) : 1 \leq n, \ 1 \leq d_i \leq 60 \text{ for } i < n, 60 < d_n \leq 100\}
   \]

   with

   \[
   P(\{(d_1, \ldots, d_n)\}) = \frac{1}{100^n}.
   \]

   This is a slight generalization of the general finite sample space called the general countable sample space (this has been added to the end of the Lecture 2 notes).

   (a) \( \left( \frac{60}{100} \right)^{j-1} \frac{40}{100} \) for \( j \geq 1 \).
(b) 
\[ \sum_{j=1}^{\infty} P(A_j) = \frac{40}{100} \sum_{j=1}^{\infty} \left( \frac{60}{100} \right)^{j-1} = \frac{40}{1 - \frac{60}{100}} = 1. \]

(c) 
\[ P(N_j | A_j) = \frac{P(N_j \cap A_j)}{P(A_j)} = \frac{P(N_j)}{P(A_j)} = \left( \frac{60}{100} \right)^{j-1} \frac{1}{\frac{100}{60}} = \frac{1}{40}. \]

(d) Let \( M \) denote the event the last roll is a 90. Then we have
\[ P(M) = \sum_{j=1}^{\infty} P(N_j | A_j) P(A_j) = \frac{1}{40} \sum_{j=1}^{\infty} P(A_j) = \frac{1}{40}. \]

Here we have applied the following countable version of the LOTP:

**Theorem 1** (Countable Law of Total Probability). If \( A_1, A_2, \ldots \) are pairwise disjoint with \( P(A_i) > 0 \) for all \( i \geq 1 \) then we have
\[ P(B) = \sum_{i=1}^{\infty} P(B | A_i) P(A_i). \]

(e) 
\[ P(D | C) = \frac{P(D \cap C)}{P(C)} = \frac{1/100}{40/100} = \frac{1}{40}. \]

**Independence**

Given two events \( A, B \), we say that they are independent if \( P(A \cap B) = P(A)P(B) \). Note that if \( P(B) > 0 \), then this is equivalent to saying \( P(A | B) = P(A) \). That is, learning \( B \) doesn’t change our beliefs on \( A \).

**Example 2.** Suppose we flip 2 coins. Then the event that the first coin is heads is independent of the event that the second coin is heads.

We can extend this to multiple events, but there is a catch.

**Definition 3** (Independence). We say that the events \( A_1, \ldots, A_n \) are independent if
\[ P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}) \]
for any \( 2 \leq k \leq n \) and \( 1 \leq i_1 < i_2 < \cdots < i_k \leq n \).

It isn’t enough to have each pair be independent. Let’s do a concrete case. Suppose \( n = 3 \). Then \( A_1, A_2, A_3 \) are independent if
\[
\begin{align*}
P(A_1 \cap A_2) &= P(A_1)P(A_2), \\
P(A_1 \cap A_3) &= P(A_1)P(A_3), \\
P(A_2 \cap A_3) &= P(A_2)P(A_3), \\
P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2)P(A_3). 
\end{align*}
\]
If we have an infinite collection of events, we say they are independent if every finite subcollection is independent.

**Independence Exercises**

1. Suppose we flip a fair coin 3 times. Let \( H_i \) denote the event that the \( i \)th flip is heads. Show that \( H_1, H_2, H_3 \) are independent.

2. Show that if \( E \) and \( F \) are independent, then \( E \) and \( F^c \) are independent.

3. A circuit consists of \( n \) components. Let \( A_i \) denote the event that the \( i \)th component fails, and let \( P(A_i) = p_i \). Assuming \( A_1, \ldots, A_n \) are independent compute the following:
   
   (a) The probability at least one of the components fails.
   
   (b) The probability all of the components fail.

4. Suppose a can has 3 coins with head-probabilities \( 1/3, 1/2, 2/3 \), respectively. We randomly pick out one coin, and flip it three times. Let \( H_i \) be the event the \( i \)th flip is heads.
   
   (a) Are the \( H_i \) independent?
   
   (b) Are the \( H_i \) independent if we condition on which coin we chose?
   
   (c) What is the probability of getting 3 heads?

**Solutions**

1. We have \( P(H_i) = 1/2 \), \( P(H_iH_j) = 1/4 \) for \( i \neq j \) and \( P(H_1H_2H_3) = 1/8 \).

   There is some sense in which independence holds here since we assumed it so. Even though we didn’t call it independence at the time, when we modeled coin flips or throwing dice we assumed the flips and dice didn’t interfere with each other, and chose probabilities accordingly. This implicitly amounted to assuming the dice/flips were independent.

2. We must prove that \( P(EF^c) = P(E)P(F^c) \). Note that

   \[
   P(E) = P(EF) + P(EF^c) = P(E)(1 - P(F^c)) + P(EF^c),
   \]

   Thus \( P(E)P(F^c) = P(EF^c) \).

