

The second-ace puzzle

Alice gets two cards from a deck with four cards:
A♠, 2♠, A♥, 2♥.

A♠ A♥	A♠ 2♠	A♠ 2♥
A♥ 2♠	A♥ 2♥	2♠ 2♥

Alice then tells Bob “I have an ace”.

She then says “I have the ace of spades”.

The situation is similar if if Alice says “I have the ace of hearts”.

Puzzle: Why should finding out which particular ace it is raise the conditional probability of Alice having two aces?

The Monty Hall Puzzle

- You're on a game show and given a choice of three doors.
 - Behind one is a car; behind the others are goats.
- You pick door 1.
- Monty Hall opens door 2, which has a goat.
- He then asks you if you still want to take what's behind door 1, or to take what's behind door 3 instead.

Should you switch?

The Second-Child Problem

Suppose that any child is equally likely to be male or female, and that the sex of any one child is independent of the sex of the other. You have an acquaintance and you know he has two children, but you don't know their sexes. Consider the following four cases:

1. You visit the acquaintance, and a boy walks into the room. The acquaintance says "That's my older child."
2. You visit the acquaintance, and a boy walks into the room. The acquaintance says "That's one of my children."
3. The acquaintance lives in a culture, where male children are always introduced first, in descending order of age, and then females are introduced. You visit the acquaintance, who says "Let me introduce you to my children." Then he calls "John [a boy], come here!"
4. You go to a parent-teacher meeting. The principal asks everyone who has at least one son to raise their hands. Your acquaintance does so.

In each case, what is the probability that the acquaintance's second child is a boy?

- The problem is to get the right sample space

Independence

Intuitively, events A and B are independent if they have no effect on each other.

This means that observing A should have no effect on the likelihood we ascribe to B , and similarly, observing B should have no effect on the likelihood we ascribe to A .

Thus, if $\Pr(A) \neq 0$ and $\Pr(B) \neq 0$ and A is independent of B , we would expect

$$\Pr(B|A) = \Pr(B) \text{ and } \Pr(A|B) = \Pr(A).$$

Interestingly, one implies the other.

$\Pr(B|A) = \Pr(B)$ iff $\Pr(A \cap B) / \Pr(A) = \Pr(B)$ iff

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B).$$

Formally, we say A and B are (*probabilistically*) *independent* if

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B).$$

This definition makes sense even if $\Pr(A) = 0$ or $\Pr(B) = 0$.

Bayes' Theorem

Bayes Theorem: Let A_1, \dots, A_n be mutually exclusive and exhaustive events in a sample space S .

- That means $A_1 \cup \dots \cup A_n = S$, and the A_i 's are pairwise disjoint: $A_i \cap A_j = \emptyset$ if $i \neq j$.

Let B be any other event in S . Then

$$\Pr(A_i|B) = \frac{\Pr(A_i) \Pr(B|A_i)}{\sum_{j=1}^n \Pr(A_j) \Pr(B|A_j)}.$$

Proof:

$$B = B \cap S = B \cap (\cup_{j=1}^n A_j) = \cup_{j=1}^n (B \cap A_j).$$

Therefore, $\Pr(B) = \sum_{j=1}^n \Pr(B \cap A_j)$.

Next, observe that $\Pr(B|A_i) = \Pr(A_i \cap B) / \Pr(A_i)$. Thus,

$$\Pr(A_i \cap B) = \Pr(B|A_i) \Pr(A_i).$$

Therefore,

$$\begin{aligned} \Pr(A_i|B) &= \frac{\Pr(A_i \cap B)}{\Pr(B)} \\ &= \frac{\Pr(B|A_i) \Pr(A_i)}{\sum_{j=1}^n \Pr(B \cap A_j)} \\ &= \frac{\Pr(B|A_i) \Pr(A_i)}{\sum_{j=1}^n \Pr(B|A_j) \Pr(A_j)} \end{aligned}$$

Example

In a certain county, 60% of registered voters are Republicans, 30% are Democrats, and 10% are Independents. 40% of Republicans oppose increased military spending, while 65% of the Democrats and 55% of the Independents oppose it. A registered voter writes a letter to the county paper, arguing against increased military spending. What is the probability that this voter is a Democrat?

$S = \{\text{registered voters}\}$

$A_1 = \{\text{registered Republicans}\}$

$A_2 = \{\text{registered Democrats}\}$

$A_3 = \{\text{registered independents}\}$

$B = \{\text{voters who oppose increased military spending}\}$

We want to know $\Pr(A_2|B)$.

