Probability and Statistics I

14. Continuous Random Variables and Probability Distributions

Revised 31/10/24

4.1 Continuous Random Variables

Probability Density Functions

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 with $a \leq b$,

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

For $f(x)$ must satisfy two properties to be a valid pdf:

1. $f(x) \geq 0$ for all $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

Cumulative Distribution Function

The cumulative distribution function $F(x)$ for a continuous random variable with pdf $f(x)$ is defined as

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

for every $x \in \mathbb{R}$.

Alternatively, if X is a continuous random variable with cdf $F(x)$ then *one possible*[1] pdf of X is given by

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

for every $x \in \mathbb{R}$ at which $F'(x)$ exists.

[1] The proposition on page 166 of the textbook is not entirely accurate. Differentiating the cdf provides one possible pdf, but the pdf is not unique. The details are beyond this course, but I'd be happy to discuss this during my office hours.

Example 14.1

Let

$$f(x) = \begin{cases} 0 & x \leq 0 \\ cx & 0 < x \leq 1 \\ 0 & x > 1 \end{cases}$$

a) Find the value of c such that $f(x)$ is a valid probability density function (pdf).

b) Find the associated cumulative density function (cdf).

c) Compute the probabilities of the following events:

  i) $X \leq .5$

  ii) $X = .5$

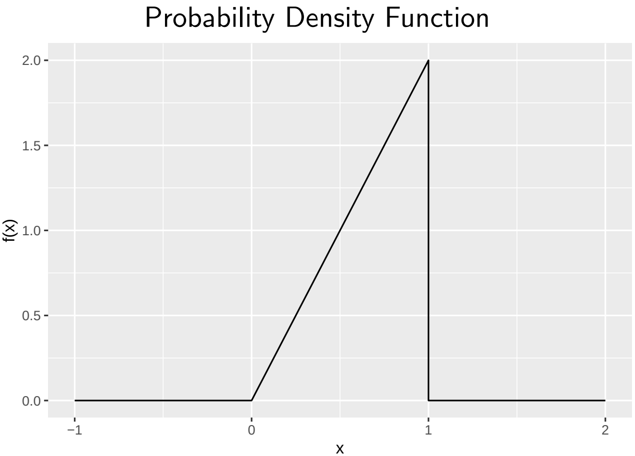
  iii) $X < .5$

  iv) $.25 \leq X \leq .75$

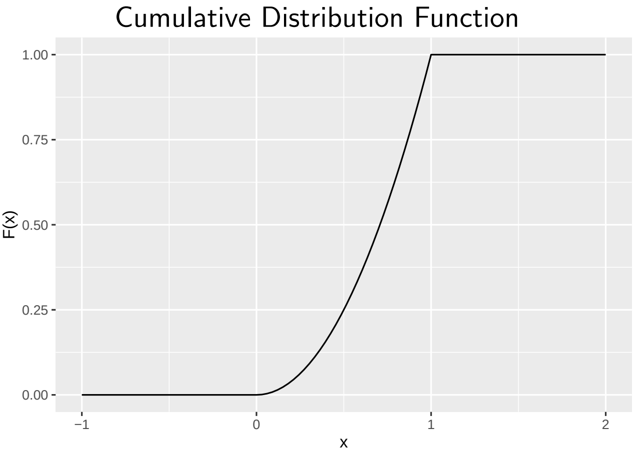
  v) $X < .25$ or $X > .75$

d) Prove that X satisfies the definition of a continuous random variable.

Example 14.1 ctd

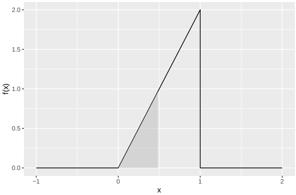


Example 14.1 ctd

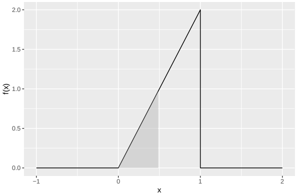


Example 14.1 ctd

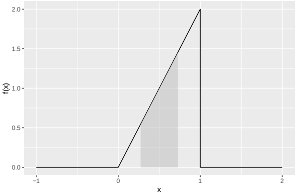
$P(X < .5)$



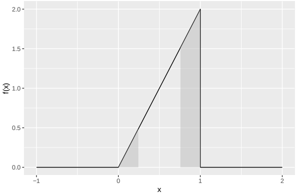
$P(X <= .5)$



$P(.25 \leq X \leq .75)$



$P(X < .25 \text{ or } X > .75)$



Percentiles

The $(100p)th$ percentile of a random variable X is the value η_p such that

$$P(X \leq \eta_p) = F(\eta_p) = p.$$

Alternatively,

$$\eta_p = F^{-1}(p).$$

Median

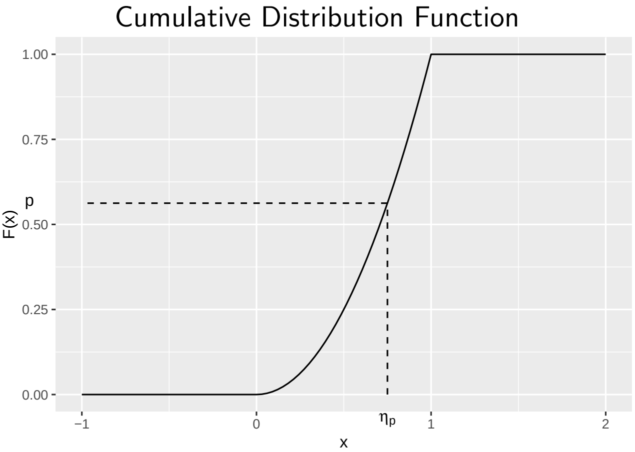
The median is the 50th percentile, $\eta_{.5}$:

$$P(X \leq \eta_{.5}) = F(\eta_{.5}) = .5.$$

Alternatively,

$$\eta_{.5} = F^{-1}(.5).$$

Probabilities and Quantiles



Example 14.2

Let

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 2x & 0 < x \leq 1 \\ 0 & x > 1 \end{cases}$$

a) Find the median of X.

b) Find the 5-th and 95-th percentiles of X.

c) What is the shortest interval, (x_1, x_2), such that $P(x_1 < X < x_2) = .90$?

Questions?

4.1 Continuous Random Variables

Some Comments

Interpreting PDFs

1. Probability
2. Limit of histogram
3. Relative probability

Interpreting PDFs

1. Probability
   For any interval, $\mathcal{I}$, with endpoints $a < b$

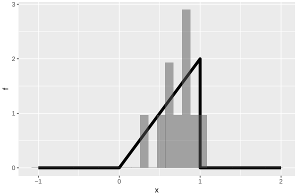
$$P(X \in \mathcal{I}) = F(b) - F(a) = \int_a^b f(u)\, du.$$

Interpreting PDFs

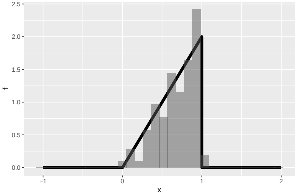
2. Limit of histogram

Limit of Histogram

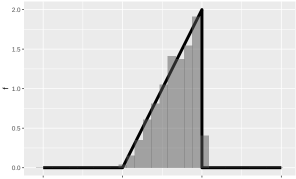
$n = 10$



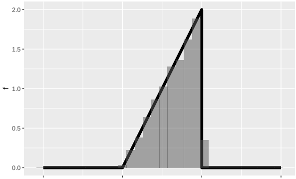
$n = 100$



$n = 1000$



$n = 10000$

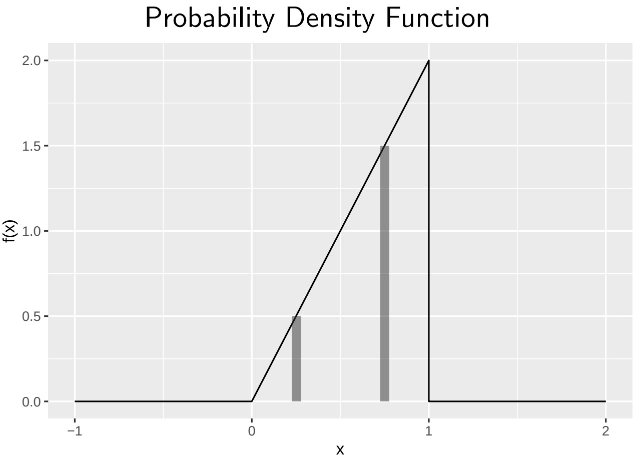


Interpreting PDFs

3. Relative probability
   For any points $x_1, x_2 \in \Re$ at which $f(x)$ is continuous:

$$\frac{P(X \in (x_1, x_1 + \epsilon))}{P(X \in (x_2, x_2 + \epsilon))} \approx \frac{\epsilon f(x_1)}{\epsilon f(x_2)} = \frac{f(x_1)}{f(x_2)}.$$

Example 14.2 ctd



Cumulative Distribution Functions

Any function $F(x)$ on $\mathbb{R}$ is a CDF if

1. $F(x)$ is non-decreasing
2. $\lim_{x \searrow -\infty} F(x) = 0$
3. $\lim_{x \nearrow \infty} F(x) = 1$
4. $F(x)$ is continuous from the right

$$\lim_{x \searrow c} F(x) = F(c)$$

for all $c \in \Re$.

Discrete vs Continuous Random Variables

A random variable/probability distribution is continuous if $F(x)$ is continuous.

A random variable/probability distribution is discrete if $F(x)$ is a step function.

Questions?

Exercise 14.1

Consider the distribution with cdf

$$F(X) = \begin{cases} 0 & x < 0 \\ \log_{10}(x+1) & 0 \leq x < 9 \\ 1 & 9 \leq x \end{cases}$$

a) Plot $F(x)$.

b) Compute the pdf, $f(x)$.

c) Plot $f(x)$.

d) Compute the following probabilities:

   i. $(X \leq \sqrt{10} - 1)$

   ii. $P(X < \sqrt{10} - 1)$

   iii. $P(X = \sqrt{10} - 1)$

   iv. $P(X > \sqrt{10} - 1)$