   A more general result holds true:

   **Theorem 4.** Let \( A_1, \ldots, A_n \) be independent events. Then \( B_1, \ldots, B_n \) are independent events, where each \( B_i \) is either \( A_i \) or \( A_i^c \).

   The above is really \( 2^n \) theorems in one, where for each version we choose to put complements on some of the \( A_i \)’s.
3. (a) \[ 1 - P(A_1^c A_2^c \cdots A_n^c) = 1 - \prod_{i=1}^n (1 - p_i). \]
(b) \[ P(A_1 A_2 \cdots A_n) = \prod_{i=1}^n p_i. \]

4. (a) No. Using LOTP:

\[
P(H_1) = \frac{1}{3^2} + \frac{1}{3 \cdot 2} + \frac{2}{3} \cdot \frac{2}{3} = \frac{1}{2}
\]
\[
P(H_1 H_2) = \frac{1}{3} \cdot \frac{1}{3^2} + \frac{1}{3} \cdot \frac{1}{2^2} + \frac{1}{3} \cdot \frac{2^2}{3^2} = \frac{29}{108}.
\]

Thus \( P(H_1)P(H_2) \neq P(H_1H_2) \).

(b) Yes, this is called conditional independence. More generally:

**Definition 5** (Conditional Independence). Let \( A_1, \ldots, A_n, B \) be events with \( P(B) > 0 \). We say that \( A_1, \ldots, A_n \) are conditionally independent given \( B \) if they are independent with respect to the measure \( P_B(E) = P(E|B) \).

By our standard method for modeling coin flips we have

\[
P(H_iH_j|C_1) = \left( \frac{1}{3} \right)^2 \quad \text{and} \quad P(H_1H_2H_3|C_1) = \left( \frac{1}{3} \right)^3,
\]

where \( C_1 \) is the event of picking the first coin. This shows that once we condition on the first coin the flips are independent. The same argument will show that conditioning on \( C_2 \) or \( C_3 \) also yields independence.

(c) By LOTP (letting \( C_i \) denote the event of choosing the \( i \)th coin):

\[
P(H_1H_2H_3) = P(H_1H_2H_3|C_1)P(C_1) + P(H_1H_2H_3|C_2)P(C_2) + P(H_1H_2H_3|C_3)P(C_3)
\]
\[
= \frac{1}{3} \left( \frac{1}{3^3} + \frac{1}{2^3} + \frac{2^3}{3^3} \right).\]

**Interesting Problems**

1. (Monty Hall) You are on a game show with 3 doors. Behind one is a car, and the other two have goats. You choose a door. Afterward the host opens a door you didn’t pick that has a goat behind it (the host knows where the car is; if you picked the car he chooses randomly with equal chance). He then offers you the option to switch choices to the remaining closed door. Do you switch?

2. A test has been developed to diagnose a disease. It has a 2% chance of a false negative, and a 2% chance of a false positive. The disease infects 1 out of every 1000 people in the population. Assuming you take the test and it unfortunately comes back positive. What is the probability you have the disease?
3. You go to a casino and have decided to play a game. Each round you have a 51% chance of losing your bet, and a 49% chance of winning an amount equal to your bet. Letting $E_i$ denote the event of winning the $i$th round, you may assume that the $E_i$ are independent. You have $500 and have decided to bet $1 each time. What is the chance you can get to $550 before you lose the money you brought?

Solutions

1. You always switch, and win the car with probability $2/3$. To see this, suppose you initially choose door 1. Let $C_i$ denote the event the car is behind door $i$, and let $W$ denote the event of winning the car assuming you always switch. Then we have

$$P(W) = P(W|C_1)P(C_1) + P(W|C_2)P(C_2) + P(W|C_3)P(C_3) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}.$$  

This shows that blindly switching gives $2/3$ chance of winning. Suppose we also take Monty’s (the host’s) chosen door into account. To that end, let $M_i$ denote the event that Monty opens door $i$. Note that

$$P(C_1|M_2) = \frac{P(M_2|C_1)P(C_1)}{P(M_2|C_1)P(C_1) + P(M_2|C_2)P(C_2) + P(M_2|C_3)P(C_3)}$$

$$= \frac{1/2 \cdot 1/3}{1/2 \cdot 1/3 + 0 \cdot 1/3 + 1 \cdot 1/3}$$

$$= \frac{1}{3}.$$ 

This shows the strategy of staying when Monty opens door 2 wins with probability $1/3$ (so switching wins with probability $2/3$). Nearly the same calculation gives the same result for when Monty opens door 3. Thus conditioning on Monty’s door choice gives the same winning probability when you switch. You can also do this with a tree diagram, and it could help illuminate what is going on. It can also help your intuition to think about the case with 1000 doors, when you choose a door, and Monty closes 998 goat-doors you haven’t picked.

