Discrete Mathematics and Probability

Lecture 15
Continuous Random Variables

Ian Stark

School of Informatics
The University of Edinburgh

Thursday 9 November 2023

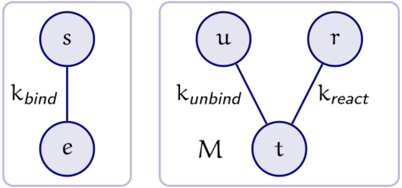
THE UNIVERSITY of EDINBURGH
informatics

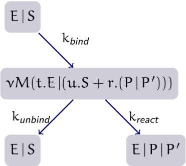
Ian Stark

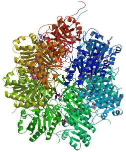
Senior Lecturer in Computer Science
Laboratory for Foundations of Computer Science
School of Informatics

Research: Mathematical models of programming languages and concurrent systems

Continuous pi-Calculus







RCSB Protein Data Bank
2GBL KaiC Circadian Clock

Morello / CHERI

https://blogs.ed.ac.uk/morello

Week 9 Topics

Continuous Random Variables and Probability Distributions

Carlton & Devore: Chapter 3

- §3.1 Probability Density Functions and Cumulative Distribution Functions
- §3.2 Expected Values (but not Moment Generating Functions)
- §3.3 The Normal (Gaussian) Distribution
- §3.4 The Exponential Distribution (but not the Gamma Distribution)

The study guide and accompanying videos indicate the examinable course content. The corresponding sections in the book are to support this. Additional sections in the book are useful extension material but not required for this course.

Two Kinds of Random Variable

Discrete Random Variable

A *discrete* random variable can take only a finite or countably infinite number of possible values.

Often these values are integers. For example, a discrete random variable taking only the values 0 and 1 is a *Bernoulli* random variable

Continuous Random Variable

A *continuous* random variable is one that takes values over a continuous range: the whole real line; an interval on the real line, perhaps infinite; or a disjoint union of such intervals.

In addition, a continuous random variable X must have the property that no possible value has positive probability: $P(X = x) = 0$ for all $x \in \mathbb{R}$.

This only applies to individual values: ranges of values like $P(X > 5)$ or $P(a < X < b)$ may have non-zero probability.

Reminder: Probability Mass Function

Definition

The *probability mass function* (PMF) of a discrete random variable X is defined for every possible value x as follows.

$$p(x) \;=\; P(X = x) \;=\; P(\{\, \omega \in \Omega \mid X(\omega) = x \,\})$$

Any probability mass function $p(x)$ takes only non-negative values, and the sum over all possible values of x will be 1.

Probability Density Function

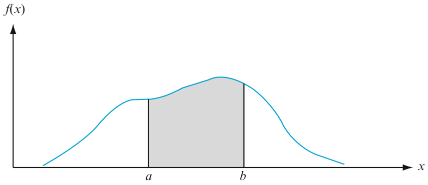
Definition

Let X be a continuous random variable. The *probability density function* (PDF) of X is a function $f(x)$ such that for any two numbers $a \leqslant b$ we have the following.

$$P(a \leqslant X \leqslant b) \;=\; \int_a^b f(x)\,dx\,.$$

For any PDF we know that $f(x) \geqslant 0$ for all values of x and the total area under the whole graph is 1.

$$\int_{-\infty}^{\infty} f(x)\,dx \;=\; 1$$



Carlton & Devore Figure 3.2

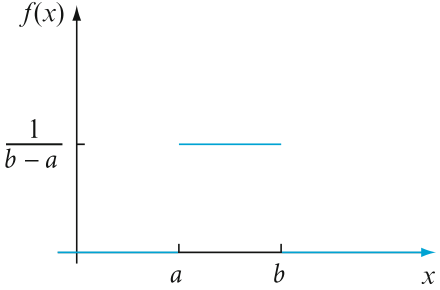
Uniform Distribution

Definition

A continuous random variable X has *uniform distribution* on the interval $[a, b]$ for values $a \leqslant b$ if it has the following PDF.

$$f(x; a, b) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

We write this as $X \sim \text{Unif}[a, b]$.



Carlton & Devore Figure 3.6

Example: Uniform Distribution

A gardening supplier sells packets of seeds. Each packet contains several hundred one-gram seeds. In practice this weight is only approximate, with an error of Y gram in each seed, where random variable Y is uniformly distributed between -0.1 and 0.3. Draw a graph of the probability density function for Y and calculate the following.

1. The probability that an individual seed is less than 0.1g overweight.
2. $P(0.2 \leqslant Y \leqslant 0.4)$

Reminder: Cumulative Distribution Function

Definition

For a discrete random variable X with PMF $p(x)$ its *cumulative distribution function* (CDF) is defined as follows.

$$F(x) = P(X \leqslant x) = \sum_{y \leqslant x} p(y)$$

For any number x, $F(x)$ is the probability that the observed value of X will be no more than x.