We have

$$\Pr(A_1) = .6 \quad \Pr(A_2) = .3 \quad \Pr(A_3) = .1$$

$$\Pr(B|A_1) = .4 \quad \Pr(B|A_2) = .65 \quad \Pr(B|A_3) = .55$$

Using Bayes' Theorem, we have:

$$\begin{aligned}\Pr(A_2|B) &= \frac{\Pr(B|A_2) \times \Pr(A_2)}{\Pr(B|A_1) \times \Pr(A_1) + \Pr(B|A_2) \times \Pr(A_2) + \Pr(B|A_3) \times \Pr(A_3)} \\ &= \frac{.65 \times .3}{(.4 \times .6) + (.65 \times .3) + (.55 \times .1)} \\ &= \frac{.195}{.49} \\ &\approx .398\end{aligned}$$

AIDS

Suppose we have a test that is 99% effective against AIDS. Suppose we also know that .3% of the population has AIDS. What is the probability that you have AIDS if you test positive?

$S = \{\text{all people}\}$ (in North America??)

$A_1 = \{\text{people with AIDS}\}$

$A_2 = \{\text{people who don't have AIDS}\}$ ($A_2 = \overline{A_1}$)

$B = \{\text{people who test positive}\}$

$$\Pr(A_1) = .003 \quad \Pr(A_2) = .997$$

Since the test is 99% effective:

$$\Pr(B|A_1) = .99 \quad \Pr(B|A_2) = .01$$

Using Bayes' Theorem again:

$$\begin{aligned} \Pr(A_1|B) &= \frac{.99 \times .003}{(.99 \times .003) + (.01 \times .997)} \\ &\approx \frac{.003}{.003 + .01} \\ &\approx .23 \end{aligned}$$

Averaging and Expectation

Suppose you toss a coin that's biased towards heads ($\Pr(\text{heads}) = 2/3$) twice. How many heads do you expect to get?

- In mathematics-speak:
What's the *expected number* of heads?

What about if you toss the coin k times?

What's the average weight of the people in this classroom?

- That's easy: add the weights and divide by the number of people in the class.

But what about if I tell you I'm going to toss a coin to determine which person in the class I'm going to choose; if it lands heads, I'll choose someone at random from the first aisle, and otherwise I'll choose someone at random from the last aisle.

- What's the expected weight?

Averaging makes sense if you use an equiprobable distribution; in general, we need to talk about *expectation*.

Random Variables

To deal with expectation, we formally associate with every element of a sample space a real number.

Definition: A *random variable* on sample space S is a function from S to the real numbers.

Example: Suppose we toss a biased coin ($\Pr(h) = 2/3$) twice. The sample space is:

- hh - Probability $4/9$
- ht - Probability $2/9$
- th - Probability $2/9$
- tt - Probability $1/9$

If we're interested in the number of heads, we would consider a random variable $\#H$ that counts the number of heads in each sequence:

$$\#H(hh) = 2; \quad \#H(ht) = \#H(th) = 1; \quad \#H(tt) = 0$$

Example: If we're interested in weights of people in the class, the sample space is people in the class, and we could have a random variable that associates with each person his or her weight.

Probability Distributions

If X is a random variable on sample space S , then the probability that X takes on the value c is

$$\Pr(X = c) = \Pr(\{s \in S \mid X(s) = c\})$$

Similarly,

$$\Pr(X \leq c) = \Pr(\{s \in S \mid X(s) \leq c\}).$$

This makes sense since the range of X is the real numbers.

Example: In the coin example,

$$\Pr(\#H = 2) = 4/9 \text{ and } \Pr(\#H \leq 1) = 5/9$$

Given a probability measure \Pr on a sample space S and a random variable X , the *probability distribution* associated with X is $f_X(x) = \Pr(X = x)$.

- f_X is a probability measure on the real numbers.

The *cumulative distribution* associated with X is $F_X(x) = \Pr(X \leq x)$.

An Example With Dice

Suppose S is the sample space corresponding to tossing a pair of fair dice: $\{(i, j) \mid 1 \leq i, j \leq 6\}$.

Let X be the random variable that gives the sum:

- $X(i, j) = i + j$

$$f_X(2) = \Pr(X = 2) = \Pr(\{(1, 1)\}) = 1/36$$

$$f_X(3) = \Pr(X = 3) = \Pr(\{(1, 2), (2, 1)\}) = 2/36$$

⋮

$$f_X(7) = \Pr(X = 7) = \Pr(\{(1, 6), (2, 5), \dots, (6, 1)\}) = 6/36$$

⋮

$$f_X(12) = \Pr(X = 12) = \Pr(\{(6, 6)\}) = 1/36$$

Can similarly compute the cumulative distribution:

$$F_X(2) = f_X(2) = 1/36$$

$$F_X(3) = f_X(2) + f_X(3) = 3/36$$

⋮

$$F_X(12) = 1$$

The Finite Uniform Distribution

The finite uniform distribution is an equiprobable distribution. If $S = \{x_1, \dots, x_n\}$, where $x_1 < x_2 < \dots < x_n$, then:

$$f(x_k) = 1/n$$

$$F(x_k) = k/n$$

The Binomial Distribution

Suppose there is an experiment with probability p of success and thus probability $q = 1 - p$ of failure.

- For example, consider tossing a biased coin, where $\Pr(h) = p$. Getting “heads” is success, and getting tails is failure.

Suppose the experiment is repeated independently n times.

- For example, the coin is tossed n times.

This is called a sequence of *Bernoulli trials*.

Key features:

- Only two possibilities: success or failure.
- Probability of success does not change from trial to trial.
- The trials are independent.

What is the probability of k successes in n trials?

Suppose $n = 5$ and $k = 3$. How many sequences of 5 coin tosses have exactly three heads?

- $hhhtt$

- $hthht$

- $hthth$

⋮

$C(5, 3)$ such sequences!

What is the probability of each one?

$$p^3(1 - p)^2$$

Therefore, probability is $C(5, 3)p^3(1 - p)^2$.

Let $B_{n,p}(k)$ be the probability of getting k successes in n Bernoulli trials with probability p of success.

$$B_{n,p}(k) = C(n, k)p^k(1 - p)^{n-k}$$

Not surprisingly, $B_{n,p}$ is called the *Binomial Distribution*.

Some Examples

Example 1: A fair die is rolled. Let X denote the number that shows up. What is the probability distribution of $Y = X^2$?

$$\begin{aligned}\{s : Y(s) = k\} &= \{s : X^2(s) = k\} \\ &= \{s : X(s) = -\sqrt{k}\} \cup \{s : X(s) = \sqrt{k}\}.\end{aligned}$$

Conclusion: $f_Y(k) = f_X(\sqrt{k}) + f_X(-\sqrt{k})$.

So $f_Y(1) = f_Y(4) = f_Y(9) = \cdots = f_Y(36) = 1/6$.

$f_Y(k) = 0$ if $k \notin \{1, 4, 9, 16, 25, 36\}$.

Example 2: A coin is flipped. Let X be 1 if the coin shows H and -1 if T . Let $Y = X^2$.

- In this case $Y \equiv 1$, so $\Pr(Y = 1) = 1$.

Example 3: If two dice are rolled, let X be the number that comes up on the first dice, and Y the number that comes up on the second.

- Formally, $X((i, j)) = i$, $Y((i, j)) = j$.

The random variable $X + Y$ is the total number showing.

Example 4: Suppose we toss a biased coin n times (more generally, we perform n Bernoulli trials). Let X_k describe the outcome of the k th coin toss: $X_k = 1$ if the k th coin toss is heads, and 0 otherwise.

How do we formalize this?

- What's the sample space?

Notice that $\sum_{k=1}^n X_k$ describes the number of successes of n Bernoulli trials.

- If the probability of a single success is p , then $\sum_{k=1}^n X_k$ has distribution $B_{n,p}$
 - The binomial distribution is the sum of Bernoullis

Independent random variables

In a roll of two dice, let X and Y record the numbers on the first and second die respectively.

- What can you say about the events $X = 3, Y = 2$?
- What about $X = i$ and $Y = j$?

Definition: The random variables X and Y are independent if for every x and y the events $X = x$ and $Y = y$ are independent.

Example: X and Y above are independent.

Definition: The random variables X_1, X_2, \dots, X_n are *mutually independent* if, for every x_1, x_2, \dots, x_n

$$\Pr(X_1 = x_1, \dots, X_n = x_n) = \Pr(X_1 = x_1) \dots \Pr(X_n = x_n)$$

Example: X_k , the success indicators in n Bernoulli trials, are independent.

Expected Value

Suppose we toss a biased coin, with $\Pr(h) = 2/3$. If the coin lands heads, you get \$1; if the coin lands tails, you get \$3. What are your expected winnings?

- $2/3$ of the time you get \$1;
 $1/3$ of the time you get \$3
- $(2/3 \times 1) + (1/3 \times 3) = 5/3$

What's a good way to think about this? We have a random variable W (for winnings):

- $W(h) = 1$
- $W(t) = 3$

The expectation of W is

$$\begin{aligned} E(W) &= \Pr(h)W(h) + \Pr(t)W(t) \\ &= \Pr(W = 1) \times 1 + \Pr(W = 3) \times 3 \end{aligned}$$

More generally, the *expected value* of random variable X on sample space S is

$$E(X) = \sum_x x \Pr(X = x)$$

An equivalent definition:

$$E(X) = \sum_{s \in S} X(s) \Pr(s)$$

Example: What is the expected count when two dice are rolled?

Let X be the count:

$$\begin{aligned} & E(X) \\ &= \sum_{i=2}^{12} i \Pr(X = i) \\ &= 2\frac{1}{36} + 3\frac{2}{36} + 4\frac{3}{36} + \cdots + 7\frac{6}{36} + \cdots + 12\frac{1}{36} \\ &= \frac{252}{36} \\ &= 7 \end{aligned}$$