One technical point. Above we said $P(C_1) = 1/3$. Strictly speaking, this isn’t known since the show doesn’t have to allocate the cars randomly. What we can say is that we randomly chose a door, and labeled the door we picked #1. Then $P(C_1) = 1/3$ makes sense.

2. Let $D$ denote the event of having the disease, $T$ denote testing positive. Then we have

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} = \frac{.98 \cdot .001}{.98 \cdot .001 + .02 \cdot .999} \approx .047.$$  

We can look at this as saying you have a prior probability of .001 of having the disease, and a positive test increased that to .047, but this number is nothing like .98.
In general, a person often has other symptoms that must be conditioned on, but this result still has real implications on the accuracy of medical tests.

Out of every 1000 people we expect 1 to be infected (.98 infected to also be diagnosed positive), and

\[ 999 \times .02 = 19.98 \approx 20 \]

people to be uninfected but get a positive test result. Thus given a positive test result, there is only about a 1 in 21 chance of being infected (actually .98 in 20.96).

3. This problem is called Gambler’s ruin. It will be helpful to solve a more general instance of this problem. Assume instead that you start with \( 0 \leq i \leq M \) dollars (\( M = 550 \) in our case). Let \( p \) denote the chance of winning each round (.49 in our case) and \( q = 1 - p \). Let \( S_i \) denote the event of getting to \( M \) dollars before going bankrupt when you being with \( i \) dollars. Let \( W \) denote the event of winning your first round. Then we have, for \( 0 < i < M \),

\[
P(S_i) = P(S_i|W)P(W) + P(S_i|W^c)P(W^c)
\]

\[
= P(S_{i+1})P(W) + P(S_{i-1})P(W^c)
\]

\[
= pP(S_{i+1}) + qP(S_{i-1}).
\]

If we let \( s_i = P(S_i) \) we have

\[ s_i = ps_{i+1} + qs_{i-1} \]

for \( 0 < i < M \). As our edge cases we have \( s_0 = 0 \) and \( s_M = 1 \). There are several ways to solve this problem, called a linear homogeneous recurrence equation. There is a general method for solving these equations:

**Theorem 6.** Let \( c_2a_{n+2} + c_1a_{n+1} + c_0a_n = 0 \) for \( n \geq 0 \) where the \( c_i \in \mathbb{R} \) are fixed constants. Let

\[ f(x) = c_2x^2 + c_1x + c_0 = c_2(x - r_1)(x - r_2) \]

be the associated polynomial with roots \( r_1, r_2 \). If \( r_1 \neq r_2 \) then solutions have the form \( a_n = \alpha r_1^n + \beta r_2^n \) for some \( \alpha, \beta \in \mathbb{R} \). Otherwise the solutions have the form \( a_n = (\alpha + \beta n)r_1^n \).

Applying this to our problem we have

\[ ps_{i+1} - s_i + qs_{i-1} = 0 \implies f(x) = px^2 - x + q = p(x - q/p)(x - 1). \]

If \( q \neq p \) then solutions have the form \( s_i = \alpha(q/p)^i + \beta \). To solve for \( \alpha, \beta \) we plug in with \( i = 0, M \):

\[ s_0 = 0 = \alpha + \beta \implies \beta = -\alpha \]

\[ s_M = 1 = \alpha(q/p)^M - \alpha = \alpha((q/p)^M - 1) \implies \alpha = \frac{1}{(q/p)^M - 1}. \]
Thus the solution for $p \neq q$

$$s_i = \frac{(q/p)^i - 1}{(q/p)^M - 1}.$$

If $p = q = 1/2$ then we have $s_i = \alpha + \beta i$. Plugging in gives

$$\begin{align*}
    s_0 &= 0 = \alpha \\
    s_M &= 1 = \beta M \implies \beta = \frac{1}{M}.
\end{align*}$$

Thus if $p = q = 1/2$ then $s_i = \frac{i}{M}$.

In summary, the probability of getting $M$ dollars before going bankrupt when starting with $i$ dollars is

$$s_i = \begin{cases} 
    \frac{(q/p)^i - 1}{(q/p)^M - 1} & \text{if } q \neq p, \\
    \frac{i}{M} & \text{if } q = p.
\end{cases}$$

When $p = .49$, $q = .51$, $M = 550$ and $i = 500$ we obtain .135. A much better strategy to use when you have unfavorable odds is bold play. In bold play you always bet the minimum of the total amount of money you have, and the amount you need to win your desired amount. In our example, the bold strategy would bet 50 dollars, and if it failed, would then bet 100 dollars, and if that failed 200 dollars, and then would have to bet the remaining 150 dollars. The odds given bold play are quite favorable (more than $p + q(p + qp) = .87$).