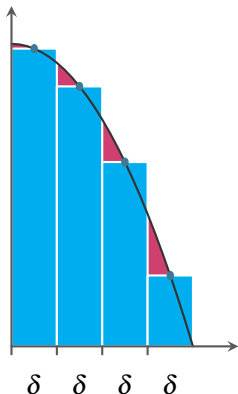


# Survey

Fill it out!!

[tinyurl.com/cs70-survey](https://tinyurl.com/cs70-survey)

# Calculus



Riemann Sum/Integral:  $\int_a^b f(x)dx = \lim_{\delta \rightarrow 0} \sum_i \delta f(a_i)$

“Area is defined as rectangles and add up some thin ones.”

Derivative (Rate of change):

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h}.$$

“Rise over run of close together points.”

Fundamental Theorem:  $F(b) - F(a) = \int_a^b F'(x)dx.$

“Area ( $F(\cdot)$ ) under  $f(x)$  grows at  $x$ ,  $F'(x)$ , by  $f(x)$ ”

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

# CS70: Continuous Probability.

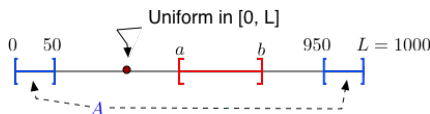
## Continuous Probability 1

1. Examples
2. Events
3. Continuous Random Variables

## Uniformly at Random in $[0, 1]$ .

Choose a real number  $X$ , uniformly at random in  $[0, 1]$ .

What is the probability that  $X$  is exactly equal to  $1/3$ ? Well, ..., 0.



What is the probability that  $X$  is exactly equal to 0.6? Again, 0.

In fact, for any  $x \in [0, 1]$ , one has  $\Pr[X = x] = 0$ .

How should we then describe 'choosing uniformly at random in  $[0, 1]$ '?

Here is the way to do it:

$$\Pr[X \in [a, b]] = b - a, \forall 0 \leq a \leq b \leq 1.$$

Makes sense:  $b - a$  is the fraction of  $[0, 1]$  that  $[a, b]$  covers.

$$F_X(x) = \Pr[X \leq x]$$

.

$$f_X(x) = \lim_{\delta \rightarrow 0} \Pr[X \in (x, x + \delta)]$$

What is true?

(A)  $F_X(x) = \int_{-\infty}^{\infty} f_X(y) dy$

(B)  $\int_{-\infty}^{\infty} f_X(x) = 1$

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

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

(E)  $\int_{-\infty}^{\infty} F_X(x) dx = 1.$

(F)  $\int_{-\infty}^{\infty} xf(x) dx = \int_{-\infty}^{\infty} (1 - F(x)) dx.$

(A) False. limits wrong. (B) cuz probability distribution.

(C) “sums up probability of rectangles”, e.g. calculus.

(D) calculus, fundamental theorem.

(F) is true since  $\int_{-\infty}^{\infty} xf(x) dx = \int_{-\infty}^{\infty} F(x) dx = E[X].$

Next lecture.

## Uniformly at Random in $[0, 1]$ .

Let  $[a, b]$  denote the **event** that the point  $X$  is in the interval  $[a, b]$ .

$$Pr[[a, b]] = \frac{\text{length of } [a, b]}{\text{length of } [0, 1]} = \frac{b - a}{1} = b - a.$$

Intervals like  $[a, b] \subseteq \Omega = [0, 1]$  are **events**.

More generally, events in this space are **unions of intervals**.

Example: the event  $A$  - “within 0.2 of 0 or 1” is  $A = [0, 0.2] \cup [0.8, 1]$ .

Thus,

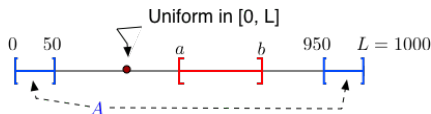
$$Pr[A] = Pr[[0, 0.2]] + Pr[[0.8, 1]] = 0.4.$$

More generally, if  $A_n$  are pairwise disjoint intervals in  $[0, 1]$ , then

$$Pr[\cup_n A_n] := \sum_n Pr[A_n].$$

Many subsets of  $[0, 1]$  are of this form. Thus, the probability of those sets is well defined. We call such sets **events**.

