THE UNIVERSITY OF TEXAS AT AUSTIN
McCOMBS SCHOOL OF BUSINESS

Introduction to predictive analytics

Lecture 1

STA 371G

Probability Theory

The Concept of Probability

What is common among the following?

- Outcome of rolling a die
- S&P500 index at the and of January
- Number of iPhone 7s to be sold over the next year
- Number of unique visitors to Amazon.com over the next week
- Lifetime of your MacBook Air

We cannot predict any of these with certainty.

Probability Theory

The Concept of Probability

Many processes in life involve randomness and the outcome is uncertain. These processes are either

- inherently random (e.g. ones that involve human behavior), or
- the underlying dynamics are so complex to take into account so we treat them as random (e.g. tossing a coin), or
- both (e.g. stock market).

Although we cannot predict with certainty, in many cases, we could assess the likelihoods of the possible outcomes of a process. This is what the probability theory is about.

Probability Theory

Definitions

Definition

An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a random experiment.

Examples

- Rolling a six-sided fair die
- Selling iPhone 7s over a year
- Buying and using a MacBook Air until it breaks down

Probability Theory

Definitions

Definition

A random variable expresses the outcome of a random experiment as a number. It is denoted by an uppercase letter.

Examples (cont'd)

- X : Number of pips on the upper side of the die
- Y : Number of iPhone 7s to be sold over a year
- Z : Lifetime of your MacBook Air

When a random variable is realized (i.e. the result is observed), its value is denoted by a lowercase letter. E.g. $x = 6$, $y = 52316673$ etc.

Probability Theory

Definitions

Multiple random variables can be defined for the same random experiment!

Examples (cont'd)

$$X_2 : \begin{cases} 1, & \text{if there are odd number of pips on the upper side,} \\ 2, & \text{if there are even number of pips on the upper side.} \end{cases}$$

$$Y_2 : \begin{cases} 1, & \text{if iPhone 7 sales exceed 100M over the next year,} \\ 0, & \text{otherwise.} \end{cases}$$

Probability Theory

Definitions

Notice that some random variables can take only discrete values whereas others can take continuous values!

Definition

A discrete random variable is a random variable with a finite (or countably infinite) range.
A continuous random variable is a random variable with an interval (either finite or infinite) of real numbers for its range.

Examples (cont'd)

- (iPhone sales) $Y \in \{0, 1, 2, \ldots\}$
- (MacBook Lifetime) $Z \in [0, \infty)$

Probability Theory

Definitions

Definition

Probability is the measure of the likelihood that a particular outcome (or set of outcomes) will be observed.

Probability is a number that is always between 0 and 1, where 0 implies impossibility and 1 implies certainty.

Examples (cont'd)

- (Rolling a die) $P(X = 5) = \frac{1}{6}$
- (MacBook Lifetime) $P(Z > 15 \text{ years}) = 0.05$

Probability Distributions

So, we have defined our random variable. How do we know what the probabilities are?

For example, what is the probability that your MacBook will break down after 5 years but before 7 years? That is, $P(5 < Y < 7) = ?$

Definition

The probability distribution of a random variable Y is a description of the probabilities associated with the possible values of Y.

Discrete random variable $\longrightarrow$ Probability Mass Function (p.m.f.)
Continuous random variable $\longrightarrow$ Probability Density Function (p.d.f.)

Probability Distributions

Discrete Random Variables

Example

X: The outcome when you roll n-sided fair die.

Since this is a fair die, the probability distribution is given by the following probability mass function:

$$f(x) = \begin{cases} \frac{1}{n} & x = 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

- $f(2) = P(X = 2)$, which is the probability of observing a "2." This interpretation will hold for continuous random variables.
- If you add up all the probabilities, you should get 1. ($n \times \frac{1}{n}$).
- This is an example of Discrete Uniform Distribution.

Probability Distributions

Continous Random Variables

Example

Y : Lifetime of your MacBook (in years)

Let's assume Y has a Continuous Uniform Distribution with a maximum of 20 years. Its probability distribution is then given by the following probability density function:

$$f(y) = \begin{cases} \frac{1}{20} & 0 \leq y \leq 20, \\ 0 & \text{otherwise.} \end{cases}$$

What is $P(Y = 5) =?$ or $P(Y = 5.5) =?$ or $P(Y = 5.551234123) =?$
They are all 0.

Probability Distributions

Continous Random Variables

Warning!

For a continuous random variable, $P(Y = a)$ is always zero, regardless of a. Because Y can take infinite number of values and the chance of particularly hitting one of those points (a) is zero! (although it will eventually take a value. Sounds like a paradox, right?)

For this reason, for continuous random variables we ask questions like "$P(a \leq Y \leq b) = ?$"

And we take integrals to find such probabilities.

Probability Distributions

Continous Random Variables

Example

Y : Lifetime of your MacBook (in years)

$$f(y) = \begin{cases} \frac{1}{20} & 0 \leq y \leq 20, \\ 0 & \text{otherwise.} \end{cases}$$

What is $P(5 < Y < 7) =?$

$$P(5 < Y < 7) = \int_5^7 \frac{1}{20} dy = \frac{y}{20}\Big|_5^7 = \frac{7}{20} - \frac{5}{20} = \frac{1}{10}$$

In general, $P(a \leq Y \leq b) = \int_a^b f(y) dy$.

Probability Distributions

Graphs Go Here

Mean, Variance and Standard Deviation

Definition

Mean or Expected Value of a random variable X is a measure of the center of its probability distribution. It is a weighted average of all possible values X can take, where the weights are the corresponding probabilities.

Discrete random variable X

$$\mu_X = E[X] = \sum_x x f(x)$$

Continuous random variable Y

$$\mu_Y = E[Y] = \int_y y f(y) dy$$

Mean, Variance and Standard Deviation

Definition

Variance of a random variable X is a measure of the dispersion, or variability in its distribution. Standard Deviation of X is the square root of its variance.

Discrete random variable X

$$\sigma_X^2 = Var(X) = E[\,(X - \mu_X)^2] = \sum_x (x - \mu_x)^2 f(x)$$

Continuous random variable Y

$$\sigma_Y^2 = Var(Y) = E[\,(Y - \mu_Y)^2] = \int_y (y - \mu_y)^2 f(y) dy$$