Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 7

Basic Probability – Part III

Recall from Last Time: Conditional Probability

Set of relevant outcomes restricted by condition

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) > 0$$

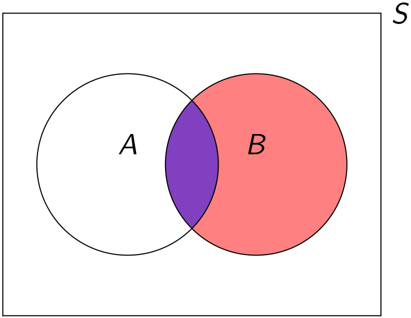


Figure: B becomes the "new sample space" so we need to re-scale by $P(B)$ to keep probabilities between zero and one.

Independence and The Multiplication Rule

The Multiplication Rule

Just rearrange the definition of conditional probability:

$P(A \cap B) = P(A|B)P(B)$

Statistical Independence

$P(A \cap B) = P(A)P(B)$

By the Multiplication Rule

Independence $\iff P(A|B) = P(A)$

Interpreting Independence

Knowing that B has occurred doesn't give us any additional information about whether A will.

Will Having 5 Children Guarantee a Boy?

A couple plans to have five children. Assuming that each birth is independent and male and female children are equally likely, what is the probability that they have at least one boy?

By Independence and the Complement Rule,

$$
\begin{aligned}
P(\text{no boys}) &= P(\text{5 girls}) \\
&= 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \\
&= 1/32
\end{aligned}
$$

$$
\begin{aligned}
P(\text{at least 1 boy}) &= 1 - P(\text{no boys}) \\
&= 1 - 1/32 = 31/32 = 0.97
\end{aligned}
$$

The Law of Total Probability

If $E_1, E_2, \ldots, E_k$ are mutually exclusive, collectively exhaustive events and A is another event, then

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \ldots + P(A|E_k)P(E_k)$$

Deriving the Law of Total Probability For $k = 2$

Since $A \cap B$ and $A \cap B^c$ are mutually exclusive and their union equals A,

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

But by the multiplication rule:

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \\ P(A \cap B^c) &= P(A|B^c)P(B^c) \end{aligned}$$

Combining,

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

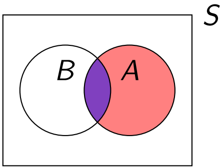


Figure: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

Connecticut, 2006 – Democratic Primary for Senate Race

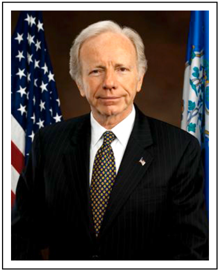


Figure: Former Senator Joe Lieberman and never-Senator Ned Lamont. In 2006 Lieberman, a popular Democratic incumbant, faced a primary challenge from Lamont, because of the former's support for George W. Bush's foreign policy agenda.

How do prediction markets work?

To learn more, see $\boxed{\text{Wolfers \& Zitzewitz (2004)}}$

> THIS CERTIFICATE ENTITLES THE BEARER TO \$10 IF DUKE WINS THE FINAL FOUR IN 2016.

Buyers – Purchase Right to Collect

Duke very likely to win $\Rightarrow$ buy for close to \$10.

Duke very unlikely to win $\Rightarrow$ buy for close to \$0.

Sellers – Sell Obligation to Pay

Duke very likely to win $\Rightarrow$ sell for close to \$10.

Duke very unlikely to win $\Rightarrow$ sell for close to \$0.

Going price of contract encodes market participants' beliefs in the form of probability:

$$\text{Price}/\$10 \approx \text{Subjective Probability}$$

Summer 2006 – Two Contracts on Tradesports

What is the missing probability?

Contract – Pays $10 if Event Occurs	Price	Prob.
Lieberman Wins CT Dem. Senate Primary	$3.20	0.32
Neither Dem. nor Rep. wins CT Senate Election	$3.01	0.3

But the real question of interest was whether Lieberman would keep his Senate seat. Can we calculate the implied probability?

$K =$ Joe Lieberman keeps his Senate seat

$N =$ Neither Dem. nor Rep. wins Senate seat (Tradesports Prob. 0.3)

$J =$ Joe Lieberman wins the Dem. Primary (Tradesports Prob. 0.32)

Law of Total Probability

$$\begin{aligned} P(K) &= P(K|J)P(J) + P(K|J^c)P(J^c) \\ &= P(K|J) \times 0.32 + P(K|J^c) \times (1 - 0.32) \end{aligned}$$

No serious Republican Challengers $\Rightarrow P(K|J) \approx 1$

$$P(K) \approx 0.32 + P(K|J^c) \times 0.68$$

K = Joe Lieberman keeps his Senate seat

N = Neither Dem. nor Rep. wins Senate seat (Tradesports Prob. 0.3)

J = Joe Lieberman wins the Dem. Primary (Tradesports Prob. 0.32)

Law of Total Probability Again

$$\begin{aligned} P(N) &= P(N|J)P(J) + P(N|J^c)P(J^c) \\ 0.3 &= P(N|J) \times 0.32 + P(N|J^c) \times 0.68 \end{aligned}$$