# Uniformly at Random in $[0, 1]$ .



Note: A **radical** change in approach.

**Finite prob. space:**  $\Omega = \{1, 2, \dots, N\}$ , with  $Pr[\omega] = p_\omega$ .

$\implies Pr[A] = \sum_{\omega \in A} p_\omega$  for  $A \subset \Omega$ .

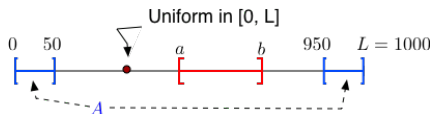
**Continuous space:** e.g.,  $\Omega = [0, 1]$ ,

*$Pr[\omega]$  is typically 0.*

Instead, start with  $Pr[A]$  for some events  $A$ .

Event  $A$  = interval, or union of intervals.

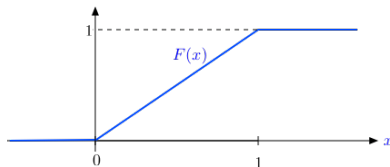
# Uniformly at Random in $[0, 1]$ .



$Pr[X \leq x] = x$  for  $x \in [0, 1]$ . Also,  $Pr[X \leq x] = 0$  for  $x < 0$ .

$Pr[X \leq x] = 1$  for  $x > 1$ .

Define  $F(x) = Pr[X \leq x]$ .

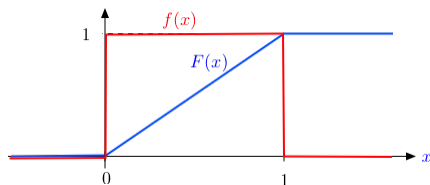


Then we have  $Pr[X \in (a, b)] = Pr[X \leq b] - Pr[X \leq a] = F(b) - F(a)$ .

Thus,  $F(\cdot)$  specifies the probability of all the events!



Uniformly at Random in  $[0, 1]$ .



$$Pr[X \in (a, b]] = Pr[X \leq b] - Pr[X \leq a] = F(b) - F(a).$$

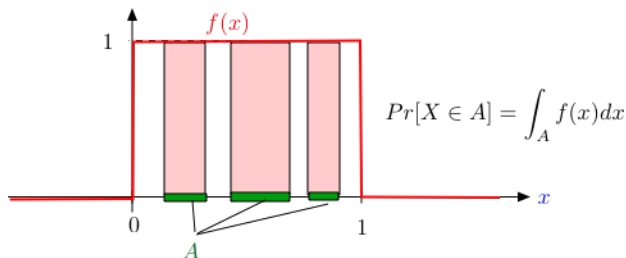
An alternative view is to define  $f(x) = \frac{d}{dx} F(x) = 1\{x \in [0, 1]\}$ . Then

$$F(b) - F(a) = \int_a^b f(x) dx.$$

Thus, the probability of an event is the integral of  $f(x)$  over the event:

$$Pr[X \in A] = \int_A f(x) dx.$$

## Uniformly at Random in $[0, 1]$ .



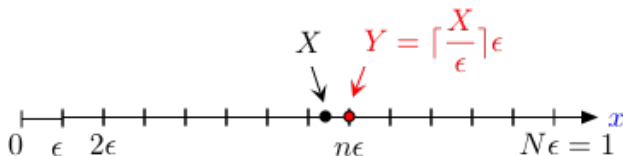
Think of  $f(x)$  as describing how  
one unit of probability is spread over  $[0, 1]$ : uniformly!

Then  $Pr[X \in A]$  is the probability mass over  $A$ .

Observe:

- ▶ This makes the probability automatically additive.
- ▶ We need  $f(x) \geq 0$  and  $\int_{-\infty}^{\infty} f(x) dx = 1$ .

Uniformly at Random in  $[0, 1]$ .



**Discrete Approximation:** Fix  $N \gg 1$  and let  $\epsilon = 1/N$ .

Define  $Y = n\epsilon$  if  $(n-1)\epsilon < X \leq n\epsilon$  for  $n = 1, \dots, N$ .

