CS 112: Computer System Modeling Fundamentals

Prof. Jenn Wortman Vaughan

April 19, 2011

Lecture 7

Reminders & Announcements

- Check your grade for homework 1 on Gradebook
- Homework 2 is due on Tuesday, April 26
- Check the homework 1 solutions to see how you should structure your own homework solutions – we'll be less generous about sloppy or confusing answers next time!!
- There is no Friday section this week

Last Time

What if we have multiple random variables associated with the same event?

- Joint PMFs
- Conditioning
- Independence

Remember, most of this was just new notation for concepts we already knew. When in doubt, think about the underlying events!

Today

A bit more about joint PMFs

- Variance of sums of independent random variables
- Covariance and correlation

Continuous random variables

- Probability density functions (PDFs)
- Uniform continuous random variables
- Exponential random variables
- Cumulative distribution functions (CDFs)

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

- This is equivalent to showing that for all k, l s.t. $p_Y(l)>0$

$$p_X(k) = p_{X|Y}(k \mid l)$$

- This is equivalent to showing that for all k, l s.t. $p_X(k)>0$

$$p_Y(l) = p_{Y|X}(l \mid k)$$

To prove or disprove independence, it is enough to prove or disprove any one of these conditions.

Why is independence useful?

If $X_1, X_2, \ldots, X_n$ are all independent, then

$$E[X_1 X_2 \ldots X_n] = ???$$

Why is independence useful?

If $X_1, X_2, \ldots, X_n$ are all independent, then

$$E[X_1 X_2 \ldots X_n] = E[X_1]E[X_2]\ldots E[X_n]$$

Why is independence useful?

If $X_1, X_2, \ldots, X_n$ are all independent, then

$$E[X_1X_2\ldots X_n] = E[X_1]E[X_2]\ldots E[X_n]$$

and

$$var(X_1 + X_2 + \ldots + X_n) = ???$$

Why is independence useful?

If $X_1, X_2, \ldots, X_n$ are all independent, then

$$E[X_1 X_2 \ldots X_n] = E[X_1]E[X_2]\ldots E[X_n]$$

and

$$\mathrm{var}(X_1 + X_2 + \ldots + X_n) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \ldots + \mathrm{var}(X_n)$$

The Sample Mean

Suppose we would like to estimate the president's approval rating. We ask n random voters whether or not they approve of the president, and use the fraction of voters who say that they approve as our estimate.

How accurate is our estimate as a function of n?

The Sample Mean

Suppose we would like to estimate the president's approval rating. We ask n random voters whether or not they approve of the president, and use the fraction of voters who say that they approve as our estimate.

How accurate is our estimate as a function of n?

These ideas will come up again later in the course…

Covariance

Covariance is a notion that measures the strength and direction of the relationship between two variables.

Covariance

Covariance is a notion that measures the strength and direction of the relationship between two variables.

Experiment: Choose a random UCLA student.

- Let X be his height and Y be his weight. These have positive covariance or are positively correlated.

Covariance

Covariance is a notion that measures the strength and direction of the relationship between two variables.

Experiment: Choose a random UCLA student.

- Let X be his height and Y be his weight. These have positive covariance or are positively correlated.
- Let X be the number of courses he is taking and Y be the average number of hours he sleeps each night. These have negative covariance or are negatively correlated.

Covariance

Covariance is a notion that measures the strength and direction of the relationship between two variables.

Experiment: Choose a random UCLA student.

- Let X be his height and Y be his weight. These have positive covariance or are positively correlated.
- Let X be the number of courses he is taking and Y be the average number of hours he sleeps each night. These have negative covariance or are negatively correlated.
- Let X be his height and Y be the last digit of his UID. These have (roughly) zero covariance or are uncorrelated.

Covariance

Formally, the covariance of two random variables X and Y is defined as

$$\mathrm{cov}(X, Y) = E[\ (X - E[X])\ (Y - E[Y])\ ]$$

Covariance

Formally, the covariance of two random variables X and Y is defined as

$$\text{cov}(X, Y) = E[\ (X–E[X])\ (Y–E[Y])\ ]$$
$$= E[XY] – E[X]E[Y]$$

Covariance

Formally, the covariance of two random variables X and Y is defined as

$$\mathrm{cov}(X, Y) = E[\,(X-E[X])\,(Y-E[Y])\,]$$
$$= E[XY] - E[X]E[Y]$$

- What is cov(X, X)?

