Probability and Statistics

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

March 20, 2019

Notes

Topics

Notes

Objectives

Objectives:

- Obtain the cumulative distribution function from a probability density function.
- Use a cumulative distribution function to find probabilities.
- Obtain the probability density function from a cumulative distribution function.

Notes

Cumulative Distribution Functions: Continuous (4.2)

- The ***cumulative distribution function*** (or ***cdf***) of a random variable X, denoted $\boldsymbol{F(x)}$, is defined for all x as

$$F(x) \;=\; P(X \leq x).$$

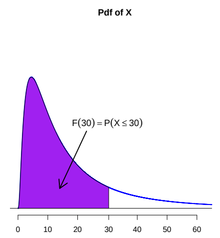
  We'll focus on the case in which X is **continuous**.

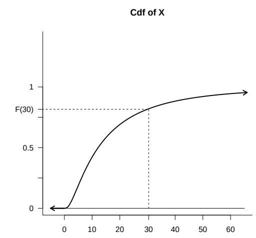
  (But we'll also briefly look at the **discrete** case on the last few slides).

- If X is **continuous** with **pdf** $f(x)$, then

$$F(x) \;=\; \int_{-\infty}^{x} f(y)\; dy.$$

Notes





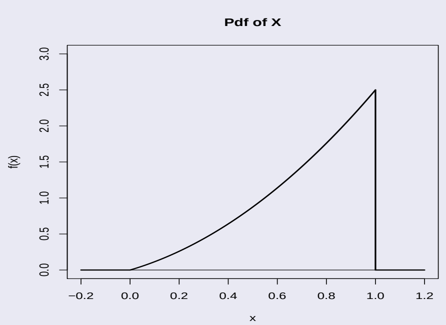
Notes

Example

Suppose (again) that X is the gain in a certain investment, in thousands of dollars, and has **pdf**

$$f(x) \;=\; \begin{cases} \frac{1}{2}(3x^2 + 2x) & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Notes



Notes

For $x < 0$, the **cdf** is

$$F(x) \;=\; 0$$

because there's no area under the **pdf** to the left of 0.

For $x > 1$,

$$F(x) \;=\; 1$$

because all of the area under the **pdf** is accumulated to the left 1.

Notes

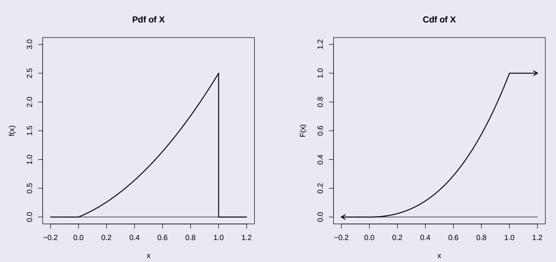
For $\mathbf{0 \leq x \leq 1}$, the **cdf** is

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(y)\,dy \\ &= \int_0^x \tfrac{1}{2}(3y^2 + 2y)\,dy \\ &= \tfrac{1}{2}(y^3 + y^2)\Big|_0^x \\ &= \tfrac{1}{2}(x^3 + x^2). \end{aligned}$$

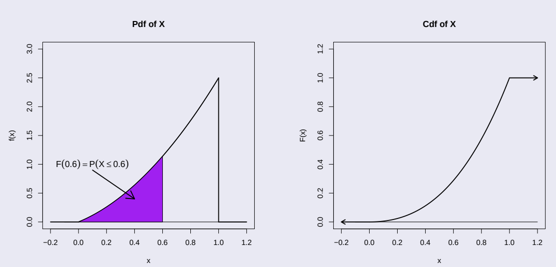
Thus the **cdf** of X is

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}(x^3 + x^2) & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}$$

