Economics 103 – Statistics for Economists

Minsu Chang

University of Pennsylvania

Lecture # 7

Basic Probability – Part III

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:

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

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:

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

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%

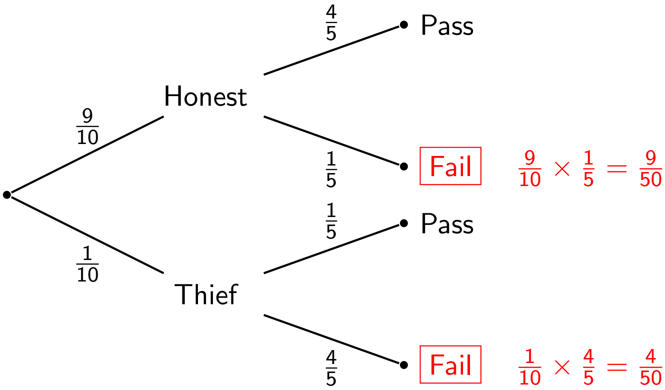

Figure: Although $\frac{9}{50} + \frac{4}{50} = \frac{13}{50}$ fail the test, only $\frac{4/50}{13/50} = \frac{4}{13} \approx 0.31$ are actually theives!

Deriving Bayes' Rule

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

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,

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

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:

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

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)

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

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

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

Solving the Lie Detector Problem with Bayes' Rule

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

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

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

Solving the Lie Detector Problem with Bayes' Rule

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

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

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

Solving the Lie Detector Problem with Bayes' Rule

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

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

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

Solving the Lie Detector Problem with Bayes' Rule

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

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

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

$$P(T|F) = \frac{0.08}{0.26} =$$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$P(W_B) = P(W_B|B)P(B) + P(W_B|G)P(G)$$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B|B)P(B) + P(W_B|G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \end{aligned}$$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

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

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

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

$$P(B|W_B) = 0.12/0.29 = 12/29 \approx 0.41$$

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, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

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

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

Random Variables

Random Variables

A random variable is neither random nor a variable.

Random Variable (RV): X

A *fixed* function that assigns a *number* to each basic outcome of a random experimnet.

Realization: x

A particular numeric value that an RV could take on. We write $\{X = x\}$ to refer to the *event* that the RV X took on the value x.

Support Set (aka Support)

The set of all possible realizations of a RV.

Random Variables (continued)

Notation

Capital latin letters for RVs, e.g. X, Y, Z, and the corresponsing lowercase letters for their realizations, e.g. x, y, z.

Intuition

You can think of an RV as a machine that spits out random numbers: although the machine is deterministic, its inputs, the outcomes of a random experiment, are not.

Example: Coin Flip Random Variable

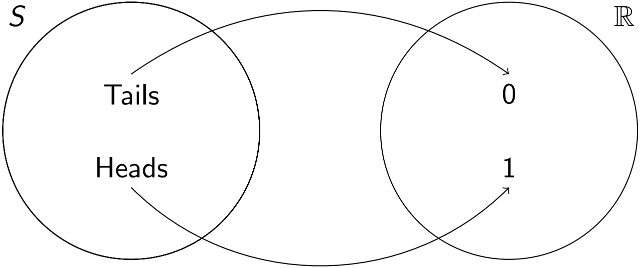

Figure: This random variable assigns numeric values to the random experiment of flipping a fair coin once: Heads is assigned 1 and Tails 0.

Which of these is a realization of the Coin Flip RV?

(a) Tails

(b) 2

(c) 0

(d) Heads

(e) $1/2$

What is the support set of the Coin Flip RV?

(a) $\{\text{Heads}, \text{Tails}\}$

(b) $1/2$

(c) 0

(d) $\{0, 1\}$

(e) 1

Let X denote the Coin Flip RV

What is $P(X = 1)$?

(a) 0

(b) 1

(c) $1/2$

(d) Not enough information to determine

Two Kinds of RVs: Discrete and Continuous

Discrete support set is discrete, e.g. $\{0, 1, 2\}$, $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$

Continuous support set is continuous, e.g. $[-1, 1]$, $\mathbb{R}$.

Start with the discrete case since it's easier, but most of the ideas we learn will carry over to the continuous case.

Discrete Random Variables I

Probability Mass Function (pmf)

A function that gives $P(X = x)$ for any realization x in the support set of a discrete RV X. We use the following notation for the pmf:

$$p(x) = P(X = x)$$

Plug in a realization x, get out a probability $p(x)$.

Probability Mass Function for Coin Flip RV

$$X = \begin{cases} 0, \text{Tails} \\ 1, \text{Heads} \end{cases}$$

$$\begin{aligned} p(0) &= 1/2 \\ p(1) &= 1/2 \end{aligned}$$

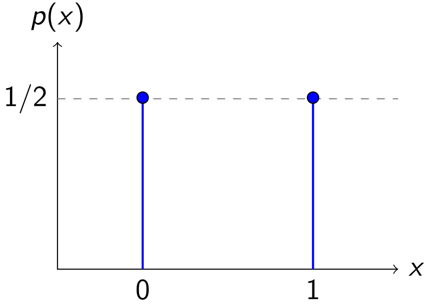

Figure: Plot of pmf for Coin Flip Random Variable

Important Note about Support Sets

Whenever you write down the pmf of a RV, it is crucial to also write down its Support Set. Recall that this is the set of *all possible realizations for a RV*. Outside of the support set, all probabilities are zero. In other words, the pmf is only defined on the support.

Properties of Probability Mass Functions

If $p(x)$ is the pmf of a random variable X, then

(i) $0 \leq p(x) \leq 1$ for all x

(ii) $\displaystyle\sum_{\text{all } x} p(x) = 1$

where "all x" is shorthand for "all x in the support of X."