Covariance

Formally, the covariance of two random variables X and Y is defined as

$$\text{cov}(X, Y) = E[\,(X - E[X])\,(Y - E[Y])\,]$$
$$= E[XY] - E[X]E[Y]$$

- What is cov(X, X)?
- Suppose X and Y are independent. What is cov(X, Y)?

Covariance

Formally, the covariance of two random variables X and Y is defined as

$$\mathrm{cov}(X, Y) = E[\,(X - E[X])\,(Y - E[Y])\,]$$
$$= E[XY] - E[X]E[Y]$$

- What is cov(X, X)?
- Suppose X and Y are independent. What is cov(X, Y)?
- Suppose cov(X, Y) = 0. Are X and Y independent?

Continuous Random Variables

What if a random variable X can take on a continuum of different values?

- Exact execution time of a task
- Exact lifetime of a component
- Time between two requests to the Google server

Continuous Random Variables

What if a random variable X can take on a continuum of different values?

- Exact execution time of a task
- Exact lifetime of a component
- Time between two requests to the Google server

PMFs don't quite make sense anymore…

Continuous Random Variables

What if a random variable X can take on a continuum of different values?

- Exact execution time of a task
- Exact lifetime of a component
- Time between two requests to the Google server

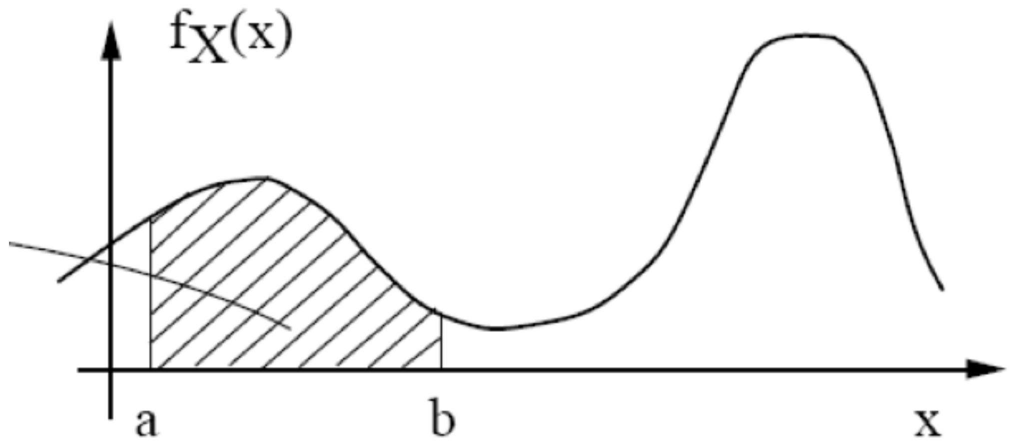
PMFs don't quite make sense anymore…

- Let X be a random variable whose value is drawn uniformly at random from [0,1]. What is $P(X = 0.5)$?

The Probability Density Function

In place of the PMF, we introduce the probability density function (or PDF), denoted f_X. For any a and any $b \geq a$,

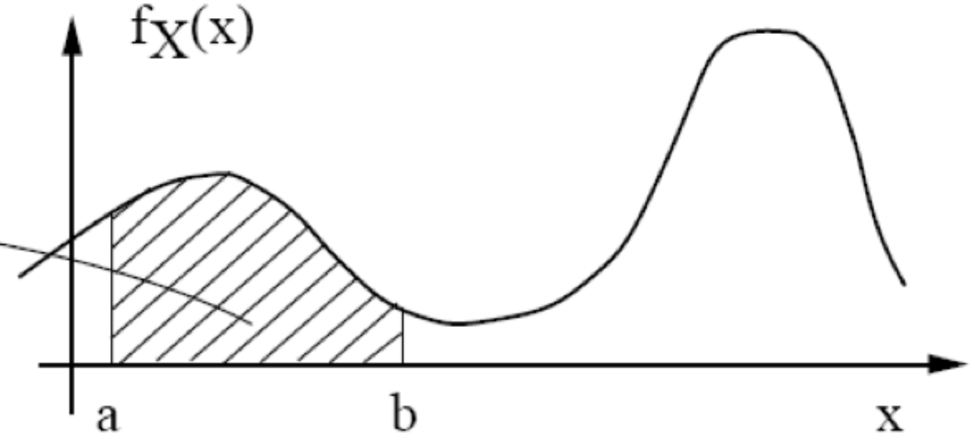
$$P(a \leq X \leq b) = \int_a^b f_X(k)\,dk$$