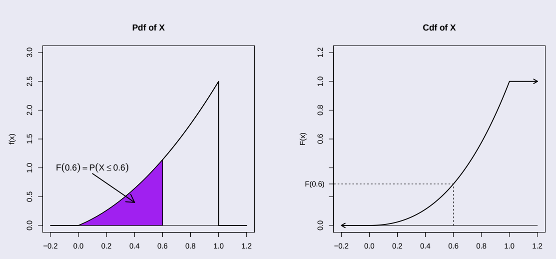
Notes



Notes



Notes



Notes

The probability $P(X \leq 0.6)$ that the investment gain will be **less than $600** is

$$\begin{aligned} P(X \leq 0.6) &= F(0.6) \\ &= \tfrac{1}{2}(0.6^3 + 0.6^2) \\ &= \mathbf{0.288}. \end{aligned}$$

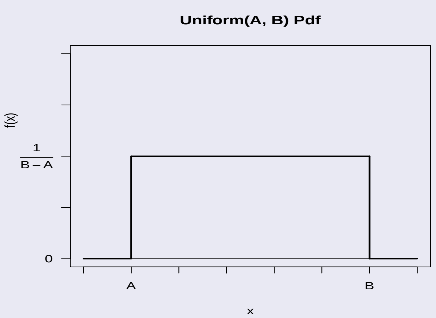
Notes

Example

Suppose X, the thickness of a certain metal sheet, follows a **uniform distribution** over the range $\boldsymbol{A}$ to $\boldsymbol{B}$.

The **pdf** is below.

Notes



Notes

For $\boldsymbol{x < A}$, the **cdf** is

$$F(x) = 0$$

because there's no area under the **pdf** to the left of A.

For $\boldsymbol{x > B}$,

$$F(x) = 1$$

because all of the area under the **pdf** is accumulated to the left B.

Notes

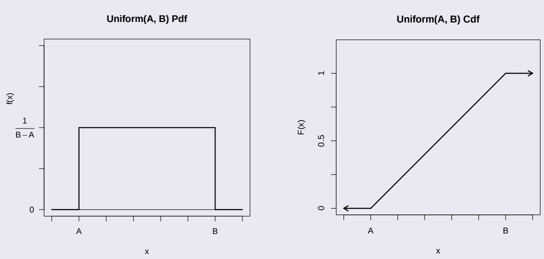
For $A \le x \le B$, the **cdf** is

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(y)\,dy \\ &= \int_{A}^{x} \frac{1}{B-A}\,dy \\ &= \left. \frac{y}{B-A} \right|_{A}^{x} \\ &= \frac{x-A}{B-A}. \end{aligned}$$

Thus the **cdf** of a **uniform** random variable is

$$F(x) = \begin{cases} 0 & \text{for } x < A \\ \frac{x-A}{B-A} & \text{for } A \le x \le B \\ 1 & \text{for } x > B \end{cases}$$

Notes



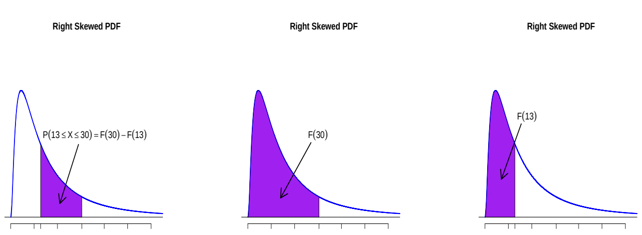
Notes

Proposition

Suppose X is a random variable with **cdf** $F(x)$. Then for any numbers a and b, with $a \le b$,

1. $P(X > a) \; = \; 1 - F(a)$
2. $P(a < X \le b) \; = \; F(b) - F(a)$

Notes



Notes

Example

Suppose (again) that X is the gain in a certain investment, in thousands of dollars.

