3 Probability Distributions

(Ch 3.4.1, 3.4.2, 4.1, 4.2, 4.3)

Probability Distribution Functions

Probability distribution function (pdf):

Function for mapping random variables to real numbers.

Discrete random variable:

Values constitute a finite or countably infinite set.

Continuous random variable:

Set of possible values is the set of real numbers **R**, one interval, or a disjoint union of intervals on the real line.

Random Variables

Notation!

1. Random variables - usually denoted by uppercase letters near the end of our alphabet (e.g. X, Y).
2. Particular value - now use lowercase letters, such as x, which correspond to the r.v. X.

Properties of PDFs

For $f(x)$ to be a legitimate pdf, it must satisfy the following two conditions:

1. $f(x) \geq 0$ for all x.

2. $\sum_{all\ x \in X} P(X = x) = 1$ for discrete distributions.

   $\int_{all\ x \in X} f(x)dx = 1$ for continuous distributions.

Discrete Random Variables

PDFs for Discrete RVs

The pdf of a discrete r.v. X describes how the total probability is distributed among all the possible range values of the r.v. X:

$f(x) = p(X=x)$, for each value x in the range of X

Example

A lab has 6 computers.

Let X denote the number of these computers that are in use during lunch hour -- {0, 1, 2… 6}.

Suppose that the probability distribution of X is as given in the following table:

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

Example

From here, we can find many things:

1. Probability that at most 2 computers are in use
2. Probability that at least half of the computers are in use
3. Probability that there are 3 or 4 computers free

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

Bernoulli Distribution

Bernoulli random variable: Any random variable whose only possible values are 0 or 1.

This is a discrete random variable – why?

This distribution is specified with a single parameter:

$$P(X = x) = \pi^{x}(1- \pi)^{(1-x)}; \ x = 0, 1$$

Bernoulli Distribution

Bernoulli random variable: Any random variable whose only possible values are 0 or 1.

This is a discrete random variable – why?

This distribution is specified with a single parameter:

$$P(X = x) = \pi^{x}(1- \pi)^{(1-x)}; \; x = 0, 1$$

Examples?

Geometric Distribution

A patient is waiting for a suitable matching kidney donor for a transplant. The probability that a randomly selected donor is a suitable match is 0.1.

What is the probability the first donor tested is the first matching donor? Second? Third?

Geometric Distribution

Continuing in this way, a general formula for the pmf emerges:

$$P(X = x) = (1 - \pi)^x \pi; \qquad x = 0,1,2, \ldots.$$

The parameter π can assume any value between 0 and 1. Depending on what parameter π is, we get different members of the ***geometric*** distribution.

<u>NOTATION</u>: We write $X \sim G(\pi)$ to indicate that X is a geometric rv with success probability π.

Binomial Distribution

The Binomial r.v. counts the **total number of successes** out of n trials, where X is the number of successes.

- Each trial must be independent of the previous experiment.
- The probability of success must be the same for each trial.

Binomial Distribution

Example: A dice is tossed four times. A "success" is defined as rolling a 1 or a 6.

- The probability of success is 1/3.
- What is $P(X = 2)$?
- What is $P(X = 3)$?

Binomial Distribution

Example: A dice is tossed four times. A “success” is defined as rolling a 1 or a 6.

- The probability of success is 1/3.
- What is $P(X = 2)$?
- What is $P(X = 3)$?

Let’s use the probabilities we calculated above to derive the binomial pdf.

Binomial Distribution

Example: A dice is tossed four times. A "success" is defined as rolling a 1 or a 6.

- The probability of success is 1/3.
- What is $P(X = 2)$?
- What is $P(X = 3)$?

Let's use the probabilities we calculated above to derive the binomial pdf.

NOTATION: We write $X \sim \text{Bin}(n, \pi)$ to indicate that X is a binomial rv based on n Bernoulli trials with success probability π.

The Negative Binomial Distribution

Consider the dice example for the binomial distribution.

Now we instead want to find the probability that we roll 3 “failures” (i.e. a 2, 3, 4, or 5) before the 2^{nd} success.

How is this related to the binomial distribution?

The Negative Binomial Distribution

Consider the dice example for the binomial distribution. What is the probability that exactly 3 successes occur before 2 failures occur?

NOTATION: We write $X \sim \text{NB}(r, \pi)$ to indicate that X is a negative binomial r.v., with x failures occurring before r successes, where the probability of success is equal to π.

