San José State University

Math 161A: Applied Probability & Statistics

Continuous distributions

Prof. Guangliang Chen

Section 4.1 Probability density functions

Section 4.2 Cumulative distribution functions and expected values

Recall that in Chapter 3 we studied discrete random variables, which can only take countably many values.

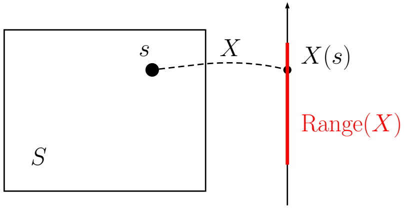
- To characterize their distributions, we introduced pmf and cdf;
- To summarize their distributions, we defined expectation and variance.

We then went through a list of named discrete distributions.

In this part we are going to learn about continuous random variables.

Definition of continuous random variables

Def 0.1. We say that a random variable X is **continuous** if its range is an interval (or a union of intervals).



Example 0.1. Typical examples include measurement of an object, life time of electronics, waiting time.

Distributions of continuous random variables

... can be fully characterized by

- **probability density functions (pdf)**, or
- **cumulative distribution functions (cdf)**

Recall that distributions of discrete random variables are described by

- probability mass function (pmf). or
- cumulative distribution functions (cdf).

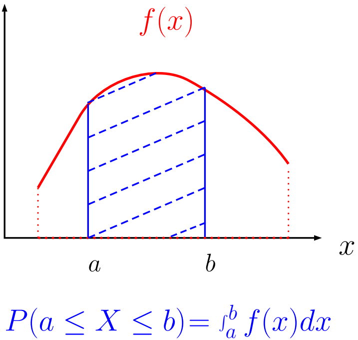
Probability density function (pdf)

Def 0.2. The pdf of a continuous random variable X is a function $f:\mathbb{R}\to\mathbb{R}$ satisfying

- $f(x) \ge 0$ for all $x \in \mathbb{R}$ (and $f(x) > 0$ over an interval, or several intervals)
- $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$

such that for any $a \le b$,

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x.$$



How to read a pdf plot:

- Range(X) is the portion where $f(x) > 0$
- $f(x) = 0$ for any x outside of the range (by default)
- The probability that X takes any particular value $c \in \mathbb{R}$ is always 0:

$$P(X = c) = P(c \leq x \leq c) = \int_c^c f(x)\,\mathrm{d}x = 0.$$

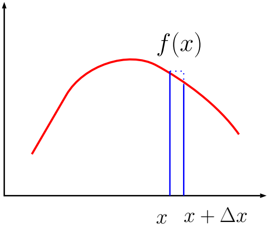
This implies that the endpoints of an interval make no effect on the probability calculations:

$$\mathbf{P(a < X < b)} = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$
$$= \int_a^b f(x)\,\mathrm{d}x.$$

Interpretation of the pdf

For any $x \in \text{Range}(X)$ ($f(x) > 0$), and small increment $\Delta x > 0$,

$$P(x \le X \le x + \Delta x)$$
$$= \int_x^{x+\Delta x} f(y)\,\mathrm{d}y \approx f(x)\Delta x$$



This implies that

- $f(x)\Delta x$ is the probability that X falls into the interval $(x, x + \Delta x)$;
- $f(x)$ alone can be thought of rate.

Example 0.2. The constant function $f(x) = 1, 0 \leq x \leq 1$ is a pdf. Find

- $P(X < -1)$,
- $P(X = 0.2)$,
- $P(X < 0.2)$,
- $P(0.2 < X < 0.5)$,
- $P(X > 0.6)$.

Example 0.3. Find the constant c such that $f(x) = c(1-x), 0 < x < 1$ is a pdf.

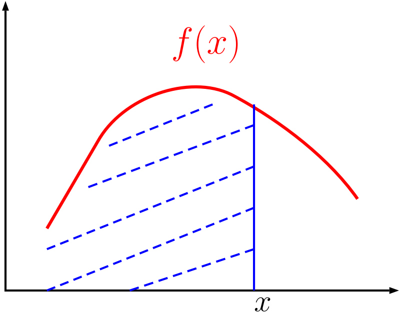
Cumulative distribution function (cdf)

Def 0.3. Let X be a continuous random variable with pdf $f(x)$. The cdf of X is defined as the function

$$F:\mathbb{R}\mapsto\mathbb{R}$$

with

$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(y)\,\mathrm{d}y$$



(Recall the discrete case: $F(x)=P(X\le x)=\sum_{i:x_i\le x} f(x_i)$)

Example 0.4. Find the cdf in each of the last two examples.

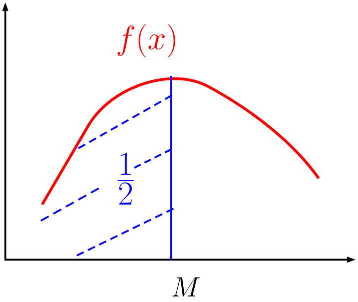
Properties of $F(x)$ (for continuous random variables)

- $F(x)$ always satisfies the following properties (and vice versa):
  - $\lim_{x\to-\infty} F(x) = 0$,
    $\lim_{x\to\infty} F(x) = 1$;
  - $F(x)$ is nondecreasing over $\mathbb{R}$;
  - $F(x)$ is continuous.
- $P(X > a) = 1 - F(a)$ and $P(a < X < b) = F(b) - F(a)$.
- $F'(x) = f(x)$ (due to the Fundamental Theorem of Calculus)

Median of a continuous distribution

Def 0.4. The **median** of the distribution of a continuous random variable X with pdf $f(x)$ is defined as the number M such that

$$\frac{1}{2} = F(M) = \int_{-\infty}^{M} f(x)\,\mathrm{d}x.$$



Remark. It is another way to define the center of the distribution (besides expected value, to be shown on next slide).

Example 0.5. For the pdf $f(x) = 2(1-x), 0 < x < 1$, show that $M = 1 - \sqrt{1/2}$.

Expected value and variance

Def 0.5. The expectation of a continuous random variable X with pdf $f(x)$ is defined as

$$\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

and its variance as

$$\begin{aligned}\sigma^2 = \mathrm{Var}(X) = \mathrm{E}((X-\mu)^2) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x \\ = \mathrm{E}(X^2) - \mu^2 &= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - \mu^2\end{aligned}$$

Example 0.6. In each of the previous examples, find the mean, variance and standard deviation of X.

Expected value of functions of X

Def 0.6. Let X be a continuous random variable with pdf $f(x)$. For any function g,

$$\mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\, \mathrm{d}x.$$

Remark. Recall that for a discrete random variable X:

$$\mathrm{E}(g(X)) = \sum_i g(x_i) f(x_i).$$

Example 0.7. Consider the random variable with pdf $f(x) = 1, 0 \leq x \leq 1$. Find $\mathrm{E}(X^k)$, where $k \geq 1$ is an integer.

Properties of expectation and variance

$\mathrm{E}(\cdot)$ and $\mathrm{Var}(\cdot)$ satisfy exactly the same properties as in the discrete case:

- For any $a, b \in \mathbb{R}$, and a continuous random variable X,

$$\mathrm{E}(a \cdot X + b) = a \cdot \mathrm{E}(X) + b$$
$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

- For any two continuous random variables X, Y,

$$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$
$$\mathrm{Var}(X + Y) \overset{\text{indep.}}{=} \mathrm{Var}(X) + \mathrm{Var}(Y)$$