Continuous Probability Distributions

Uniform Distribution

Important Terms & Concepts Learned

- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
- Complementary Cumulative Distribution Function (CCDF)

- Expected value
- Mean
- Variance
- Standard deviation

- Uniform distribution
- Bernoulli distribution/trial
- Binomial distribution
- Poisson distribution
- Geometric distribution
- Negative binomial distribution

Which distribution is this?

$$\binom{n}{x} p^x (1-p)^{n-x}$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

Which distribution is this?

$$\binom{n}{x} p^x (1-p)^{n-x}$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

Which distribution is this?

$$\binom{x-1}{r-1}(1-p)^{x-r}p^r$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

Which distribution is this?

$$\binom{x-1}{r-1}(1-p)^{x-r}p^r$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

Which distribution is this?

$$\frac{e^{-\lambda}\lambda^{x}}{x!}$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

Which distribution is this?

$$\frac{e^{-\lambda}\lambda^{x}}{x!}$$

A. Uniform
B. Binomial
C. Geometric
D. Negative Binomial
E. Poisson

Get your i-clickers

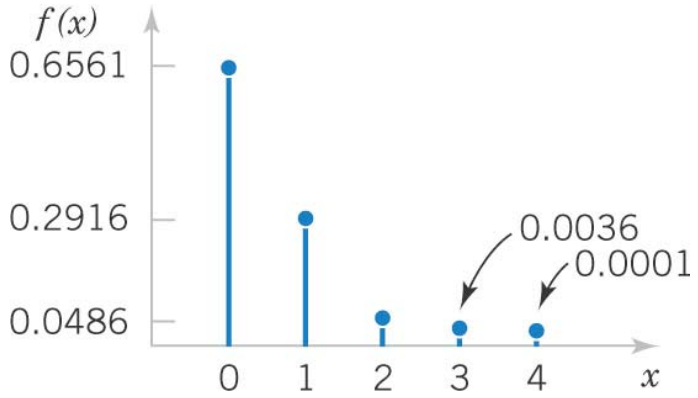
Name	Probability Distribution	Mean	Variance
Discrete			
Uniform	$\frac{1}{n}, a \le b$	$\frac{(b+a)}{2}$	$\frac{(b-a+1)^2-1}{12}$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n, 0 \le p \le 1$	np	$np(1-p)$
Geometric	$(1-p)^{x-1}p$, $x = 1, 2, \ldots, 0 \le p \le 1$	$1/p$	$(1-p)/p^2$
Negative binomial	$\binom{x-1}{r-1}(1-p)^{x-r}p^r$ $x = r, r+1, r+2, \ldots, 0 \le p \le 1$	r/p	$r(1-p)/p^2$
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots, 0 < \lambda$	λ	λ

Continuous & Discrete Random Variables

- A discrete random variable is usually integer number
  - N – the number of proteins in a cell
  - D- number of nucleotides different between two sequences
- A continuous random variable is a real number
  - C=N/V – the concentration of proteins in a cell of volume V
  - Percentage D/L*100% of different nucleotides in protein sequences of different lengths L (depending on set of L's may be discrete but dense)

Probability Mass Function (PMF)

- *X* – discrete random variable



Probability Mass Function for the # of mismatches in 4-mers

- Probability Mass Function: f(x)=P(X=x) – the probability that X is exactly equal to x

P(*X* =0) =	0.6561
P(*X* =1) =	0.2916
P(*X* =2) =	0.0486
P(*X* =3) =	0.0036
P(*X* =4) =	0.0001
$\sum_x$ P(X=x)=	1.0000

Probability Density Function (PDF)

Density functions, in contrast to mass functions, distribute probability continuously along an interval

