Chapter 11: Taking Chances

Many daily decisions are based on probability

e.g., we might decide that, depending on whether it rains, we'll go to a movie or a picnic

> But we have to pack the picnic gear, and buy food or buy tickets for a popular show well in advance of the event
>
> So we need to know how likely it is that it will rain, how much we like movies vs. picnics, etc.

Examples of probability in real-life circumstances:

- Fred is playing roulette in a Macau casino. The roulette wheel has 36 numbers (ignoring the zero), of which half are red and half are black. Fred reasons as follows: In the last ten spins, all the winning numbers have been red. But on average, only half the winning numbers are red. So to even things out, there must be more black numbers than red numbers coming up. So I stand a better chance of winning if I bet on black this time.

- The chance of the 6/49 numbers being exactly the same two days in a row is extremely small. So to maximize my chances of winning today, I should not choose last week's winning numbers.
- Suppose I am at the Metrotown bus terminus, waiting for the number 10 bus to leave for City Hall. The number 10 leaves from here every 8 minutes. So the longer I wait for the bus, the higher the probability that it will leave in the next minute.
- A city has a crackdown on speeding drivers, and the number of traffic fatalities falls by 12%. The local government claims that the increased enforcement has saved lives. But the crackdown was started because of a sudden increase in traffic fatalities the prior year. After an unusually high value, the number of deaths is likely to fall the following year anyway. So there is no reason to think that the crackdown caused the decrease in fatalities.
- "BALTIMORE (AP) A Maryland woman this week gave birth to triplets for the second time in less than two years, defying odds of about one in 50 million, hospital officials said." The reasoning here is that since only about one birth in seven thousand is of triplets, the odds of having two sets of triplets in a row is about one in $7000 \times 7000 = 49{,}000{,}000$, which is about one in 50 million.

The Law of Large Numbers: If a trial--e.g. tossing a coin--is repeated many, many times, the average will get closer and closer to the true value--e.g. 50% heads

Gambler's Fallacy (misunderstanding the Law of Large Numbers)**:** The belief that probability is a force that brings "deviant" (lucky or unlucky) sequences back to the average—a "law of averages"

e.g., I've rolled three sixes in *Risk*; so to prevent me from "unbalancing" the true proportion, the dice will not give sixes for awhile

e.g., if numbers haven't come up for a few lotteries, they are "due" to come up now

Strange things happen: Even things with a very low probability happen: e.g., people do win the Lotto Max, even though the odds are one in 85 million.

If enough coins are tossed, one will come up heads 19 times in a row

(1 : 524,288 chance)

To see this: Fill a truck with a million pennies, dump them. Keep the approximately 500,000 that came up heads, and dump them; keep the approximately 250,000 that came up heads, and dump them, and so on.

There's a very good chance of having at least one coin in the 19th batch

Financial advice and strange things

Heuristic: "a general strategy for solving a problem or coming to a decision" (279)

A heuristic won't always solve the problem, but often solves it well enough for most uses, and usually more quickly than a perfect solution would.

Some examples of common heuristics:

If the water in one glass is higher than another, there's more water in the first glass (young children)

If someone is taller they're smarter (young children)

When in doubt, push a pawn (chess rule)

Better safe than sorry

The Representativeness heuristic: People tend to judge the probability of an event by assimilating it to a more common, “representative” event and assuming that the probabilities will be similar.

e.g., consider the following two sequences of 13 coin tosses:

A: H T H H T T T H T H T T H

B: T T T T T T T H H H H H H

Which is more probable? Why?

“A” closely resembles the typical, “representative” jumble of H’s and T’s we usually get, whereas “B” seems to resemble a carefully organized series of heads and tails, not something that should happen at random

The tale of Linda:

“Linda is 31, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations”

Given the above, which of the following are more probable?

- Linda is an elementary school teacher
- Linda works in a bookstore and takes yoga classes.
- Linda is active in the feminist movement.
- Linda is a psychiatric social worker.
- Linda is a member of the League of Women Voters.
- Linda is a bank teller.
- Linda is an insurance salesperson.
- Linda is a bank teller and is active in the feminist movement.

