Lecture 15: Continuous Random Variables

Terminology, Use, and Common Distributions

Course page: imouzon.github.io/stat305

What is a Continuous Random Variable?

Background on Continuous Random Variable

Along with discrete random variables, we have continuous random variables. While discrete random variables take one specific values from a *discrete* (aka countable) set of possible real-number values, continous random variables take values over intervals of real numbers.

> **def: Continuous random variable**
> A continuous random variable is a random variable which takes values on a continuous interval of real numbers.

The reason we treat them differently has mainly to do with the differences in how the math behaves: now that we are dealing with interval ranges, we change summations to integrals.

Terminology and Usage

Probability Density Function

Since we are now taking values over an interval, we can not "add up" probabilities with our probability function anymore. Instead, we need a new function to describe probability:

> **def: probability density function**
> A probability density function (pdf) defines the way the probability of a continuous random variable is distributed across the interval of values it can take. Since it represents probability, the probability function must always be non-negative. Regions of higher density have higher probability.

Any function that satisfies the following can be a probability density function:

1. $\int_{-\infty}^{\infty} f(x)dx = 1$
2. $f(x) \geq 0$ for all x in $(-\infty, \infty)$

Background

Terms and Use

pdf

Probability Density Function

With continuous random variables, we use pdfs to get probabilities as follows:

> For a continuous random variable X with probability density function $f(x)$,
>
> $$P(a \leq X \leq b) = \int_a^b f(x)dx$$
>
> for any real values a, b such that $a \leq b$

Drawing: Figure 5.6 in book

Cumulative Density Function

We also have the cumulative density function for continuous random variables:

> **def: Cumulative density function (cdf)** For a continous random variable, X, with pdf f(x) the cumulative density function $F(x)$ is defined as the probability that X takes a value less than or equal to x which is to say
>
> $$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

TRUE FACT: the Fundamental Theorem of Calculus applies here:

$$\frac{d}{dx}F(x) = f(x)$$

Probability Density Function: Example

Suppose that X is a continuous random variable which can take any value on the interval $(0, 10)$ and that it's density is spread uniformly across that interval. In other words,

$$f(x) = \begin{cases} c & 0 < x < 10 \\ 0 & o.w. \end{cases}$$

1. Find the value of c that makes $f(x)$ a valid probability function.

2. Sketch the probability density function.

3. Sketch the cumulatiive density function.

4. Find $P(2 \leq X \leq 4)$.

Expected Value and Variance

As with discrete random variables, continuous random variables have expected values and variances:

> **def: Expected Value of Continuous Random Variable**
> For a continous random variable, X, with pdf f(x) the expected value (also known as the mean) is defined as
>
> $$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

We often use the symbol μ for the mean of a random variable, since writing $E(X)$ can get confusing when lots of other parenthesis are around. We also sometimes write EX.

Background

Terms and Use

pdf

cdf

E(X), V(X)

Expected Value and Variance

> **def: Variance of Continuous Random Variable**
> For a continous random variable, X, with pdf f(x) and expected value μ, the variance is defined as
>
> $$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x) dx$$
>
> which is identical to saying
>
> $$V(X) = E(X^2) - E(X)^2$$

We will sometimes use the symbol σ^2 to refer to the variance and you may see the notation $VarX$ or VX as well.

Expected Value and Variance

We can also use the variance to get the standard deviation of the random variable:

> **def: Standard Deviation of Continuous Random Variable**
> For a continous random variable, X, with pdf f(x) and expected value μ, the standard deviation is defined as:
>
> $$\sigma = \sqrt{\sigma^2} = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx}$$

Probability Density Function: Example

Suppose that X is a continuous random variable which can take any value on the interval $(0, 10)$ and that it's density is spread uniformly across that interval. In other words,

$$f(x) = \begin{cases} 0.10 & 0 < x < 10 \\ 0 & o.w. \end{cases}$$

1. Find the standard deviation of X.

Common Distributions

Background

Terms and Use

Common Dists

Uniform

Common Distributions

Uniform Distribution

For cases where we only know/believe/assume that a value will be between two numbers but know/believe/assume *nothing* else.