No serious Independent challengers $\Rightarrow P(N|J) \approx 0$

$$\begin{aligned} 0.3 &\approx P(N|J^c) \times 0.68 \\ P(N|J^c) &\approx 0.44 \end{aligned}$$

K = Joe Lieberman keeps his Senate seat

N = Neither Dem. nor Rep. wins Senate seat (Tradesports Prob. 0.3)

J = Joe Lieberman wins the Dem. Primary (Tradesports Prob. 0.32)

Since Lieberman will run as an Independent if he loses the primary and there are no serious Independent challengers:

$$P(K|J^c) \approx P(N|J^c) \approx 0.44$$

We *can* calculate the missing probability!

K = Joe Lieberman keeps his Senate seat
N = Neither Dem. nor Rep. wins Senate seat (Tradesports Prob. 0.3)
J = Joe Lieberman wins the Dem. Primary (Tradesports Prob. 0.32)

Combining the Tradesports prices with the assumptions of no serious Republican or Independent Challengers:

$$
\begin{aligned}
P(K) &= P(K|J)P(J) + P(K|J^c)P(J^c) \\
&\approx 0.32 + P(K|J^c) \times 0.68 \\
&\approx 0.32 + 0.44 \times 0.68 \\
&\approx 0.62
\end{aligned}
$$

Wait, isn't that just $0.3 + 0.32$?

$K =$ Joe Lieberman keeps his Senate seat

$N =$ Neither Dem. nor Rep. wins Senate seat (Tradesports Prob. 0.3)

$J =$ Joe Lieberman wins the Dem. Primary (Tradesports Prob. 0.32)

You may have noticed that the answer we ended up with was 0.62, which is *identical* to simply adding up the Tradesports probabilities of N and J.

Essentially, the assumptions we made are equivalent to assuming that $K = N \cup J$ *and* that N and J are mutually exclusive events. Since the probabilities of mutually exclusive events sum: $P(K) = P(N \cup J) = P(N) + P(J)$.

What Actually Happened?

Ned Lamont won the Democratic Primary with 52% of the vote, but Joe Liberman ran as an independent in the General Election and won. Lieberman retired from the Senate in January 2013.

Why are calculations like this interesting?

Statistical Arbitrage

If the probabilities implied by the prices of prediction market contracts violate any of the rules we have studied in this class, there is a *pure arbitrage opportunity*, i.e. a way make to make a guaranteed, risk-free profit.

A Simple Example of Statistical Arbitrage

Courtesy of Eric Crampton

November 5th, 2012

- \$2.30 for contract paying \$10 if Romney wins on BetFair
- \$6.58 for contract paying \$10 if Obama wins on InTrade

Implied Probabilities

- BetFair: $P(Romney) \approx 0.23$
- InTrade: $P(Obama) \approx 0.66$

What's Wrong with This?

Violates complement rule! $P(Obama) = 1 - P(Romney)$ but the implied probabilities here don't sum up to one!

A Simple Example of Statistical Arbitrage

Courtesy of Eric Crampton

November 5th, 2012

- \$2.30 for contract paying \$10 if Romney wins on BetFair
- \$6.58 for contract paying \$10 if Obama wins on InTrade

Arbitrage Strategy

Buy Equal Numbers of Each

- Cost = \$2.30 + \$6.58 =\$8.88 per pair
- Payout if Romney Wins: \$10
- Payout if Obama Wins: \$10
- Guaranteed Profit: \$10 - \$8.88 = \$1.12 per pair

Four Volunteers Please!

The Lie Detector Problem

From accounting records, we know that 10% of employees in the store are stealing merchandise.

The managers want to fire the thieves, but their only tool in distinguishing is a lie detector test that is 80% accurate:

$$\begin{array}{rcl} \text{Innocent} & \Rightarrow & \text{Pass test with 80\% Probability} \\ \text{Thief} & \Rightarrow & \text{Fail test with 80\% Probability} \end{array}$$

What is the probability that someone is a thief *given* that she has failed the lie detector test?

Monte Carlo Simulation – Roll a 10-sided Die Twice

Managers will split up and visit employees. Employees roll the die twice but keep the results secret!

First Roll – Thief or not?

$0 \Rightarrow$ Thief, $1 - 9 \Rightarrow$ Innocent

Second Roll – Lie Detector Test

$0, 1 \Rightarrow$ Incorrect Test Result, $2 - 9$ Correct Test Result

	0 or 1	2–9
Thief	Pass	Fail
Innocent	Fail	Pass

What percentage of those who failed the test are guilty?

Who Failed Lie Detector Test:

Of Thieves Among Those Who Failed:

Base Rate Fallacy – Failure to Consider Prior Information

Base Rate – Prior Information

Before the test we know that 10% of Employees are stealing.

People tend to focus on the fact that the test is 80% accurate and ignore the fact that only 10% of the employees are theives.