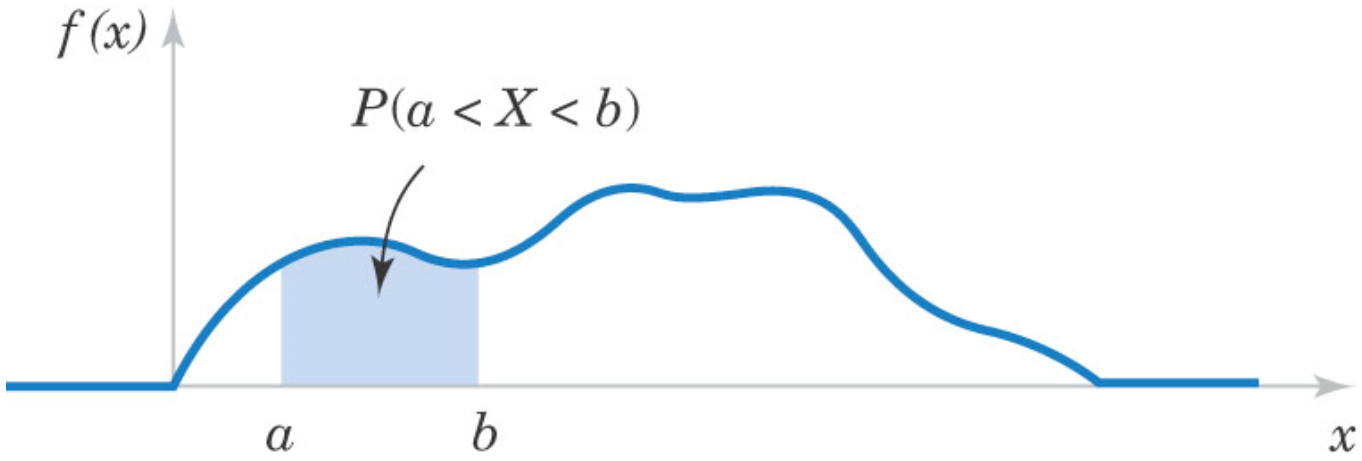


Figure 4-2 Probability is determined from the area under $f(x)$ from a to b.

Probability Density Function

For a continuous random variable X,

a probability density function is a function such that

(1) $f(x) \geq 0$ means that the function is always non-negative.

(2) $\int_{-\infty}^{\infty} f(x)dx = 1$

(3) $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx =$ area under $f(x)dx$ from a to b

EXAMPLE Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the value of C?

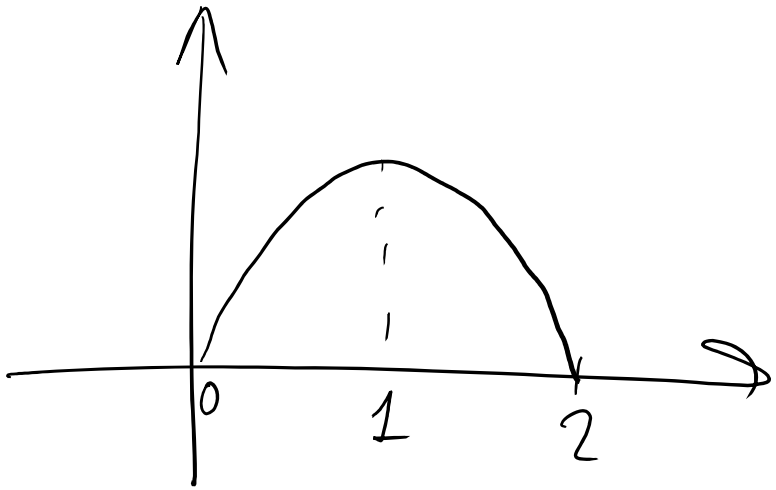
(b) Find $P\{X > 1\}$.

EXAMPLE Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the value of C?

(b) Find $P\{X > 1\}$.



(a) $\int_0^2 C(4x - 2x^2)\,dx = 1$

$C \cdot \left(4 \cdot \frac{2^2}{2} - 2 \cdot \frac{2^3}{3}\right) = 1$

$C \cdot \frac{16}{6} = 1$

$C = \frac{3}{8}$

(b) $P(X > 1) = \int_1^2 f(x)\,dx = \frac{1}{2}$ by symmetry

Histogram approximates PDF

A histogram is graphical display of data showing a series of adjacent rectangles. Each rectangle has a base which represents an interval of data values. The height of the rectangle creates an area which represents the probability of X to be within the base.

When base length is narrow, the histogram approximates f(x) (PDF): height of each rectangle = its area/length of its base.