Cumulative Distribution Function

Definition

For a continuous random variable X with PDF $f(x)$ its *cumulative distribution function* (CDF) is defined as follows.

$$F(x) \;=\; P(X \leqslant x) \;=\; \int_{-\infty}^{x} f(y)\,dy$$

For any number x, $F(x)$ is the probability that the observed value of X will be no more than x.

Proposition

If X is a continuous random variable with PDF $f(x)$ and CDF $F(x)$ then at every x where the derivative $F'(x)$ is defined we have

$$F'(x) \;=\; f(x)\;.$$

Computing Probabilities with a CDF

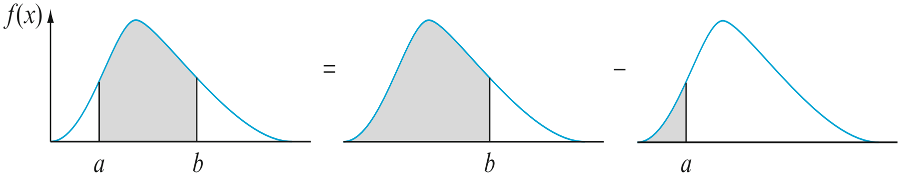
Proposition

Let X be a continuous random variable with PDF $f(x)$ and CDF $F(x)$. Then for any value a we have

$$P(X \leqslant a) = F(a) \qquad P(X > a) \; = \; 1 - F(a)$$

and for any two values $a < b$ we have

$$P(a \leqslant X \leqslant b) \; = \; F(b) - F(a) \; .$$



Carlton & Devore Figure 3.8

Example: PDF and CDF

Random variable T is distributed with the following probability density function.

$$f(t) = \begin{cases} kt(t-1) & 0 \leqslant t \leqslant 1 \\ 0 & t < 0 \text{ or } t > 1 \end{cases}$$

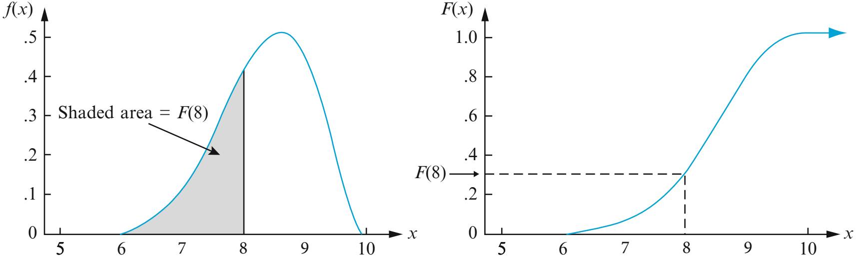
Calculate the value of k, sketch a graph of this PDF, and calculate the cumulative distribution function $F(t)$ for the random variable T. Use this to calculate $P(1/3 < T < 2/3)$.

Review: PDF and CDF

Give a continuous random variable X with probability density function (PDF) $f(x)$, its cumulative distribution function (CDF) is written $F(x)$ and defined as follows.

$$F(x) \;=\; P(X \leqslant x) \;=\; \int_{-\infty}^{x} f(y)\,dy$$

Conversion between PDF and CDF gives different ways to calculate the probabilities involved.



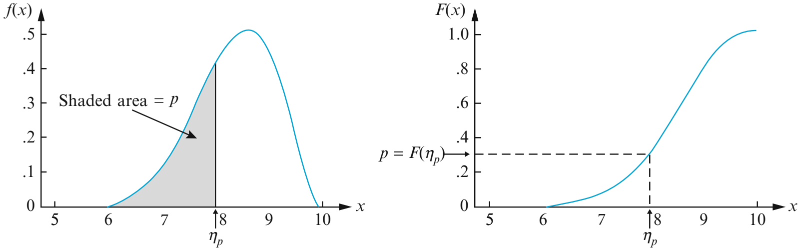
Carlton & Devore Figure 3.5

Percentiles of a Continuous Distribution

Definition

Let X be a continuous random variable with PDF $f(x)$ and CDF $F(x)$ and p any real value between 0 and 1. The $(100p)$*th percentile* of X is the value η_p such that $P(X \leqslant \eta_p) = p$.

So we have $p \;=\; \int_{-\infty}^{\eta_p} f(x)\,dx \;=\; F(\eta_p)$ and $\eta_p \;=\; F^{-1}(p)$.



Carlton & Devore Figure 3.10

Percentiles of a Continuous Distribution

Definition

Let X be a continuous random variable with PDF $f(x)$ and CDF $F(x)$ and p any real value between 0 and 1. The $(100p)$*th percentile* of X is the value η_p such that $P(X \leqslant \eta_p) = p$.

$$\text{So we have } p = \int_{-\infty}^{\eta_p} f(x)\,dx = F(\eta_p) \qquad \text{and } \eta_p = F^{-1}(p)\,.$$

For example, the 35th percentile of a distribution is $\eta_{0.35} = F^{-1}(0.35)$ and the 60th percentile is $\eta_{0.6} = F^{-1}(0.6)$.

The *median* of a distribution is the 50th percentile, $\eta_{0.5} = F^{-1}(0.5)$, sometimes written simply η.