We found that the **cdf** of X is

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}(x^3 + x^2) & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}$$

Notes

The probability $P(X > 0.6)$ that it will be **more than $600** is

$$\begin{aligned} P(X > 0.6) &= 1 - F(0.6) \\ &= 1 - \tfrac{1}{2}(0.6^3 + 0.6^2) \\ &= 1 - 0.288 \\ &= \mathbf{0.712}. \end{aligned}$$

Notes

The probability $P(0.4 \leq X \leq 0.8)$ that the investment gain will be **between $400 and $800** is

$$\begin{aligned} P(0.4 \leq X \leq 0.8) &= F(0.8) - F(0.4) \\ &= \tfrac{1}{2}(0.8^3 + 0.8^2) - \tfrac{1}{2}(0.4^3 + 0.4^2) \\ &= \mathbf{0.464}. \end{aligned}$$

Notes

- We obtain the **pdf** $\boldsymbol{f(x)}$ by taking the **derivative** of the **cdf** $\boldsymbol{F(x)}$.

Proposition

Suppose X is a continuous random variable with **cdf** $F(x)$ and **pdf** $f(x)$. Then at every value x for which the derivative $F'(x)$ exists,

$$F'(x) = f(x).$$

This is a consequence of the Fundamental Theorem of Calculus, which says

$$\frac{d}{dx} \int_{\infty}^{x} f(y)\, dy = f(x).$$

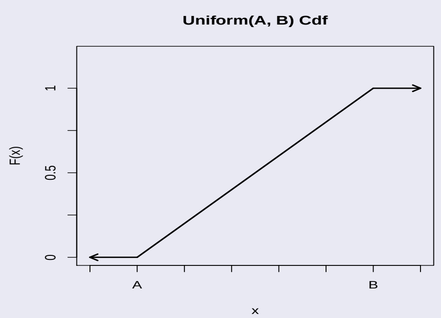
Notes

Example

Recall that the **cdf** of a **uniform** random variable over the range A to B is

$$F(x) = \begin{cases} 0 & \text{for } x < A \\ \frac{x-A}{B-A} & \text{for } A \le x \le B \\ 1 & \text{for } x > B \end{cases}$$

Notes



Notes

The **pdf** is obtained from the **cdf** by taking its derivative:

$$f(x) = F'(x) = \begin{cases} \frac{d}{dx}(0) & \text{for } x < A \\ \frac{d}{dx}\left(\frac{x-A}{B-A}\right) & \text{for } A < x < B \\ \frac{d}{dx}(1) & \text{for } x > B \end{cases}$$

$$= \begin{cases} \frac{1}{B-A} & \text{for } A < x < B \\ 0 & \text{otherwise} \end{cases}$$

Notes

Cumulative Distribution Functions: Discrete (3.2)

- Recall that the ***cumulative distribution function*** (or ***cdf***) of a random variable X, denoted $\boldsymbol{F(x)}$, is defined for all x as
$$F(x) = P(X \le x).$$
- If X is **discrete** with **pmf** $p(x)$, then
$$F(x) = \sum_{y \le x} p(y).$$
- In the **discrete** case, $F(x)$ is a **step function**, with a step of size $p(x)$ at each of the possible values x of the random variable X.

Notes

Example

Suppose the **pmf** of a **discrete** random variable X is

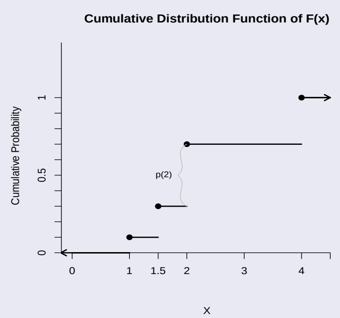
x	1	1.5	2	4
$p(x)$	0.1	0.2	0.4	0.3

Then the **cdf** $F(x)$ is a **step function**, with a step of size $p(x)$ at each of the values $x = 1$, 1.5, 2, and 4. We can write this as

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ 0.1 & \text{for } 1 \leq x < 1.5 \\ 0.3 & \text{for } 1.5 \leq x < 2 \\ 0.7 & \text{for } 2 \leq x < 4 \\ 1 & \text{for } x \geq 4 \end{cases}$$

A graph of $F(x)$ is shown on the next slide.

Notes



Notes

Notes

Notes