Thief (Y/N), Lie Detector (P/F)

	0	1	2	3	4	5	6	7	8	9
0	YP	YP	YF	YF	YF	YF	YF	YF	YF	YF
1	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
2	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
3	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
4	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
5	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
6	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
7	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
8	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
9	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP

Table: Each outcome in the table is equally likely. The 26 given in red correspond to failing the test, but only 8 of these (YF) correspond to being a thief.

Base Rate of Thievery is 10%

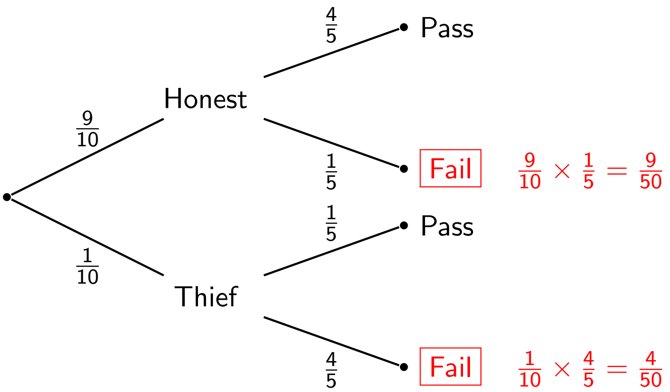


Figure: Although $\frac{9}{50} + \frac{4}{50} = \frac{13}{50}$ fail the test, only $\frac{4/50}{13/50} = \frac{4}{13} \approx 0.31$ are actually theives!

Deriving Bayes' Rule

Intersection is symmetric: $A \cap B = B \cap A$ so $P(A \cap B) = P(B \cap A)$

By the definition of conditional probability,

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

And by the multiplication rule:

$$P(B \cap A) = P(B|A)P(A)$$

Finally, combining these

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

Understanding Bayes' Rule

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

Reversing the Conditioning

Express $P(A|B)$ in terms of $P(B|A)$. *Relative magnitudes* of the two conditional probabilities determined by the ratio $P(A)/P(B)$.

Base Rate

$P(A)$ is called the "base rate" or the "prior probability."

Denominator

Typically, we calculate $P(B)$ using the law of toal probability

In General $P(A|B) \neq P(B|A)$

Question

Most college students are Democrats. Does it follow that most Democrats are college students? (A = YES, B = NO)

Answer

There are many more Democracts than college students:

$$P(\text{Dem}) > P(\text{Student})$$

so $P(\text{Student}|\text{Dem})$ is small even though $P(\text{Dem}|\text{Student})$ is large.

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief}, \; F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$\begin{aligned} P(F) &= P(F|T)P(T) + P(F|T^c)P(T^c) \\ &= 0.8 \times 0.1 + 0.2 \times 0.9 \\ &= 0.08 + 0.18 = 0.26 \end{aligned}$$

$$P(T|F) = \frac{0.08}{0.26} = \frac{8}{26} = \frac{4}{13} \approx 0.31$$

“Odd” Question # 5

There are two kinds of taxis: green cabs and blue cabs. Of all the cabs on the road, 85% are green cabs. On a misty winter night a taxi sideswiped another car and drove off. A witness says it was a blue cab. The witness is tested under conditions like those on the night of the accident, and 80% of the time she correctly reports the color of the cab that is seen. That is, regardless of whether she is shown a blue or a green cab in misty evening light, she gets the color right 80% of the time.

> Given that the witness said she saw a blue cab, what is the probability that a blue cab was the sideswiper?

Solving The Taxi Problem

G = Taxi is Green, $P(G) = 0.85$
B = Taxi is Blue, $P(B) = 0.15$
W_B = Witness says Taxi is Blue, $P(W_B \mid B) = 0.8, P(W_B \mid G) = 0.2$

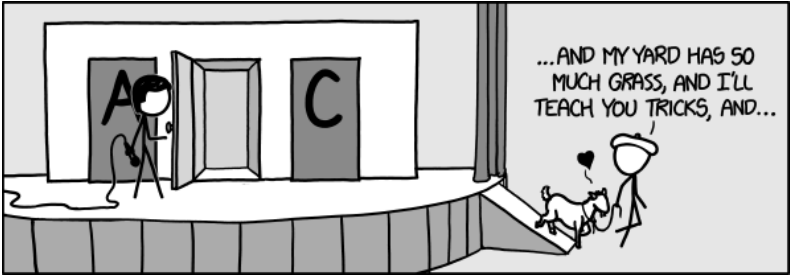
$$P(B \mid W_B) = P(W_B \mid B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B \mid B)P(B) + P(W_B \mid G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \\ &= {\color{red}0.12} + 0.17 = {\color{red}0.29} \end{aligned}$$

$$\begin{aligned} P(B \mid W_B) &= 0.12/0.29 = 12/29 \approx 0.41 \\ P(G \mid W_B) &= 1 - (12/19) \approx 0.59 \end{aligned}$$

The Monty Hall Problem!

What is the probability that you win if you switch?



Key Point – Monte doesn't choose a door randomly: he *always* shows you a goat.

Without loss of generality, suppose you chose door #1