The Poisson Probability Distribution

A Poisson r.v. describes the total number of events that happen in a certain <u>time period</u>.

Examples:

- # of vehicles arriving at a parking lot in one week
- # of gamma rays hitting a satellite per hour
- # of cookies chips in a length of cookie dough

The Poisson Probability Distribution

A Poisson r.v. describes the total number of events that happen in a certain time period.

Examples:

- # of vehicles arriving at a parking lot in one week
- # of gamma rays hitting a satellite per hour
- # of cookies sold at a bake sale in 1 hour

The Poisson Probability Distribution

A Poisson r.v. describes the total number of events that happen in a certain time period.

A discrete random variable X is said to have a **Poisson distribution** with parameter λ ($\lambda > 0$) if the pdf of X is

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}; \qquad x = 0, 1, 2, \ldots.$$

NOTATION: We write $X \sim P(\lambda)$ to indicate that X is a Poisson r.v. with parameter λ.

Example

Let X denote the number of mosquitoes captured in a trap during a given time period.

Suppose that X has a Poisson distribution with $\lambda = 4.5$.

What is the probability that the trap contains 5 mosquitoes?

Example problem

Cumulative Distribution Functions

Definition: The **cumulative distribution function** (*cdf*) is denoted with $F(x)$.

For a discrete *r.v.* X with pdf $f(x) = P(X = x)$, $F(x)$ is defined for every real number x by

$$F(x) = P(X \le x) = \sum_{y : y \le x} P(X = y)$$

For any number x, the cdf $F(x)$ is the probability that the observed value of X will be at most x.

Example

Suppose we are given the following pmf:

$$p(x) = \begin{cases} .500 & x = 0 \\ .167 & x = 1 \\ .333 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

Then, calculate:
F(0), F(1), F(2)

What about F(1.5)? F(20.5)?

Is P(X < 1) = P(X <= 1)?

Continuous Random Variables

Continuous Random Variables

A random variable X is **continuous** if possible values comprise either a single interval on the number line or a union of disjoint intervals.

Example: If in the study of the ecology of a lake, X, the r.v. may be depth measurements at randomly chosen locations.

Cumulative Distribution Functions

Definition: The **cumulative distribution function** (*cdf*) is denoted with $F(x)$.

For a discrete *r.v.* X with pdf $f(x)$, $F(x)$ is defined for every real number x by

$$F(x) = \int_{y:y \le x} f(y)dy$$

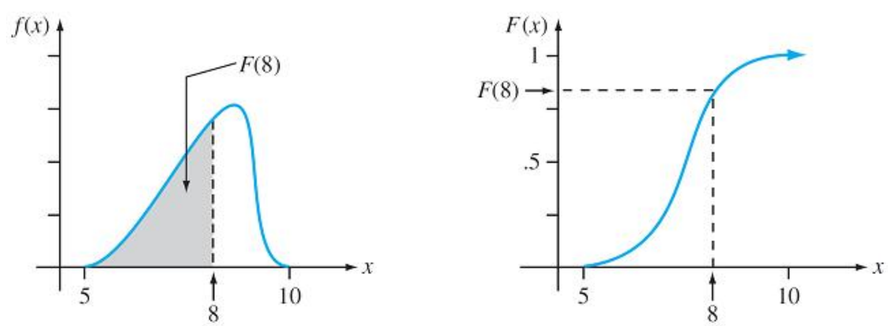
Cumulative Distribution Functions

Definition: The **cumulative distribution function** (*cdf*) is denoted with *F*(*x*).

For a discrete *r.v. X* with pdf *f(x), F(x)* is defined for every real number *x* by

