Lecture 1: Conditional Probability

Jeff Rouder

May, 2024

Follow along at:

https://github.com/rouderj/freiburg24

Consider the dreaded disease **Dipsidoodleitis**. Dipsidoodleitis is detected by a blood test which measures dipsidoodlamines in the blood. Values for people without the disease are normally distributed with $M = 70$ and $sd = 10$. Value for people with the disease are $M = 84$, $sd = 14$. Baserates are know as follows: 10% of people have the disease.

You happen to have a dipsidoodleamine value of 88. What is the probability that you are diseas free?

- This is a workshop on doing, meaning generating, solving, programming
- I will not lecture much

Probability

Two Domains:

- Rules
- Soul (yes, probability has a soul)

Rules of Probability, Kolmogorov Axioms

Let X be the sample space (it'a set). Let $A, B \subseteq X$ be events in X.

1. $Pr(A) \geq 0$,
2. $Pr(X) = 1$, and
3. if $A \cap B = \emptyset$, then $Pr(A \cup B) = Pr(A) + Pr(B)$.

Your Turn

Suppose we flip a coin twice. Here are the probabilities on the outcomes. $P(\{(H,H)\}) = .5$, $P(\{(H,T)\}) = .2$, $P(\{(T,H)\}) = .15$. Use the Kolmogov axioms to find:

1. The probabilities on all events.
2. Probability of at least one tail
3. Probability of a trick coin

Probability

- Rules describe relative weight
- Works say on the weight of my body divided into overlapping parts
- But, we need more

Soul of Probability

Frequentist (Conventional):

- Probability is proportion in the long run
- Probability is a property of the natural world
  - Probability of coin is a property of the coin, much like its radius, weight, or alloy composition.
- Probability cannot be changed
- Probability is only knowable exactly in the long run
- Probability is not placed on unknown values that are fixed, say a parameter or a model.

Soul of Probability:

Bayesian

- Probability is degree of plausibility
- Probability is subjective
- Probability is our mental construct, not part of the world per se
- Probabilities may be placed on parameters, models, individual coin flips, etc.

Updating Probability?

There are three horses in a race. Each has a 1/3rd chance of winning. We are the house and we are going to let people make bets. Each bet costs $1, and if you pick the winning horse, the bet pays $2.90. We make 10 cents per bet. Now, right before betting opens, Horse 3 eats a bad apple, gets sick, and withdraws. How should we adjust probability and bets so that we make 10 cents per bet?

Bayesian Goal:

- Treat beliefs as probabilities
- Update beliefs rationally in light of data using rules of conditional probability

Your Turn

- What is a random variable?

My Turn

Suppose X=(rain,sun,snow) with

- Pr({rain})=.8
- Pr({sun})=.19
- Pr({snow})=.01

Suppose:

- Y({rain})=-1
- Y({sun})=1
- Y({snow})=5

Then

- General: $Pr(Y = y) = Pr(Y^{-1}(y))$
- Example: $Pr(Y = 1) = Pr(\{\text{sun}\}) = .19$
- Example: $Pr(Y > 0) = Pr(\{\text{sun}, \text{snow}\}) = .2$

My Turn

- Technical A random variable maps from one probability space to another.
- A random variable maps from the events in the world into real numbers. Probabilities on numbers correspond to probabilities on events in the world.

Probability Mass function

$$Pr(Y = y) = \begin{cases} .8, & y = -1, \\ .19, & y = 1, \\ .01, & y = 5, \\ 0, & \text{otherwise.} \end{cases}$$

We can also write it as $Pr_Y(y)$ or $Pr(y)$, or sometimes, $f(y)$.

Your Turn

- The above random variables were discrete.
- What is a continuous random variable?
- What is a density function?
- Draw an example for weight of people
- What is $\Pr(W{=}60\text{kg})$?
- What is the unit on the y-axis?

Joint Probability

- Foundational, Starting Point
- Example: Two coin flips.
  - $X_1 = 0, 1$, first flip
  - $X_2 = 0, 1$, second flip
- Joint
  - $Pr(X_1 = x_1 \cap X_2 = x_2) = Pr(X_1 = x_1, X_2 = x_2) = Pr(x_1, x_2)$
  - $Pr(0, 0) = .2$
  - $Pr(0, 1) = .3$
  - $Pr(1, 0) = .4$
  - $Pr(1, 1) = .1$
- How many "free" probabilities here? 3
- Why, they sum to 1.0

Your Turn

- $Pr(X_1 = 0)$, $Pr(X_1 = 1)$, $Pr(X_2 = 0)$, $Pr(X_2 = 1)$ are called marginal probabilities
- What is Pr(X_1=0)?
- How do you know this? What is the Law?