Origin: We know a the random variable will take a value inside a certain range, but we don't have any belief that one part of that range is more likely than another part of that range.

> **Definition: Uniform random variable**
> The random variable U is a uniform random variable on the interval $[a, b]$ if it's density is constant on $[a, b]$ and the probability it takes a value outside $[a, b]$ is 0. We say that U follows a uniform distribution or $U \sim uniform(a, b)$.

Uniform Distribution

Definition: Uniform pdf
If U is a uniform random variable on $[a, b]$ then the probability density function of U is given by

$$f(u) = \begin{cases} \frac{1}{b-a} & a \leq u \leq b \\ 0 & o.w. \end{cases}$$

With this, we can find the for any value of a and b, if $U \sim uniform(a, b)$ the mean and variance are:

$$E(U) = \frac{1}{2}(b-a)$$

$$Var(U) = \frac{1}{12}(b-a)^2$$

Background

Terms and Use

Common Dists

Uniform

Uniform Distribution

Definition: Uniform cdf
If U is a uniform random variable on $[a, b]$ then the cumulative density function of U is given by

$$F(u) = \begin{cases} 0 & u < a \\ \dfrac{u-a}{b-a} & a \leq u \leq b \\ 1 & u > b \end{cases}$$

Uniform Distribution

A few useful notes:

- The most commonly used uniform random variable is $U \sim Uniform(0, 1)$.
- Again, this is useful if we want to use a random variable that takes values within an interval, but we don't think it is likely to be in any certain region.
- The values a and b used to determine the range in which $f(u)$ is not 0 are parameters of the distribution.

Exponential Distribution

Origin: Often used in to take the values of waiting times until some specific occurance where the idea of "memorylessness" is important - meaning that if you expect to see the occurance in 1 minute after waiting 5 minutes you still expect to see the occurance in 1 minute.

> **Definition: Exponential random variable**
> The random variable X is a exponential random variable on the interval $[0, \infty]$ if it's density drops exponentially with rate $\frac{1}{\alpha}$ as you move away from 0. We say that X follows a exponential distribution or $X \sim exponential(\alpha)$.

Exponential Distribution

> **Definition: Exponential pdf**
> If X is an exponential random variable with rate $\frac{1}{\alpha}$ then the probability density function of X is given by
>
> $$f(u) = \begin{cases} \frac{1}{\alpha} e^{-\frac{x}{\alpha}} & x \geq 0 \\ 0 & o.w. \end{cases}$$

From this, we can derive:

$$E(X) = \alpha$$

$$Var(X) = \alpha^2$$

Uniform Distribution

Definition: Exponential cdf
If X is a exponential random variable with rate $1/\alpha$ then the cumulative density function of X is given by

$$F(x) = \begin{cases} 1 - exp(-x/\alpha) & 0 \leq x \\ 0 & x < 0 \end{cases}$$

We sometimes write the exponential part as $exp(-x/\alpha)$ for clarity.

Exponential Distribution

The memoryless property of exponential random variables

Exponential random variables have what is called a memory less property. That is, that if we know the random variable is greater than s then the probability that it is greater than $s + t$ can be found by the following:

$$P(X > s + t | X > s) = P(X > t)$$

In other words, if we knowing that a process has not occured *yet* does not change the probability that it will occur *soon*. The waiting time is forgotten. This can be both good and bad, depending on what the event you are waiting for is.

Example:

an electrical surge passes through a certain resistor randomly following an exponential distribution with a rate of once every hour. It has been two hours since the last surge. How long should we wait expect to wait for the next surge?

Background

Terms and Use

Common Dists

Uniform

Exponential

Exponential Distribution

Example

Suppose that $X \sim exponential(.5)$. Find the probability that $X > \mu$.