$$F(x) = \int_{y:y \le x} f(y)dy$$

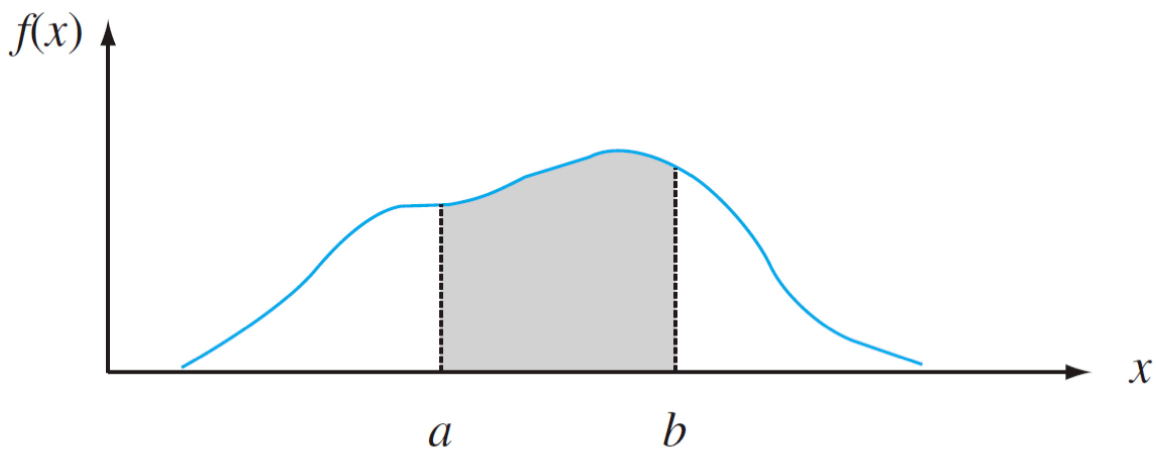
This is illustrated below, where *F*(*x*) increases smoothly as *x* increases.



A pdf and associated cdf

PDFs for Continuous RVs

The probability that X takes on a value in the interval $[a, b]$ is the area above this interval and under the graph of the density function:



$P(a \le X \le b)$ = the area under the density curve between a and b

Example

Consider the reference line connecting the valve stem on a tire to the center point.

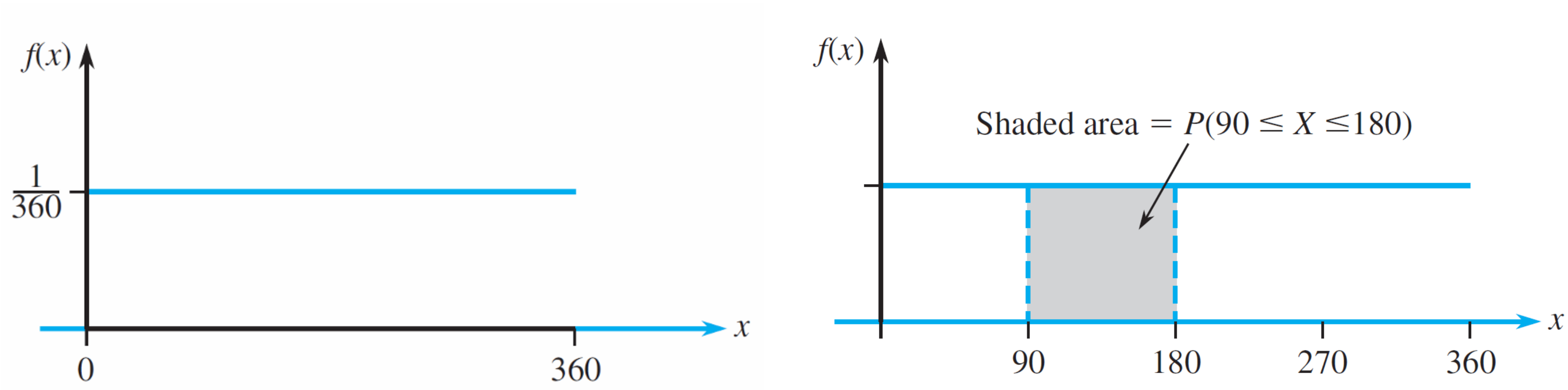
Let *X* be the angle measured clockwise to the location of an imperfection. The pdf for *X* is

$$f(x) = \begin{cases} \dfrac{1}{360} & 0 \leq x < 360 \\ 0 & \text{otherwise} \end{cases}$$

Example

cont' d

The pdf is shown graphically below:



The pdf and probability from example on previous slide.

Example

cont' d

Clearly $f(x) \geq 0$. How can we show that the area of this pdf is equal to 1?

How do we calculate P(90 <= X <= 180)?

What is the probability that the angle of occurrence is within 90° of the reference line? (The reference line is at 0 degrees.)

Uniform Distribution

The previous problem was an example of the **uniform distribution**.

Definition: A continuous rv *X* is said to have a **uniform distribution** on the interval [*a*, *b*] if the pdf of *X* is

$$f(x; a, b) = \frac{1}{b - a}, \quad a \le x \le b$$

Uniform Distribution

The previous problem was an example of the **uniform distribution**.