Most tended to think it is least likely that Linda was a bank teller

In fact, most believed it more likely that Linda is *both* a bank teller and an active feminist than that she was simply a bank teller

This belief, however, amounts to the “conjunction fallacy,” the belief that the probability of both of two, separate events occurring can be greater than the probability of either separately

Explanation: We believe that the conjunction says something “representative” about the Linda we know, but the simple claim that she is a bank teller seems at odds with this

Why do we make this “mistake”?

Prototypes: We often identify things by comparing them with well-known examples

e.g., to decide whether an animal is a bear, we compare it to a classic image of a bear

In the “Linda” case, we judge which of the list is “more like” the classic Linda

The AVAILABILITY HEURISTIC: People judge probability using available facts, and these may not be representative

the _ _ _ _ *i n g* vs. _ _ _ _ _ *n* _ example

Schwartz (1991) asked subjects for six examples when they had been assertive (most could think of six). He then asked other people for twelve examples, which only a few people could do. He then asked both groups how assertive they were.

The ‘six’ people seemed to score themselves as more assertive than average, because their available data had a greater *proportion* of their being assertive.

Most people believe they are in the safest 25% of drivers.

The point: Our brains seem to be programmed with simple rules (heuristics) for calculating probability, but these rules can easily lead us astray in non-standard situations

So, we should only trust these heuristics in fairly simple situations, but use more advanced probability techniques in non-standard cases

PROBABILITY SCALE: If there's no chance of something, X, happening, we say that X has zero probability

If something is guaranteed to happen, we say X has a 100 per cent probability (in terms of fractions of a 100), or a probability of 1 (on the decimal scale)

A priori probability

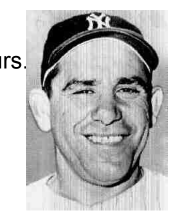
People's intuitive probability calculations are often very inaccurate

Yogi Berra: Success is 1/3 talent + 75% hard work

a priori method: we assume that the probability e.g. of any card in a deck's being drawn is the same

We know that a standard card deck has 52 cards, with 13 different card types, each with four suits; so there are four 5-cards in 52 cards

Yogi Berra-isms (http://en.wikiquote.org/wiki/Yogi_Berra)



Always go to other peoples' funerals; otherwise they won't go to yours.

Half the lies they tell me aren't true.

I'd give my right arm to be ambidextrous.

I'd find the fellow who lost it, and, if he was poor, I'd return it.

(When asked what he would do if he found a million dollars)

If people don't want to come out to the ballpark, nobody's going to stop them.

It ain't over 'til it's over.

It's like déjà vu all over again.

It's tough to make predictions, especially about the future.

Ninety percent of this game is mental, and the other half is physical.

Slump? I ain't in no slump! I just ain't hitting.

The future ain't what it used to be.

(At a dinner in an Italian restaurant, when asked how many slices should be cut in Yogi's pizza.)

You better make it four. I don't think I could eat eight.

Hence, if the chance of getting any specific 5-card, say a 5 ♥ is 1/52, the chance of getting any one of the four suits (♥ or ♦ or ♣ or ♠) is:

$P(5) = 4 \times 1/52 = 1/13 \approx 0.07692307692307693$

The *a priori* approach uses plausible assumptions, together with reasoning (math) to determine probabilities

General rule: We can describe any occurrence as a *hypothesis* that the event occurs: e.g., the event that heads comes up means that "the *hypothesis that heads comes up* is true"

If we suppose all outcomes to be equally probable (equiprobable), then:

$Pr(h)$ = favourable outcomes/total outcomes

Tossing a coin: let h = Tails comes up; then

$Pr(h)$ (assuming a "fair" coin) = ?

Tossing a pair of dice: let h = six comes up; then $Pr(h)$ (assuming "fair" dice) = ? (p284)

Using two dice, what is:

$Pr(5)$?

Which number has the highest probability of being thrown?

Its probability?

$Pr(11)$?

$Pr(7 \vee 11)$?

Which is more likely, $Pr(5)$ or $Pr(8)$?

More likely, $Pr(5 \vee 8)$ or $Pr(2 \vee 7)$?

Laws of probability: It is guaranteed that either an event will occur or it won't, and this covers all possible events: e.g. the probability that it will rain or that it won't is 100 %

Suppose the probability that it will rain is 60%; what's the chance it won't rain?

In general terms, $\Pr(h) + \Pr(\text{not-}h) = 1$

Rule 1: The probability that an event will not occur is 1 minus the probability that it will occur; or, $\Pr(\text{not } h) = 1 - \Pr(h)$

What's the probability of *not* tossing a "6", using one die?

