Continuous Random Variables

Definition: A random variable is continuous if its set of possible values is an entire interval of numbers.

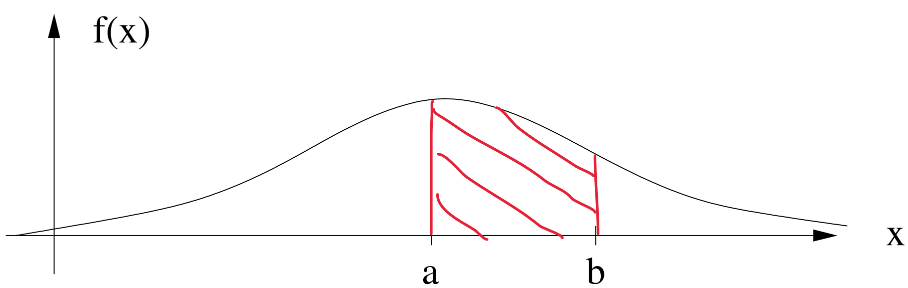
Examples:

- The weight of a random fish caught in the SF Bay.
- Your exact height on your 80^{th} birthday.
-

Definition: Let X be a continuous random variable. Then the probability density function (PDF) of X is a function $f(x)$ such that for any two numbers a and b

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

That means, the probability that X takes on a value between a and b is represented by the area under the curve $f(x)$ between a and b.



Poll Question 7.1

Proposition:

- If X is a continuous random variable, then $P(X = c) = 0$ for every c.

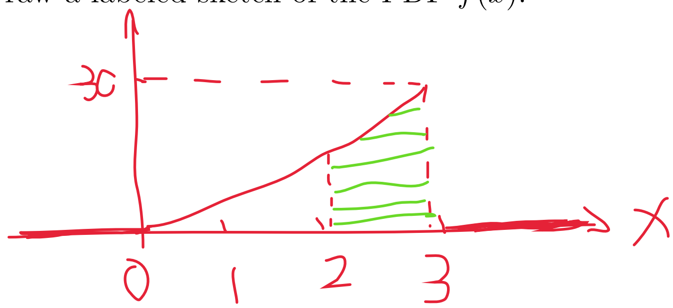
  Note that assigning probability zero to the possible outcome c *does not* rule out c as a possible value. But it implies that the chances of observing *exactly* c are very small.

- The density $f(x)$ is defined on the entire real line, but it may be zero on an interval. $f(c) = 0$ *does* rule out c as a possible value.
- $f(x) \geq 0$ for all $x \in \mathbb{R}$.
- $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Example: Let X be a continuous random variable with PDF

$$f(x) = \begin{cases} c \cdot x & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Draw a labeled sketch of the PDF $f(x)$.



(b) Find the value of the constant c.

PDF(X)=1

4.5c = 1

Poll Question 7.2

(c) Compute $P(2 < X \leq 4)$.

$P(2<X<4)=P(2<X<3)+P(3<X<4)=5/9$

Cumulative Distribution Function

DEFINITION: The cumulative distribution function (CDF) for a continuous random variable is defined in the same way as the CDF for a discrete random variable:

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

In the density curve picture, $F(x)$ is the area under the curve $f(x)$ to the left of x.

Poll Question 7.3

PROPOSITION: You can obtain the PDF $f(x)$ from the CDF $F(x)$

$$f(x) = F'(x) \text{ (Fundamental Theorem of Calculus)}$$

PROPOSITION: The CDF can be used to compute probabilities for a continuous random variable:

$$P(X \leq a) = F(a)$$

$$P(X > a) = 1 - F(a)$$

$$P(a \leq X \leq b) = F(b) - F(a)$$

Example: (cont.)

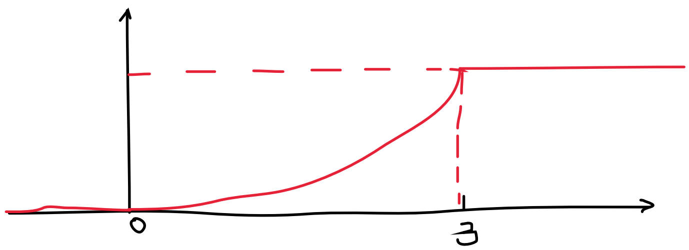
Let X be the same continuous random variable from before:

$$f(x) = \begin{cases} \frac{2}{9}x & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(d) Find the CDF of X.

$$F(X) \begin{cases} 0 & x<0 \\ x\text{^}2/9 & 0<=x<=3 \\ 1 & x>3 \end{cases}$$

(e) Draw a labeled sketch of the CDF of X.



(f) Use the CDF to compute $P(X > 2)$.

P(X>2)=1-P(X<2)=1-F(2)=5/9

Poll Question 7.4

Percentiles

DEFINITION: Let p be a number between 0 and 1. The 100pth percentile of the distribution of a continuous random variable X, denoted by $\eta(p)$, is defined by

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(x)dx$$

That is, $\eta(p)$ is the number on the x-axis, such that the shaded area under $f(x)$ to the left of $\eta(p)$ has size p.

DEFINITION: The 50^{th} percentile is also called the MEDIAN.

Example: (cont.)

(g) Find the 75^{th} percentile for the continuous random variable X above.

p=0.75=

$\eta = \sqrt{9 \times 0.75}$

Expected Value and Variance

The major difference in computation of expected values and variances between discrete and continuous random variables is that where sums were used for discrete random variables, we will now use integrals:

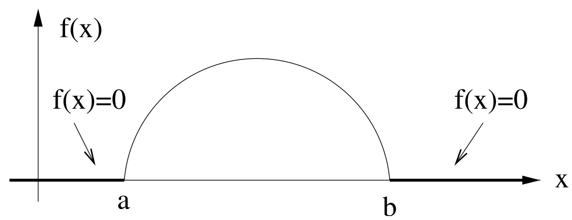
DEFINITION: Let X be a continuous random variable. Then the MEAN or EXPECTED VALUE of X is defined as

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

Remark:

- Compare this to the discrete formula $E(X) = \sum x \cdot p(x)$. The sum is replaced by an integral and the PMF $p(x)$ is replaced by the PDF $f(x)$.
- In most cases, this integral will be finite, but it does not have to be. There are random variables which have infinite expectation.
- If $f(x)$ is zero outside of an interval (a, b) we have to integrate *only* over that interval.

$$E(X) = \int_a^b x \cdot f(x)dx$$



PROPOSITION: If X is a continuous random variable with PDF $f(x)$ and $h(X)$ is any function of X, then

$$E[h(X)] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) \cdot f(x)dx$$

Especially, it is $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$.

Poll Question 7.5

PROPOSITION: All other formulas concerning expectation and variance still hold:

$$V(X) = E(X^2) - E(X)^2$$

$$E(aX + b) = aE(X) + b$$

$$V(aX + b) = a^2 V(X)$$

Example: Let X be a continuous random variable with PDF

$$f(x) = \begin{cases} \frac{3}{8}x^2 & 0 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Sketch a graph of the PDF $f(x)$.

Poll Question 7.6

(b) Find $E(X)$.

(c) Find $E(X^2)$.

(d) Find $V(X)$.