The Probability Density Function

In place of the PMF, we introduce the probability density function (or PDF), denoted f_X. For any a and any $b \geq a$,

$$P(a \leq X \leq b) = \int_a^b f_X(k)\,dk$$



To satisfy normalization, we need $\int_{-\infty}^{\infty} f_X(k)\,dk = 1$

The Probability Density Function

For very small values of δ, we can approximate

$$P(a \leq X \leq a + \delta) = \int_{a}^{a+\delta} f_X(k)dk \approx f_X(a)\delta$$

This gives us a nice interpretation of $f_X(a)$ as the probability mass per unit length near a

The Probability Density Function

For very small values of δ, we can approximate

$$P(a \le X \le a+\delta) = \int_a^{a+\delta} f_X(k)dk \approx f_X(a)\delta$$

This gives us a nice interpretation of $f_X(a)$ as the probability mass per unit length near a

Need to be careful with this interpretation though!!! Note that, for example, $f_X(a)$ can be bigger than 1…

Expectations

Expectation and variance of continuous random variables are similar to the discrete case.

$$E[X] = ???$$

Expectations

Expectation and variance of continuous random variables are similar to the discrete case.

$$E[X] = \int_{-\infty}^{\infty} k f_X(k)\,dk$$

Expectations

Expectation and variance of continuous random variables are similar to the discrete case.

$$E[X] = \int_{-\infty}^{\infty} k f_X(k) dk$$

$$E[g(X)] = ???$$

Expectations

Expectation and variance of continuous random variables are similar to the discrete case.

$$E[X] = \int_{-\infty}^{\infty} k f_X(k) dk$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(k) f_X(k) dk$$

Expectations

Expectation and variance of continuous random variables are similar to the discrete case.

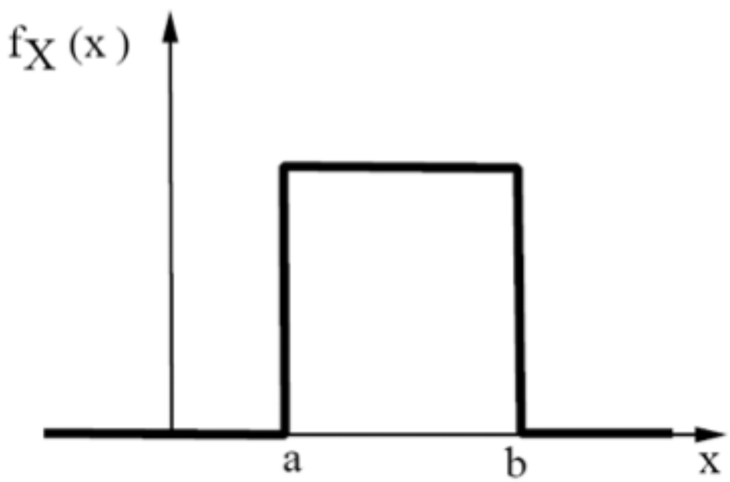
$$E[X] = \int_{-\infty}^{\infty} k f_X(k)\,dk$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(k) f_X(k)\,dk$$

$$\mathrm{var}(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (k - E[X])^2 f_X(k)\,dk$$