Law of Total Probability

Suppose there is a sequence of sets $B_1, B_2, \ldots, B_J$ that partitions sample space X ($X \subseteq \bigcup_j B_j$, $B_i \cap B_j = \emptyset,\ i \neq j$)

$$Pr(A) = Pr(A \cap B_1) + \ldots + Pr(A \cap B_J) = \sum_j Pr(A \cap B_j)$$

- $Pr(X_1 = 0) = Pr(X_1 = 0, X_2 = 0) + Pr(X_1 = 0, X_2 = 1) = .2 + .3 = .5$

Your turn

- Consider marginals $Pr(X_1 = 0)$, $Pr(X_1 = 1)$, $Pr(X_2 = 0)$, $Pr(X_2 = 1)$.
- There are 4 in total. How many "free" marginal probabilities are there?
- What are the sums-to-one constraints?

Defintion of Conditional Probability

$$Pr(X_1|X_2) = \frac{Pr(X_1 \cap X_2)}{Pr(X_2)}$$

also

$$Pr(X_1 \cap X_2) = Pr(X_1|X_2)Pr(X_2)$$

Your Turn

- $Pr(X_1 = 1)$
- $Pr(X_1 = 1, X_2 = 0)$
- $Pr(X_1 = 1|X_2 = 0)$
- How many free parameters in $Pr(X_1|X_2)$?

Law of Total Probability (conditional format)

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$$

We will use this form a lot soon

Law of Conditional Probability

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

Bernoulli

Frequentist:

$$Y(\theta) \sim \text{Bernoulli}(\theta)$$

Bayesian:

$$Y|\theta \sim \text{Bernoulli}(\theta)$$

Short Cut:

$$Y \sim \text{Bernoulli}(\theta)$$

- The Bernoulli is a discrete distribution, it lives on two points: $y = 0$ and $y = 1$.
- The Bernoulli describes dichotomous events, like coin flips.
- The Bernoulli has a single parameter, the probability of success.

Probabilities for Bernoulli

Frequentist:

$$Pr(Y = y, \theta) = \theta^y (1 - \theta)^{(1-y)}, \quad y = 0, 1$$

Bayesian:

$$Pr(Y = y \mid \theta) = \theta^y (1 - \theta)^{(1-y)}, \quad y = 0, 1$$

Normal Distribution

Everyone knows the normal.

Freq.

$$Y(\mu, \sigma^2) \sim \text{Normal}(\mu, \sigma^2)$$

Bayes.

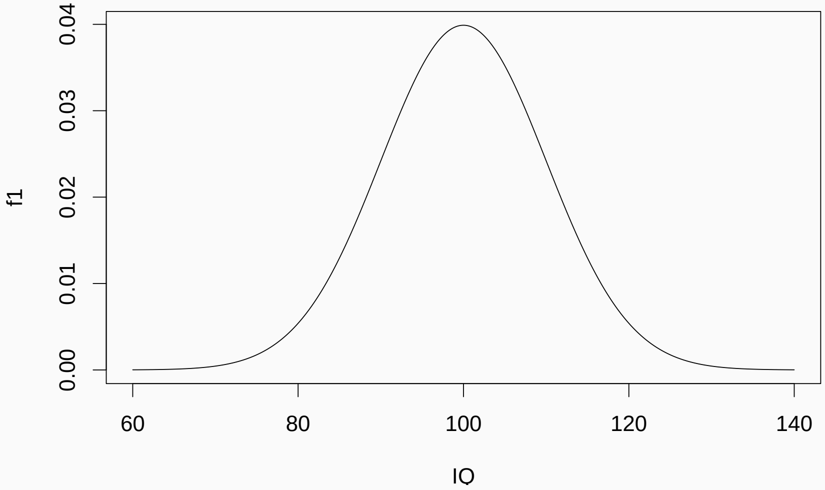
$$Y|\mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$$

Short Cut.

$$Y \sim \text{Normal}(\mu, \sigma^2)$$

$$f(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

Normal Distribution



Normal Distribution

1. The normal is a continuous distribution, it lives on the interval $(-\infty, \infty)$
2. The normal has a location parameter μ and a variance parameter σ^2
3. The normal is used because of its convenience.
4. Effects shift the distribution.

Working the Problem

Consider the dreaded disease **Dipsidoodleitis**. Dipsidoodleitis is detected by a blood test which measures dipsidoodlamines in the blood. Values for people without the disease are normally distributed with $M = 70$ and $sd = 10$. Value for people with the disease are $M = 84$, $sd = 14$. Baserates are know as follows: 10% of people have the disease.

You happen to have a dipsidoodleamine value of 88. What is the probability that you are diseas free?

Working The Problem

Notation Step: Let Let Y denote be a random variable that denotes the dipsidoodlamine level. Let D be a random variable that denotes whether an individual has the disease or not. $(D = 0, 1$ for no disease and disease, respectively).

Model Step What kind of model statements can we make? Are they conditional or marginal statements?

- $D \sim \text{Bernoulli}(.1)$
- $Y|D = 0 \sim \text{Normal}(70, 10^2)$
- $Y|D = 1 \sim \text{Normal}(84, 14^2)$