This is the value such that $P(X < \eta) = P(X > \eta) = 1/2$.

Expected Value of a Continuous Random Variable

For a discrete random variable the mean or expected value is an average over all possible values of the variable, weighted by their probabilities. For a continuous random variable we replace this with integration to get a continuous weighted average by probability density.

Definition

Let X be a continuous random variable with PDF $f(x)$. The *expected value* $E(X)$ is calculated as a weighted integral.

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx$$

This is also known as the *mean* of the distribution and written μ_X or simply μ.

Expected Value of a Function of a Continuous Random Variable

Proposition

Let X be a continuous random variable with PDF $f(x)$. If $h(X)$ is any real-valued function of X then we can calculate an expected value for that, too.

$$\mu_{h(X)} = \mathsf{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) \cdot f(x)\, dx$$

The textbook says "This is sometimes called the *Law of the Unconscious Statistician*"; although I can't find anyone actually doing this except in other textbooks.

This is a proposition, not a definition, because $h(X)$ is itself a random variable for which we do not (yet) know the PDF: so calculating $\mathsf{E}(h(X))$ directly is not straightforward.

Note that $\mathsf{E}(h(X))$ and $h(\mathsf{E}(X))$ will not necessarily have the same value.

Example: Expectation of a Function of a Random Variable

Random variable X is distributed with the following PDF.

$$f(x) = \begin{cases} x & \frac{1}{2} \leqslant x \leqslant \frac{3}{2} \\ 0 & \text{otherwise} \end{cases}$$

Sketch the graph of this PDF. Calculate the expected values of X, $1/X$, and X^2. Compare $\mathrm{E}(1/X)$ with $1/\mathrm{E}(X)$ and $\mathrm{E}(X^2)$ with $\mathrm{E}(X)^2$.

Variance and Standard Deviation for a Continuous Random Variable

Definition

Let X be a continuous random variable with PDF $f(x)$ and mean μ. Its *variance* $\mathrm{Var}(X)$ is the expected value of the squared distance to the mean.

$$\mathrm{Var}(X) \;=\; E\left((X-\mu)^2\right) \;=\; \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)\,dx$$

The *standard deviation*, written $\mathrm{SD}(X)$ or σ_X or just σ, is the square root of the variance.

$$\sigma_X \;=\; \mathrm{SD}(X) \;=\; \sqrt{\mathrm{Var}(X)}$$

Properties 1/2

Let X be a continuous random variable with PDF $f(x)$, mean μ, and standard deviation σ.

Variance Shortcut

$$\mathsf{Var}(X) \;=\; \mathsf{E}(X^2) - \mu^2 \;=\; \int_{-\infty}^{+\infty} x^2 \cdot f(x)\,dx - \left(\int_{-\infty}^{+\infty} x \cdot f(x)\,dx\right)^2$$

Chebyshev Inequality

For any constant value $k \geqslant 1$, the probability that X is more than k standard deviations away from the mean is no more than $1/k^2$.

$$P\big(|X - \mu| \geqslant k\sigma\big) \;\leqslant\; \frac{1}{k^2}$$

Properties 2/2

Let X be a continuous random variable with PDF $f(x)$, mean μ, and standard deviation σ.

Linearity of Expectations

For any functions $h_1(X)$ and $h_2(X)$ and constants a_1, a_2, and b, the expected value of these in linear combinations is the linear combination of the expected values.

$$\mathsf{E}\big(a_1 \cdot h_1(X) + a_2 \cdot h_2(X) + b\big) \;=\; a_1 \cdot \mathsf{E}\big(h_1(X)\big) + a_2 \cdot \mathsf{E}\big(h_2(X)\big) + b$$

Rescaling

For any constants a and b the mean, variance and standard deviation of $(aX + b)$ can be calculated from the corresponding values for X.

$$\mathsf{E}(aX + b) \;=\; a\,\mathsf{E}(X) + b \qquad \mathsf{Var}(aX + b) \;=\; a^2\,\mathsf{Var}(x) \qquad \mathsf{SD}(aX + b) \;=\; |a|\,\mathsf{SD}(X)$$

Example: Proving Properties of Expected Values

Prove that expected value scales linearly: $\mathsf{E}(aX + b) = a\,\mathsf{E}(x) + b$

(Use the "Law of the Unconscious Statistician".)

Prove the shortcut for calculating variance: $\mathsf{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}(X)^2$.

(Use the linearity of expectations for expected value of functions of a random variable.)

Summary

Topics

- Continuous random variables
- Uniform distribution
- PDF, CDF, and converting between them
- Percentiles; median; expected values; variance and standard deviation
- Variance shortcut; Chebyshev inequality; linearity of expectations; rescaling

Study Guide

- Continuous Random Variables
- Numerical Properties of Continuous Distributions

Reading

Chapter 3, §§3.1.1–3.1.5 and §3.2.1; Pages 147–157 and 162–166.

Exercises

Chapter 3, Exercises 1–31; Pages 158–162 and 168–170.