Definition: A continuous rv *X* is said to have a **uniform distribution** on the interval [*a*, *b*] if the pdf of *X* is

$$f(x; a, b) = \frac{1}{b - a}, \quad a \leq x \leq b$$

<u>**NOTATION**</u>: We write *X* ~ *U(a, b)* to indicate that *X* is a uniform rv with a lower bound equal to *a* and an upper bound equal to *b*.

Exponential Distribution

The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines to describe **time-to-event data**.

Examples?

Exponential Distribution

The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines to describe **time-to-event data**.

Definition: *X* is said to have an exponential distribution with the rate parameter λ ($\lambda > 0$) if the pdf of *X* is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Exponential Distribution

The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines to describe **time-to-event data**.

Definition: *X* is said to have an exponential distribution with the rate parameter λ ($\lambda > 0$) if the pdf of *X* is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

NOTATION: We write $X \sim P(\lambda)$ to indicate that *X* is an Exponential r.v. with parameter λ.

Exponential Distribution

The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines to describe **time-to-event data**.

A partial reason for the popularity of such applications is the **memoryless property** of the Exponential distribution.

The Exponential Distributions

Suppose a light bulb's lifetime is exponentially distributed with parameter λ.

What is the probability that the lifetime of the light bulb lasts less than t hours?

What is the probability that the lifetime of the light bulb lasts more than t hours?

The Exponential Distributions

Suppose a light bulb's lifetime is exponentially distributed with parameter λ.

Now say you turn the light bulb on and then leave. You come back after t_0 hours to find it still on. What is the probability that the light bulb will last for *at least additional t* hours?

In symbols, we are looking for $P(X \geq t + t_0 \mid X \geq t_0)$.

How would we calculate this?

The Weibull Distribution

The Weibull distribution is used similarly to the exponential distribution to model times to an event, but with an extra parameter included for flexibility.

The Weibull Distribution

The Weibull distribution is used similarly to the exponential distribution to model times to an event, but with an extra parameter included for flexibility.

Definition

A random variable X is said to have a **Weibull distribution** with parameters α and β ($\alpha > 0$, $\beta > 0$) if the pdf of X is

$$f(x;\ \alpha, \beta) = \ \alpha\beta^{-\alpha}x^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}}, \qquad x\ \geq 0$$

The Weibull Distribution

The Weibull distribution is used similarly to the exponential distribution to model times to an event, but with an extra parameter included for flexibility.

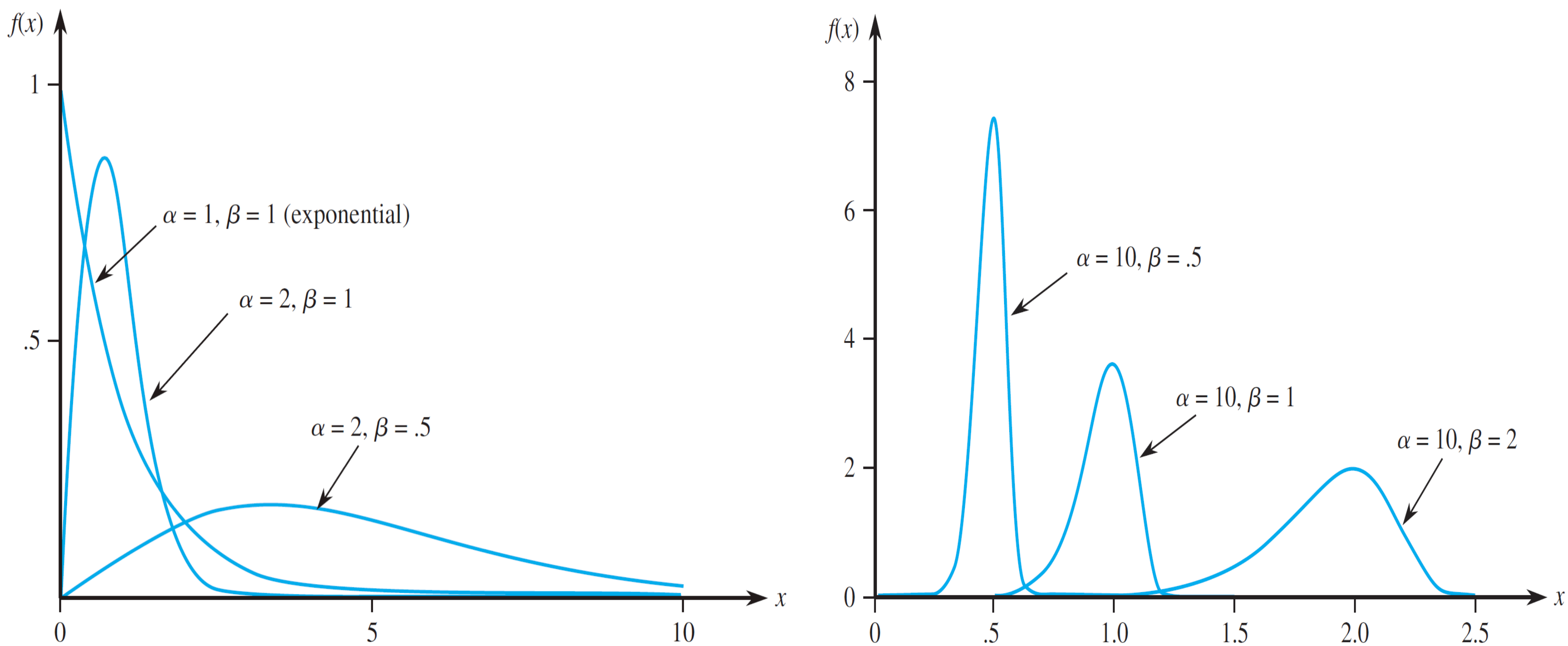
Definition
A random variable X is said to have a **Weibull distribution** with parameters α and β ($\alpha > 0$, $\beta > 0$) if the pdf of X is

