Mathematics 136 – Calculus 2
More on Distribution Functions (pdfs)
March 17, 2020

Background

A function $f(x)$ that is non-negative for all x and that satisfies the condition that *the total area between the graph $y = f(x)$ and the x-axis is one*, i.e.

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

is called a *probability density function* (pdf). A random variable X is said to have pdf $f(x)$ (or have a distribution described by $f(x)$) if the probability that $a \leq X \leq b$ is given by

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\, dx \tag{2}$$

for all a, b.

The *mean, or expected value* of a random variable X with distribution described by the pdf $f(x)$ is computed by

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx, \tag{3}$$

provided that this improper integral exists. This can be thought of as something like the *weighted average* of x, with weights given by the pdf $f(x)$.

Note that the integrals in (1), (3) are *improper integrals*, and the integral in (2) can be improper as well. This means that we will need to apply the techniques we learned for improper integrals to compute them if we need to.

Questions

A) Find the constant c so that the function

$$f(x) = \begin{cases} ce^{-3x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is a pdf. Use your value of c to determine $P(1 < X < 3)$ if X is a random variable with a distribution described by this pdf. (These are called *exponential* pdfs. They are used to model the distributions of random quantities like the lifetime of a randomly selected electronic component of a particular type, or the time between random incoming calls to customer service call centers, etc.)

Solution: We must have

$$1 = c\int_0^\infty e^{-3x}\ dx = c\lim_{b\to\infty}\frac{1}{3}(1-e^{-3b}) = c\cdot\frac{1}{3}.$$

Therefore $c = 3$. Then

$$P(1 < X < 3) = \int_1^3 3e^{-3x}\ dx = -e^{-9} + e^{-3} \doteq .0497$$

(about a 5% chance).

B) Find the expected value of X from question A. What is the probability that $X > E(X)$? What is the value x_m where $P(X > x_m) = \frac{1}{2}$ (the median)? How are the mean and median related in this example?

Solution: The expected value is

$$\begin{aligned}\int_0^\infty 3xe^{-3x}\ dx &= \lim_{b\to\infty}\left(-xe^{-3x} - \frac{1}{3}e^{-3x}\Big|_0^b\right)\\ &= \lim_{b\to\infty}\left(\frac{-b}{e^{3b}} - \frac{1}{3e^{3b}} + \frac{1}{3}\right)\\ &= \frac{1}{3}\end{aligned}$$

(The limit of the first term is zero by L'Hopital's rule).

The probability that $X > E(X)$ is

$$P(X > 3) = \int_3^\infty 3e^{-3x}\ dx = e^{-9} \doteq .0001234$$

We have

$$\frac{1}{2} = \int_{x_m}^\infty 3e^{-3x}\ dx = \int_0^{x_m} 3e^{-3x}\ dx = 1 - e^{-3x_m}$$

when $e^{-3x_m} = \frac{1}{2}$ or

$$x_m = \frac{\ln(.5)}{-3} \doteq .231$$

The median is less than the mean or expected value in this case. (This is expected because of the "long right tail" of the graph of the pdf.)

C) Find the c so that $f(x) = \dfrac{c}{x^2+1}$ (domain all real numbers) is a pdf.

Solution: We must have

$$1 = c \int_{-\infty}^{\infty} \frac{1}{x^2+1}\, dx$$

But

$$\int \frac{1}{x^2+1}\, dx = \tan^{-1}(x) + C$$

and

$$lim_{c\to\pm\infty} \tan^{-1}(x) = \pm\frac{\pi}{2}$$

. Hence $1 = c \cdot \pi$ and $c = \frac{1}{\pi}$.

D) Does the expected value exist for a random variable with pdf as in part C? Why or why not? Does the median exist?

Solution: The expected value would be computed by the integral

$$\int_{-\infty}^{\infty} \frac{1}{\pi} \cdot \frac{x}{x^2+1}\, dx$$

Splitting at 0 (and omitting the constant factor $\frac{1}{\pi}$ for simplicity) we must have that both improper integrals

$$\int_{-\infty}^{0} \frac{x}{x^2+1}\, dx \quad \text{and} \quad \int_{0}^{\infty} \frac{x}{x^2+1}\, dx$$

must exist. However,

$$\int_0^{\infty} \frac{x}{x^2+1}\, dx = \lim_{b\to\infty} \frac{1}{2} \ln(x^2+1)\Big|_0^b = \lim_{b\to\infty} \frac{1}{2} \ln(b^2+1)$$

does not exist. Therefore the expected value does not exist. The median, however, does exist. By the symmetry about the y-axis of the pdf graph

$$y = \frac{1}{\pi} \cdot \frac{1}{x^2+1}$$

we have

$$\int_{-\infty}^{0} \frac{1}{\pi} \cdot \frac{1}{x^2+1}\, dx = \frac{1}{2} = \int_{0}^{\infty} \frac{1}{\pi} \cdot \frac{1}{x^2+1}\, dx$$

Therefore, the median is 0.