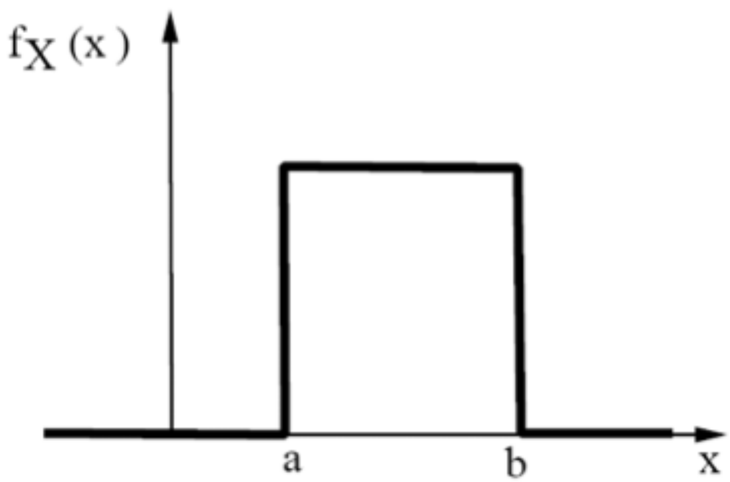
Uniform PDFs

Let X be a random variable whose value is drawn uniformly at random from the range [a, b]



Uniform PDFs

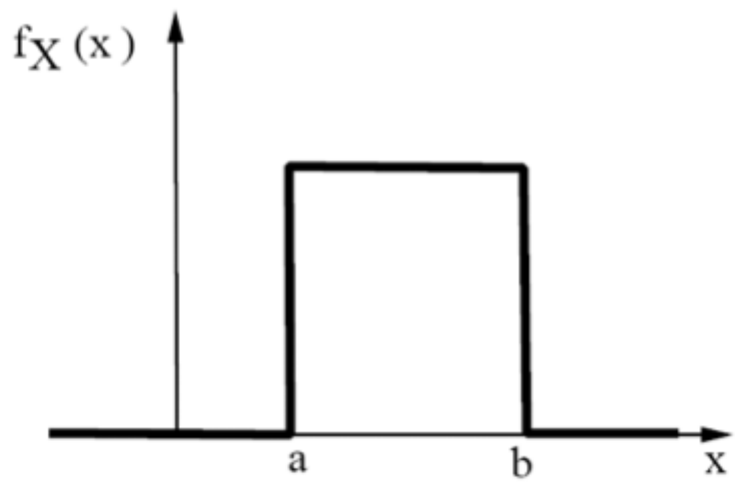
Let X be a random variable whose value is drawn uniformly at random from the range [a, b]



- What is f_X? What is E[X]? What is var(X)?

Uniform PDFs

Let X be a random variable whose value is drawn uniformly at random from the range $[a, b]$



- What is f_X? What is E[X]? What is var(X)?
- What if X was chosen from the range (a, b) instead?

Exponential Random Variables

Exponential random variables model the amount of time until an incident of interest takes place

- Length of time before a message arrives at the computer
- Length of time before a light bulb burns out

Exponential Random Variables

Exponential random variables model the amount of time until an incident of interest takes place

- Length of time before a message arrives at the computer
- Length of time before a light bulb burns out

PDF: $f_X(k) = \lambda e^{-\lambda k}$

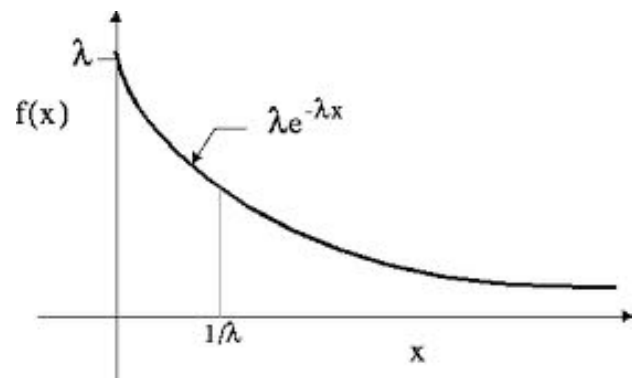


Figure 6. Exponential *pdf*

Exponential Random Variables

Exponential random variables model the amount of time until an incident of interest takes place

- Length of time before a message arrives at the computer
- Length of time before a light bulb burns out