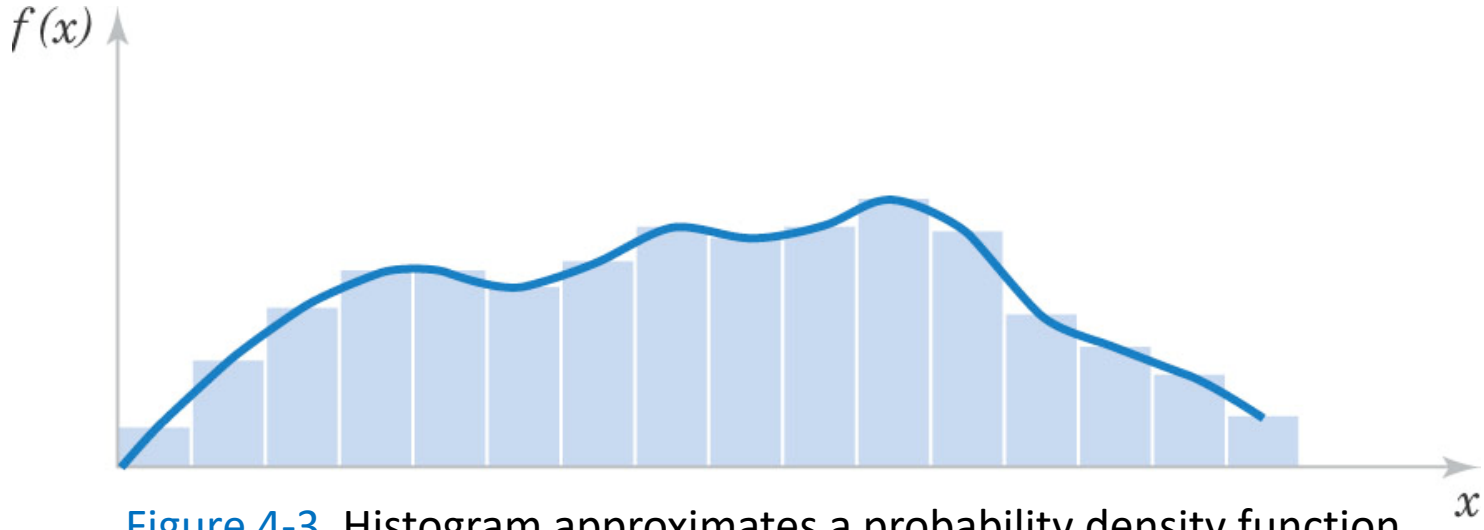


Figure 4-3 Histogram approximates a probability density function.

Cumulative Distribution Functions (CDF & CCDF)

The cumulative distribution function (CDF)

of a continuous random variable X is,

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du \quad \text{for} \quad -\infty < x < \infty \qquad (4\text{-}3)$$

One can also use the inverse cumulative distribution function

or complementary cumulative distribution function (CCDF)

$$F_{>}(x) = P(X > x) = \int_{x}^{\infty} f(u)\,du \quad \text{for} \quad -\infty < x < \infty$$

Definition of CDF for a continous variable is the same as for a discrete variable

Density vs. Cumulative Functions

- The probability density function (PDF) is the derivative of the cumulative distribution function (CDF).

$$f(x) = \frac{dF(x)}{dx} = -\frac{dF_{>}(x)}{dx}$$

as long as the derivative exists.

Mean & Variance

Suppose X is a continuous random variable with probability density function $f(x)$. The mean or expected value of X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \qquad (4\text{-}4)$$

The variance of X, denoted as $V(X)$ or σ^2, is

$$\sigma^2 = V(X) \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2$$

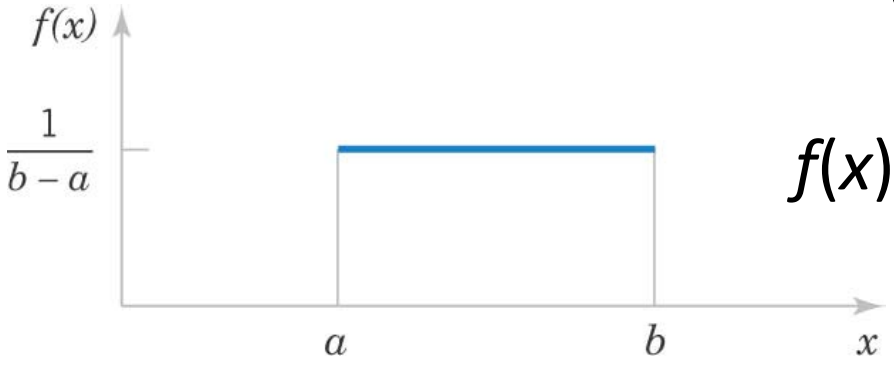
The standard deviation of X is $\sigma = \sqrt{\sigma^2}$.

