Wednesday, October 18 **Fall 2023**

Section 5: Continuous Distributions

Srihari Ganesh

Based on section note formatting template by Rachel Li and Ginnie Ma '23

1 Summary

1.1 Continuous Random Variables

Definition 1 (Continuous r.v.)**.** A **continuous** random variable has an interval for its support.

More precisely: a continuous random variable has uncountable support, while discrete random variable shave finite/countably infinite support.

Definition 2 (Cumulative distribution function)**.** The **cumulative distribution function (CDF)** of a random variable X is the function $F : \mathbb{R} : [0, 1]$ defined by $F(x) = P(X \leq x)$.

☣ **3.** For a continuous random variable X, *any* value of $x \in \mathbb{R}$ (including those in the support) has $P(X = x) = 0$. This also means that the CDF can be defined in multiple ways since
$P(X \leq x) = P(X < x) + P(X = x) = P(X < x)$.

Definition 4 (Probability density function)**.** For a continuous random variable X with CDF F, the **probability density function (PDF)** is $f : \mathbb{R} \to \mathbb{R}^{\geq 0}$ defined as

$$f(x) = \frac{d}{dx}F(x). \tag{1}$$

Probability density functions satisfy the following condition:

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

☣ **5.** Probability densities don't play nearly as nicely as probabilities. One common mistake is the following: if you know the PDF of a random variable X, the PDF *does not* over to $g(X)$, i.e.,

$$f_X(x) \neq f_{g(X)}(g(x)).$$

1.1.1 Uses of CDFs and PDFs

For *any* random variable X (continuous or discrete), you can use the CDF to calculate the following:

$$P(X > x) = 1 - P(X \leq x) = 1 - F(x)$$

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1)$$

You can assume that the CDF of a continuous random variable is differentiable, so that the PDF can actually be defined. We can find the probabilities of intervals by integrating the PDF and adjusting the bounds:

$$P(X \leq x) = P(X < x) = \int_{-\infty}^{x} f(x)dx$$
$$P(X \geq x) = P(X > x) = \int_{x}^{\infty} f(x)dx$$
$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx.$$

1.1.2 Continuous analogs of all of our tools

The general rules are:

- Integrals instead of sums.
- PDFs instead of PMFs.
- When asked to find the distribution of a continuous random variable, it's much easier to work with the CDF than the PDF.

So here's a table with the tools we've talked about:

Tool	Discrete	Continuous
Expectation	$E(X) = \sum_x xP(X = x)$	$E(X) = \int_{-\infty}^{\infty} xf_X(x)dx$
LOTUS	$E(g(X)) = \sum_x g(x)P(X = x)$	$E(g(X)) = \int_{-\infty}^{\infty} g(x)f_X(x)dx$
Bayes' rule	$P(X = x \mid Y = y) = \frac{P(Y=y \mid X=x)P(X=x)}{P(A)}$	$f_{X\mid Y=y}(x) = \frac{f_{Y\mid X=x}(y)f_X(x)}{f_Y(y)}$

1.2 Uniform

Definition 6 (Uniform distribution)**.** For any interval (a, b), a random variable $U \sim \text{Unif}(a, b)$ has a uniform distribution (i.e., constant PDF) over the support (a, b). There is no uniform whose support is the full real line.

The PDF and CDF can be derived:

$$f_U(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & x \notin (a, b) \end{cases}$$
$$F_U(x) = P(U < x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & x \in (a, b) \\ 1 & x \geq b \end{cases}$$

2 Practice Problems

1. Suppose X is a random variable with the following PDF:

$$f_X(x) = \begin{cases} \frac{1}{x^2} & x \geq 1, \\ 0 & x < 1. \end{cases}$$

   (a) What is the expected value of X?

Solution

$$E(X) = \int_1^\infty x\frac{1}{x^2}dx = \int_1^\infty \frac{1}{x} = \ln x|_1^\infty = \infty.$$

   (b) What is the expected value of $1/X$?

Solution
Using LOTUS,

$$E(1/X) = \int_1^\infty \frac{1}{x}\frac{1}{x^2}dx = -\frac{1}{x^2}|_1^\infty = 1.$$

   (c) What is the distribution of X^2?