PDF: $f_X(k) = \lambda e^{-\lambda k}$

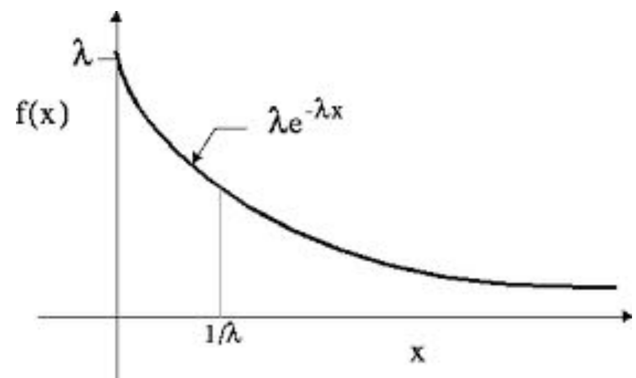


Figure 6. Exponential *pdf*

$E[X] = \lambda^{-1}$ $\quad$ $var(X) = \lambda^{-2}$

These can be verified using integration by parts.

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \leq k)$$

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \leq k)$$

- For discrete random variables,

$$F_X(k) = ???$$

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \leq k)$$

- For discrete random variables,

$$F_X(k) = \sum_{x \leq k} p_X(x)$$

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \leq k)$$

- For discrete random variables,

$$F_X(k) = \sum_{x \leq k} p_X(x)$$

- For continuous random variables,

$$F_X(k) = ???$$

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \leq k)$$

- For discrete random variables,

$$F_X(k) = \sum_{x \leq k} p_X(x)$$

- For continuous random variables,

$$F_X(k) = \int_{-\infty}^{k} f_X(x)dx$$

Cumulative Distribution Functions

A cumulative distribution function (CDF), denoted F_X, "accumulates" probability up to a certain value of X

$$F_X(k) = P(X \le k)$$

- For discrete random variables,

$$F_X(k) = \sum_{x \le k} p_X(x)$$

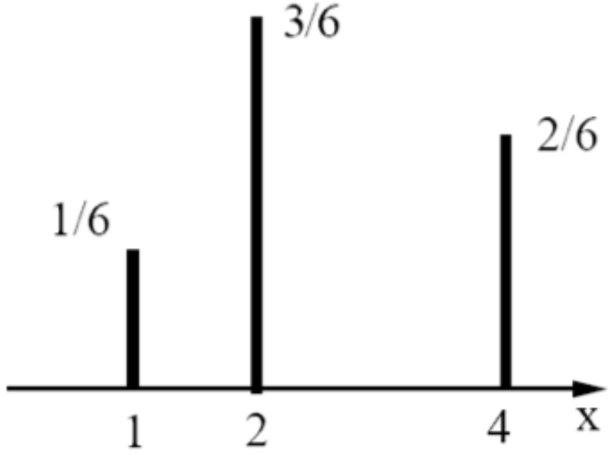
- For continuous random variables,

$$F_X(k) = \int_{-\infty}^{k} f_X(x)dx \quad \rightarrow \quad f_X(k) = \frac{dF_X}{dk}(k)$$

Cumulative Distribution Functions

Consider a discrete random variable X

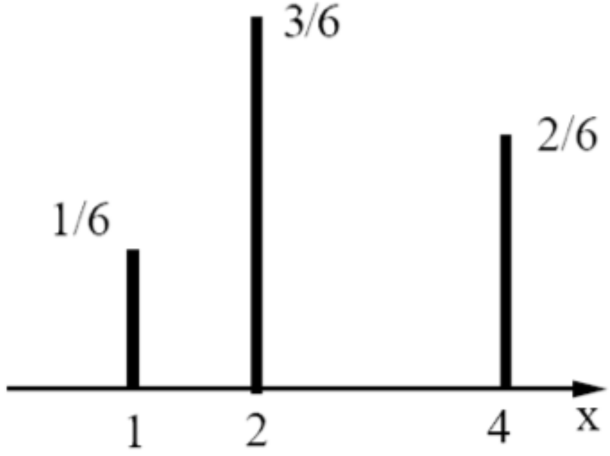
PMF:



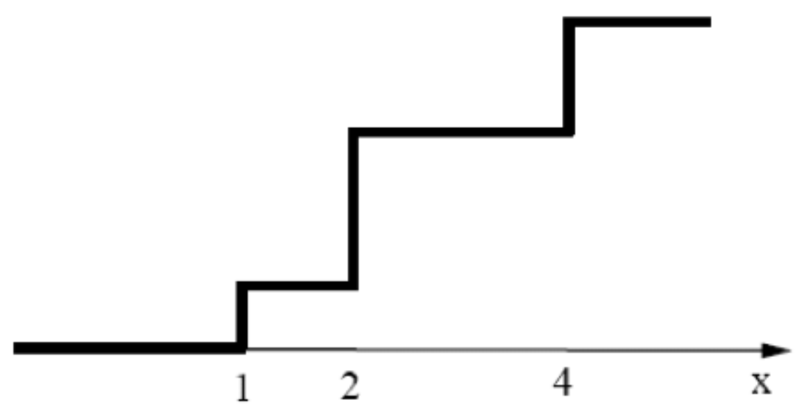
Cumulative Distribution Functions

Consider a discrete random variable X

PMF:

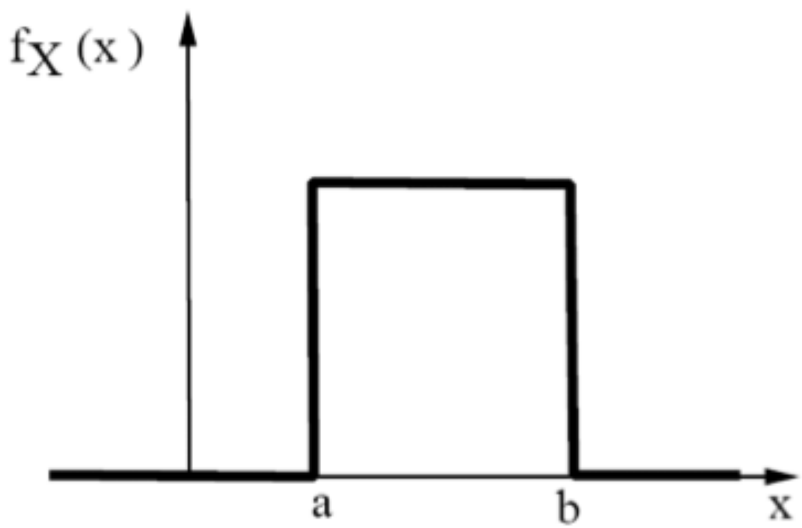


CDF:



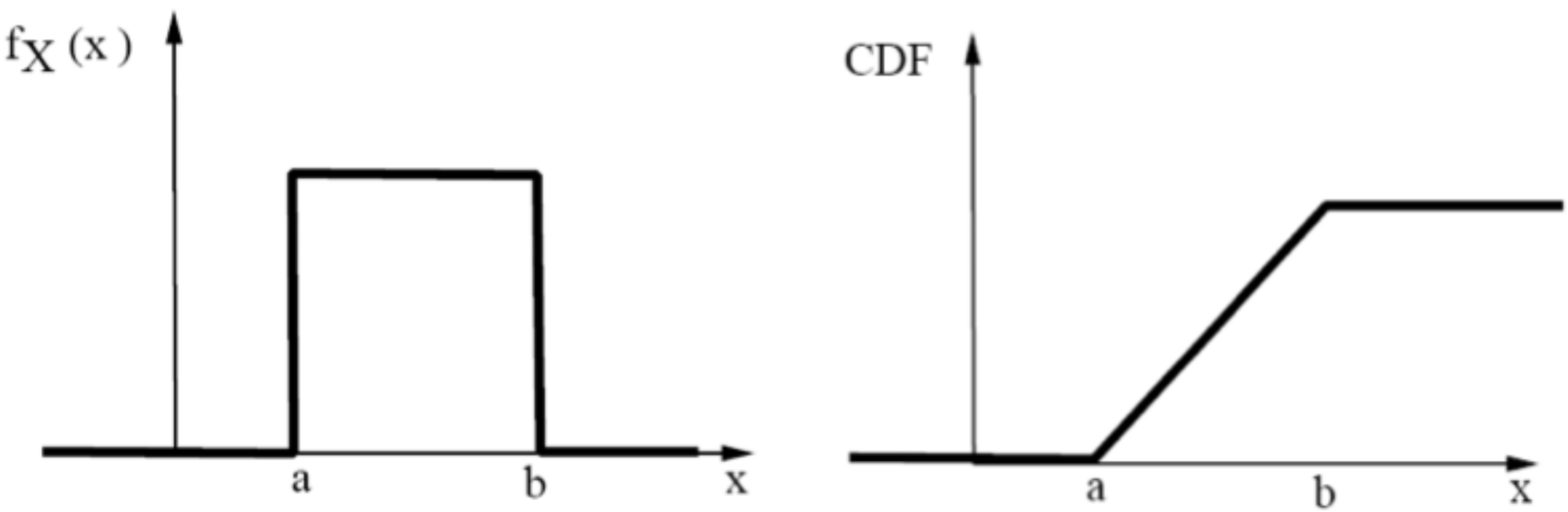
Cumulative Distribution Functions

Consider a uniform continuous random variable X



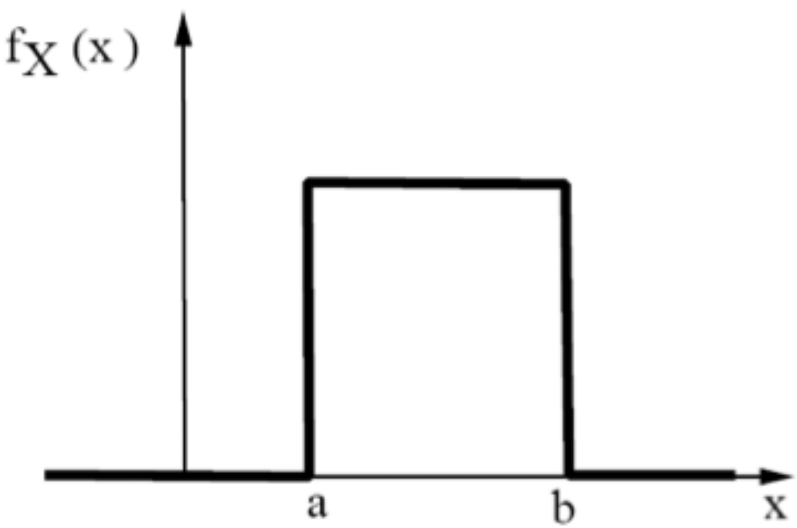
Cumulative Distribution Functions

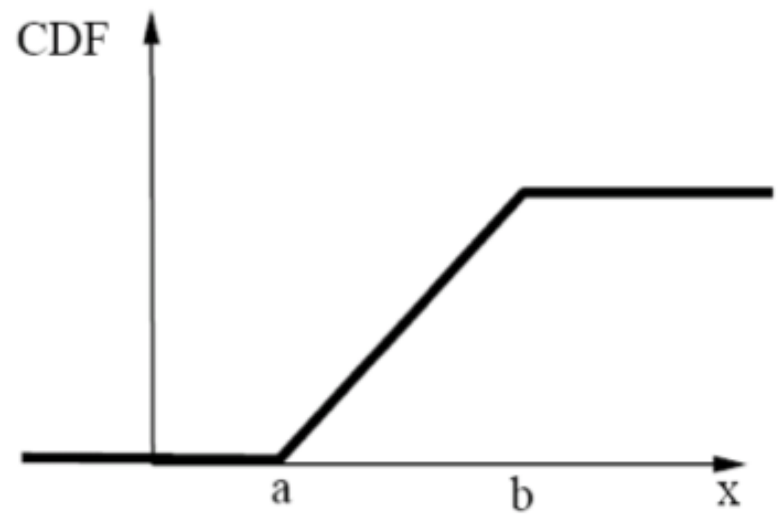
Consider a uniform continuous random variable X



Cumulative Distribution Functions

Consider a uniform continuous random variable X





Note that CDFs are always monotonically non-decreasing

Cumulative Distribution Functions

What is the cumulative distribution function of an exponential random variable with $f_X(k) = \lambda e^{-\lambda k}$?

(we'll start with this next time...)