Gallery of Useful Continuous Probability Distributions

Continuous Uniform Distribution

- This is the simplest continuous distribution and analogous to its discrete counterpart.
- A continuous random variable *X* with probability density function

$$f(x) = 1 / (b-a) \text{ for } a \le x \le b \qquad (4\text{-}6)$$



Compare to discrete

$f(x) = 1/(b-a+1)$

Figure 4-8 Continuous uniform PDF

Comparison between Discrete & Continuous Uniform Distributions

Discrete:

- PMF: $f(x) = 1/(b-a+1)$
- Mean and Variance:

  $\mu = E(x) = (b+a)/2$

  $\sigma^2 = V(x) = [(b-a+1)^2-1]/12$

Continuous:

- PMF: $f(x) = 1/(b-a)$
- Mean and Variance:

  $\mu = E(x) = (b+a)/2$

  $\sigma^2 = V(x) = (b-a)^2/12$

X is a continuous random variable with a uniform distribution between 0 and 3.

What is P(X=1)?

A. 1/4
B. 1/3
C. 0
D. Infinity
E. I have no idea

Get your i-clickers

X is a continuous random variable with a uniform distribution between 0 and 3.

What is P(X=1)?

A. 1/4
B. 1/3
C. 0
D. Infinity
E. I have no idea

Get your i-clickers

X is a continuous random variable with a uniform distribution between 0 and 3.

What is $P(X<1)$?

A. 1/4
B. 1/3
C. 0
D. Infinity
E. I have no idea

Get your i-clickers

X is a continuous random variable with a uniform distribution between 0 and 3.

What is P(X<1)?

A. 1/4
B. 1/3
C. 0
D. Infinity
E. I have no idea

Get your i-clickers

Matlab exercise:

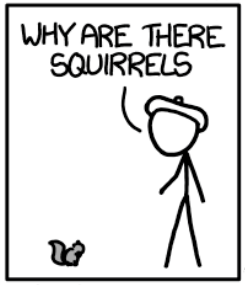
- generate 100,000 random numbers drawn from uniform distribution between 3 and 7
- plot histogram approximating its PDF
- calculate mean, standard deviation and variance

Matlab template: Uniform PDF

- **Stats=????;**
- **r2=???+???.*rand(Stats,1);**
- **disp(mean(r2));**
- **disp(var(r2));**
- **disp(std(r2));**
- **step=0.1;**
- **[a,b]=hist(r2,0:step:8);**
- **pdf_e=a./sum(a).??? (* or /) step;**
- **figure; plot(b,pdf_e,'ko-');**

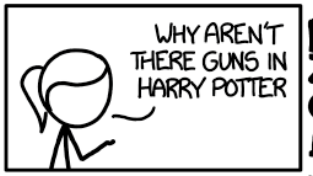
Credit: XKCD comics

QUESTIONS
FOUND IN GOOGLE AUTOCOMPLETE

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD
WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KLINGONS DIFFERENT
WHY ARE THERE SQUIRRELS

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL
WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY
WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD
WHY IS SEX SO IMPORTANT

WHY ARE THERE SLAVES IN THE BIBLE
WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT
WHY AREN'T MY ARMS GROWING

WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY
WHY IS YKK ON ALL ZIPPERS
WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS
WHY IS THERE LAVA
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE
WHY ARE THERE ANTS IN MY LAPTOP
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN
WHY ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47S SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS
WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA
WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE
WHY AREN'T THERE GUNS IN HARRY POTTER

WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG
WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND
WHY IS LIFE SO BORING
WHY ARE THERE FEMALE MR MIMES
WHY IS GPS FREE
WHY IS THERE HELL IF GOD FORGIVES
WHY DO IGUANAS DIE
WHY AREN'T THERE DINOSAUR GHOSTS

Matlab exercise: Uniform PDF

- **Stats=100000;**
- **r2=3+4.*rand(Stats,1);**
- **disp(mean(r2));**
- **disp(var(r2));**
- **disp(std(r2));**
- **step=0.1;**
- **[a,b]=hist(r2,0:step:8);**
- **pdf_e=a./sum(a)./step;**
- **figure; plot(b,pdf_e,'ko-');**