Independent events: Two events are *independent*, if the occurrence of one event does not affect the probability that that other event will occur

e.g. drawing an ace from a shuffled deck, then shuffling it back in

Getting another ace is independent of what you got on the previous draw

Rule 2: Given two *independent* events, the probability of both occurring is the product of their individual probabilities.

$$\Pr(h_1 \,\&\, h_2) = \Pr(h_1) \times \Pr(h_2)$$

Note: In Rule 2, the conjoined events occur in a *particular order = permutation*

In other words, $\Pr(h_1 \,\&\, h_2) = \Pr(h_1 \textbf{ then } h_2) = \Pr(h_1) \times \Pr(h_2)$

If we want to know the probability of all *combinations* of h_1 and h_2, whatever their order, we will have to **add** the probability of each permutation

Conditional events: Those whose probability of occurring *IS* affected by the occurrence of other events

Pr(Ace) before any cards are drawn?

Suppose you remove an ace and keep it out of the deck: What's the probability that you will get another ace on your next draw?

Rule 2G. Given two events, the probability of their both occurring is the probability of the first times the probability of the second, given that the first has occurred.

$$\Pr(h_1 \,\&\, h_2) = \Pr(h_1) \times \Pr(h_2/h_1)$$

Without reshuffling, $\Pr(A \,\&\, A) = 1/13 \times \Pr(A_{on2nddraw}/A_{on1stdraw}) = 1/13\times3/51 = 3/663$

Suppose we want to determine the probability of either of two (or more) events occurring, where no two of these events can occur at the same time (i.e., if the events are *mutually exclusive*)?

E.g., what's the probability of getting either a total of 5 or 6, on a toss of two dice?

Favourable outcomes/total outcomes = ?/36

There are four ways we can total 5, and 5 ways to reach 6, and any of these is favourable; so

$$= 4/36 + 5/36 = 9/36 = 1/4 = .25$$

RULE 3: The probability that at least one of two **mutually exclusive** events will occur is the sum of the probabilities that each of them will occur.

$$\Pr(h_1 \text{ or } h_2) = \Pr(h_1) + \Pr(h_2)$$

If events are not mutually exclusive, some events will be both h_1 and h_2.

These "double" events will be counted twice: once in $\Pr(h_1)$ and once in $\Pr(h_2)$

So to get the true probability of $\Pr(h_1 \text{ or } h_2)$ in such cases, we need to subtract the twice-counted cases. We get:

RULE 3G: $\Pr(h_1 \text{ or } h_2) = \Pr(h_1) + \Pr(h_2) - \Pr(h_1 \,\&\, h_2)$

E.g., what's the probability that someone in this class is either male or more than 5' 5"?

When h_1 and h_2 are mutually exclusive, $\Pr(h_1 \& h_2) = 0$, as should be, by Rule 3

Problem: What's the probability of drawing *at least one* spade in a series of 16 consecutive draws, where the card is returned to the deck after each draw and the deck is reshuffled?

You might think as follows: On each draw, there's a 1/4 chance of spades; so after 4 draws, you would have a 100% chance of spades.

But if so, after 16 draws, you'd have a 400% chance of a spade?!

This answer can't be right, since

(a) the **maximum** probability is 100%, and, similarly,

(b) it's **possible** that you get *no* spades after 16 draws.

Why did it *seem* right at first? We seem to be using Rule 3:

$\Pr(♠_1 \vee ♠_2 \vee \ldots \vee ♠_{16}) = \Pr(♠_1) + \Pr(♠_2) + \ldots + \Pr(♠_{16}) = 4???$

Recall, though, that **Rule 3** only works if the probabilities are *mutually exclusive*—in other words, if you get a spade on one draw, then you don't get one on another draw

However, its perfectly possible to get spades on any, all or none of the draws:

e.g. at least one spade in two draws: $\Pr[(♠\&♠) \vee (♠\&\text{not-}♠) \vee (\text{not-}♠\&♠)] = 7/16$

Writing out the formula using **Rule 3** in this way gets enormously complicated, as the number of draws increase

It's better to reformulate the question: We know that the probability of at least one spade has to be less than one, and we have a simple way of calculating the probability that we get *no* spades on 16 draws

$\Pr(\text{no spades}) = \Pr(\sim♠_1 \ \& \sim♠_2 \ \& \ldots \& \sim♠_{16})$

Since no draw is correlated (shuffled deck)

$$= (3/4 \times 3/4 \times \ldots \times 3/4) \text{ (16 times)} = (3/4)^{16} = 43046721/4294967296 = .01$$

Because Pr(at least one spade) = 1- Pr(no spades) = 1- .01 = .99

or *very nearly* 1.