$$f(x;\ \alpha,\beta) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}}, \qquad x \geq 0$$

What is this distribution if alpha = 1?

The Weibull Distribution

Both α and β can be varied to obtain a number of different-looking density curves, as illustrated below.



Weibull density curves

The Beta Distribution

So far, all families of continuous distributions (except for the uniform distribution) had positive density over an infinite interval.

The beta distribution provides positive density only for *X* in an interval from 0 to 1.

The beta distribution is commonly used to model variation in the **proportion or percentage** of a quantity occurring in different samples.

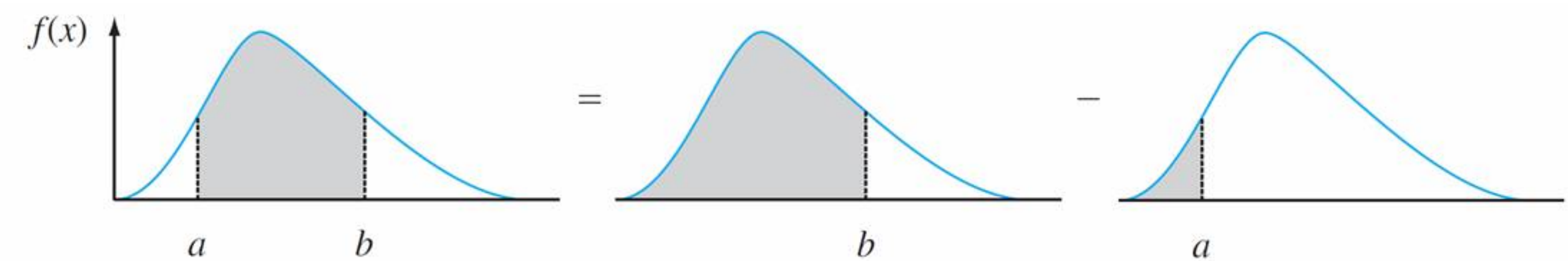
Examples?

The Beta Distribution

Definition
A random variable *X* is said to have a **beta distribution** with parameters α, β (both positive), if the pdf of *X* is:

$$f(x;\ \alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha)}\, x^{\alpha-1}(1-x)^{\beta-1};\ 0\ \leq x\ \leq 1$$

Using $F(x)$ to Compute Probabilities



Example

The distribution of the amount of gravel (in tons) sold by a particular construction supply company in a given week is a continuous rv *X* with pdf

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the cdf of sales for any x? How do you use this to find the probability that X is less than .25? What about the probability that X is greater than .75? What about P(.25 < X < .75)?

Percentiles of a Distribution

Definition

The **median** of a continuous distribution is the 50th percentile of the distribution.

How can we express this in terms of *$F(x)$*?

Example

X is a r.v. such that:

$$f(x) = a^*exp(-ax), x > 0$$

Calculate F(x).

How would you find the median of this distribution?

How do we find the 40th percentile of this distribution?

Example

You flip and unfair coin four times, where the probability of flipping a head is 60%.

What is the median number of heads for this distribution?