STAT 414 / 415 Probability Theory and Mathematical Statistics

Cumulative Distribution Functions

Printer-friendly version (https://onlinecourses.science.psu.edu/stat414/print/book/export/html/98)

You might recall that the cumulative distribution function is defined for discrete random variables as:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

Again, $F(x)$ accumulates all of the probability less than or equal to x. The cumulative distribution function for continuous random variables is just a straightforward extension of that of the discrete case. All we need to do is replace the summation with an integral.

Definition. The **cumulative distribution function** ("**c.d.f.**") of a continuous random variable X is defined as:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

for $-\infty < x < \infty$.

You might recall, for discrete random variables, that $F(x)$ is, in general, a non-decreasing *step* function. For continuous random variables, $F(x)$ is a non-decreasing *continuous* function.

Example

Let's return to the example in which X has the following probability density function:

$$f(x) = 3x^2$$

for $0 < x < 1$. What is the cumulative distribution function $F(x)$?

Example

Let's return to the example in which X has the following probability density function:

$$f(x) = \frac{x^3}{4}$$

for $0 < x < 2$. What is the cumulative distribution function of X?

Example

Suppose the p.d.f. of a continuous random variable X is defined as:

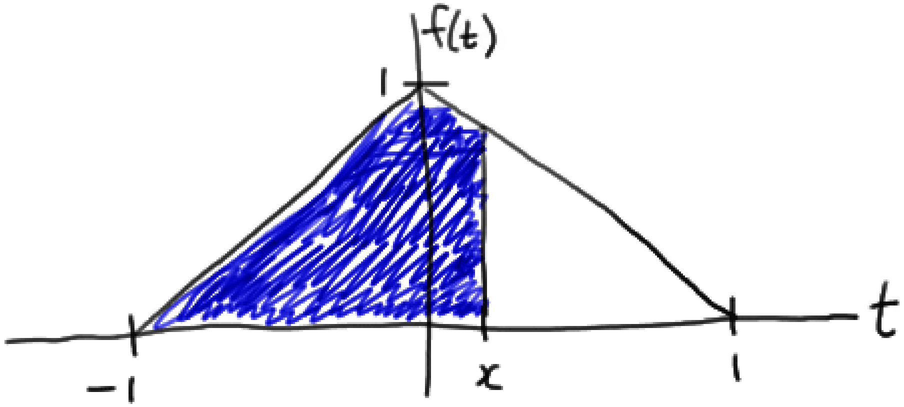
$$f(x) = x + 1$$

for $-1 < x < 0$, and

$$f(x) = 1 - x$$

for $0 \leq x < 1$. Find and graph the c.d.f. $F(x)$.

Solution. If we look at a graph of the p.d.f. $f(x)$:



we see that the cumulative distribution function $F(x)$ must be defined over four intervals — for $x \le -1$, when $-1 < x \le 0$, for $0 < x < 1$, and for $x \ge 1$. The definition of $F(x)$ for $x \le -1$ is easy. Since no probability accumulates over that interval, $F(x) = 0$ for $x \le -1$. Similarly, the definition of $F(x)$ for $x \ge 1$ is easy. Since all of the probability has been accumulated for x beyond 1, $F(x) = 1$ for $x \ge 1$. Now for the other two intervals:

In summary, the cumulative distribution function defined over the four intervals is:

$$F(x) = \begin{cases} 0, & \text{for } x \le -1 \\ \frac{1}{2}(x+1)^2, & \text{for } -1 < x \le 0 \\ 1 - \frac{(1-x)^2}{2}, & \text{for } 0 < x < 1 \\ 1, & \text{for } x \geq 1 \end{cases}$$

The cumulative distribution function is therefore a concave up parabola over the interval $-1 < x \le 0$ and a concave down parabola over the interval $0 < x < 1$. Therefore, the graph of the cumulative distribution function looks something like this: