Probability Distributions: Continuous

Introduction to Data Science Algorithms
Jordan Boyd-Graber and Michael Paul
SEPTEMBER 29, 2016
Administrivia

- HW1 Grading Underway
- Shockingly similar submissions
  - As an exercise, at some point we’ll compute the probability
- All code must come from your fingers
- HW2 released
- Python review
Continuous random variables

Today we will look at continuous random variables:

- Real numbers: \( \mathbb{R}; (-\infty, \infty) \)
- Positive real numbers: \( \mathbb{R}^+; (0, \infty) \)
- Real numbers between -1 and 1 (inclusive): \([-1, 1]\)

The sample space of continuous random variables is uncountably infinite.
Continuous distributions

• Last time: a discrete distribution assigns a probability to every possible outcome in the sample space
• How do we define a continuous distribution?
• Suppose our sample space is all real numbers, \( \mathbb{R} \).
  ○ What is the probability of \( P(X = 20.1626338) \)?
  ○ What is the probability of \( P(X = -1.5) \)?
Continuous distributions

- Last time: a discrete distribution assigns a probability to every possible outcome in the sample space.
- How do we define a continuous distribution?
- Suppose our sample space is all real numbers, \( \mathbb{R} \).
  - What is the probability of \( P(X = 20.1626338) \)?
  - What is the probability of \( P(X = -1.5) \)?
- The probability of any continuous event is always 0.
  - Huh?
  - There are infinitely many possible values a continuous variable could take. There is zero chance of picking any one exact value.
  - We need a slightly different definition of probability for continuous variables.
A *probability density function* (PDF, or simply *density*) is the continuous version of probability mass functions for discrete distributions.

The density at a point \( x \) is denoted \( f(x) \).

Density behaves like probability:

- \( f(x) \geq 0 \), for all \( x \)
- \( \int_{x} f(x) = 1 \)

Even though \( P(X = 1.5) = 0 \), density allows us to ask other questions:

- Intervals: \( P(1.4999 < X < 1.5001) \)
- Relative likelihood: is 1.5 more likely than 0.8?
Probability of intervals

- While the probability for a specific value is 0 under a continuous distribution, we can still measure the probability that a value falls within an interval.
  - $P(X \geq a) = \int_{x=a}^{\infty} f(x)$
  - $P(X \leq a) = \int_{x=-\infty}^{a} f(x)$
  - $P(a \leq X \leq b) = \int_{x=a}^{b} f(x)$

- This is analogous to the disjunction rule for discrete distributions.
  - For example if $X$ is a die roll, then $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$
  - An integral is similar to a sum
Likelihood

- The *likelihood function* refers to the PMF (discrete) or PDF (continuous).
- For discrete distributions, the likelihood of \( x \) is \( P(X = x) \).
- For continuous distributions, the likelihood of \( x \) is the density \( f(x) \).
- We will often refer to likelihood rather than probability/mass/density so that the term applies to either scenario.