RULE 4: The probability that an event will occur ***at least once*** in a series of *n* independent trials is simply 1 minus the probability that it will *not* occur in that number of trials.

$$1 - \Pr(\text{not } h)^n$$

Problem similar to exam problem

p291: "You are presented with two bags, one containing two ham sandwiches and the other containing a ham sandwich and a cheese sandwich. You reach in one bag and draw out a ham sandwich. What is the probability that the other sandwich in the bag is also a ham sandwich?"

> You might think as follows: There are two bags with a ham sandwich in them. I could have selected either of them with an equal probability of 50%.
>
> > So there's a 50% chance the next sandwich is ham (or cheese)

However, the situation is under-described.

The probability that the next sandwich is ham depends on which bag you've chosen, and given that you've drawn a ham sandwich on your first draw provides some information about which bag you've selected

What is the probability that I've selected the first bag, given that I've selected a ham sandwich? the second bag?

Pr(ham_next) = Pr(ham_next/bag1) x Pr(bag1/ham_first) + Pr(ham_next/bag2) x Pr (bag2/ham_first)

= 1 x 2/3 + 0 x 1/3 = **2/3**

BAYES'S THEOREM

Consider Wendy, who has tested positive for colon cancer

Because therapy is painful and dangerous, the doctor wants to know how likely it is that Wendy truly has colon cancer.

Previous studies have shown:

(1) The probability that someone in the larger population has colon cancer, **Pr(C)**= .003; so

(2) the probability that someone doesn't have colon cancer, **Pr(~C)** = **1- Pr (C)** = .997

(3) If a person has colon cancer, then the probability of a positive test, **Pr(P/C)** = 0.9

(4) If a person doesn't have colon cancer, the probability of a positive test, **Pr (P/~C)** = .03

What is the probability that Wendy has colon cancer, given that she has tested positive = **P(C/P)**?

The most common estimate by doctors and others is well above 50%

To find the answer to this difficult problem we try to determine: #"favourable"/ total outcomes = #true positives/[true positives + false positives]

Let's suppose the larger population's size is 100,000 (to deal with the decimal places) By (1), Pr(C) = .003; so, 300 of this population (of 100,000) has colon cancer

by (2), Pr(P/C) = 90% = .9 of those with colon cancer (300) will test positive

True positives = Pr(C) × Pr(P/C) = 300 × (.9) = 270 true positives

Since Pr(~C) = .997, in a 100,000 people, we get 99,700 who don't have colon cancer

From (3), Pr(P/~C) = .03

So, of those who don't have colon cancer, 99,700 × .03 = 2991 will test positive, → 2991 false positives

Thus: #true pos/[true pos + false pos] = 270/(270 + 2991) = 270/3261 = .083 or 8.3%

If we remove the 100,000 multiplier (common to both numerator and denominator), we get:

$$P(C/P) = \frac{Pr(C) \times Pr\,(P/C)}{[Pr(C) \times Pr\,(P/C)] + [Pr(\sim C) \times Pr\,(P/\sim C)]}$$

Which is a simple form of Bayes's theorem.

More generally, Bayes's theorem is used to find the probability that a hypothesis is true, given the evidence = Pr(h/e).

For Wendy's case, the hypothesis, h, is that Wendy has colon cancer (C)

The "evidence," e, is the positive test result (P)

General form of Bayes's Theorem:

$$\Pr(h/e) = \frac{\Pr(h) \times \Pr(e/h)}{[\Pr(h) \times \Pr(e/h)] + [\Pr(\sim h) \times \Pr(e/\sim h)]}$$

Bayes's Theorem is useful for determining Pr(h/e), but only when we also know Pr(h), Pr(e/h) and Pr(e/~h)

Putting the information into a table:

	Colon cancer	No Colon Cancer	Total
Test positive	270	?	?
Don't test positive	30	?	?
Total	300	99700	100000

EXERCISES: IN CLASS WORKSHEET