Solution
The support of X^2 is $[1, \infty)$. We can find the CDF, assuming $y \geq 1$:

$$P(X^2 \leq y) = P(X \leq \sqrt{y}) = \int_1^{\sqrt{y}} \frac{1}{x^2}dx = -\frac{1}{x}|_1^{\sqrt{y}} = 1 - \frac{1}{\sqrt{y}}$$

Given that $P(X^2 \leq y) = 0$ if $y < 1$, we have defined our CDF. That is sufficient for defining the distribution.
If you're curious about the PDF:

$$f_{X^2}(y) = \begin{cases} \frac{1}{2y^{3/2}} & y \geq 1 \\ 0 & y < 1 \end{cases}$$

   (d) Now suppose X is the side length of a square. What is the distribution of the area of the square?

Solution
The square has area X^2, so the distribution of the area is the same as the distribution of X^2, as found in part (a)

2. Suppose Y is a random variable with the following PDF:

$$f_Y(y) = \frac{1}{2}e^{-|y|}.$$

(a) What is the distribution of $|Y|$?

Solution
We can derive the CDF: for $y \geq 0$ we get

$$\begin{aligned} P(|Y| < y) = P(-y < Y < y) &= \int_{-y}^{y} \frac{1}{2}e^{-|x|}dx \\ &= 2\int_0^y \frac{1}{2}e^{-|x|}dx = \int_0^y e^{-x}dx \\ &= -e^{-x}|_0^y = 1 - e^{-y}. \end{aligned}$$

So the full CDF is

$$P(|Y| < y) = \begin{cases} 1 - e^{-y} & y \geq 0, \\ 0 & y < 0. \end{cases}$$

As we'll see soon, $|Y| \sim \text{Expo}(1)$.

(b) What is the expected value of Y, given that the expectation does exist (i.e., is not infinite)? *(hint: use symmetry)*

Solution

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} \frac{1}{2}ye^{-|y|}dy \\ &= \int_0^{\infty} \frac{1}{2}ye^{-y}dy + \int_{-\infty}^{0} \frac{1}{2}ye^{y}dy \\ &= \frac{1}{2}\int_0^{\infty} ye^{-y}dy + \frac{1}{2}\int_0^{\infty} -ye^{-y}dy \end{aligned}$$

The two terms are identical up to a negative, so they actually cancel out to give $E(Y) = 0$.

(c) What is the expected value of $e^{-|Y|}$? Use LOTUS on the distribution of $|Y|$ that you found in part (a).

Solution
If we let $Z = |Y|$, then we find that the PDF is $f(z) = e^{-z}$ for $z \geq 0$ and $f(z) = 0$ for $z < 0$. We can then use LOTUS:

$$E(e^{-Z}) = \int_0^{\infty} e^{-z}e^{-z}dz = \int_0^{\infty} e^{-2z}dz$$

We know that $\int_0^{\infty} e^{-z}dz = 1$ since Z has a valid PDF. So let's set $u = 2z$ and usubstitute:

$$E(e^{-Z}) = \frac{1}{2}\int_0^\infty e^{-u}du = \frac{1}{2}.$$

3. Suppose $X \sim \text{Unif}(0,1)$.

(a) Find $P(X = 0.3)$.

Solution
Since X is continuous, $P(X = 0.3) = 0$.

(b) Find $P(X < 0.3)$.

Solution
Using the CDF, $P(X < 0.3) = 0.3$.

(c) Find $P(0.3 < X < 0.7)$.

Solution
We can use the CDF:

$$P(0.3 < X < 0.7) = P(X < 0.7) - P(X < 0.3) = 0.7 - 0.3 = 0.4.$$

We could also integrate the PDF:

$$P(0.3 < X < 0.7) = \int_{0.3}^{0.7} 1dx = 0.7 - 0.3 = 0.4.$$

(d) Find $P(0.3 \leq X \leq 0.7)$.

Solution
This is the same as the previous part since the extra equalities don't matter.