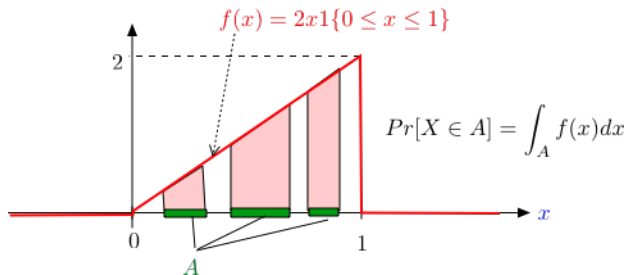
Then  $|X - Y| \leq \epsilon$  and  $Y$  is discrete:  $Y \in \{\epsilon, 2\epsilon, \dots, N\epsilon\}$ .

Also,  $\Pr[Y = n\epsilon] = \frac{1}{N}$  for  $n = 1, \dots, N$ .

Thus,  $X$  is ‘almost discrete.’

Calculus view:  $\Pr[Y = n\epsilon]$  is area of rectangle in Riemann sum.

## Nonuniformly at Random in $[0, 1]$ .



This figure shows a different choice of  $f(x) \geq 0$  with  $\int_{-\infty}^{\infty} f(x) dx = 1$ .

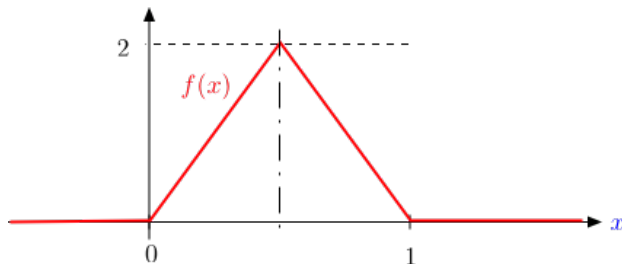
It defines another way of choosing  $X$  at random in  $[0, 1]$ .

Note that  $X$  is more likely to be closer to 1 than to 0.

One has  $Pr[X \leq x] = \int_{-\infty}^x f(u) du = x^2$  for  $x \in [0, 1]$ .

Also,  $Pr[X \in (x, x + \varepsilon)] = \int_x^{x+\varepsilon} f(u) du \approx f(x)\varepsilon$ .

## Another Nonuniform Choice at Random in $[0, 1]$ .



This figure shows yet a different choice of  $f(x) \geq 0$  with  $\int_{-\infty}^{\infty} f(x) dx = 1$ .

It defines another way of choosing  $X$  at random in  $[0, 1]$ .

Note that  $X$  is more likely to be closer to  $1/2$  than to 0 or 1.

For instance,  $Pr[X \in [0, 1/3]] = \int_0^{1/3} 4x dx = 2[x^2]_0^{1/3} = \frac{2}{9}$ .

Thus,  $Pr[X \in [0, 1/3]] = Pr[X \in [2/3, 1]] = \frac{2}{9}$  and  $Pr[X \in [1/3, 2/3]] = \frac{5}{9}$ .

## General Random Choice in $\mathfrak{R}$

Let  $F(x)$  be a nondecreasing function with  $F(-\infty) = 0$  and  $F(+\infty) = 1$ .

Define  $X$  by  $Pr[X \in (a, b]] = F(b) - F(a)$  for  $a < b$ . Also, for  $a_1 < b_1 < a_2 < b_2 < \dots < b_n$ ,

$$\begin{aligned} Pr[X \in (a_1, b_1] \cup (a_2, b_2] \cup (a_n, b_n]] \\ &= Pr[X \in (a_1, b_1]] + \dots + Pr[X \in (a_n, b_n]] \\ &= F(b_1) - F(a_1) + \dots + F(b_n) - F(a_n). \end{aligned}$$

Let  $f(x) = \frac{d}{dx} F(x)$ . Then,

$$Pr[X \in (x, x + \varepsilon]] = F(x + \varepsilon) - F(x) \approx f(x)\varepsilon.$$

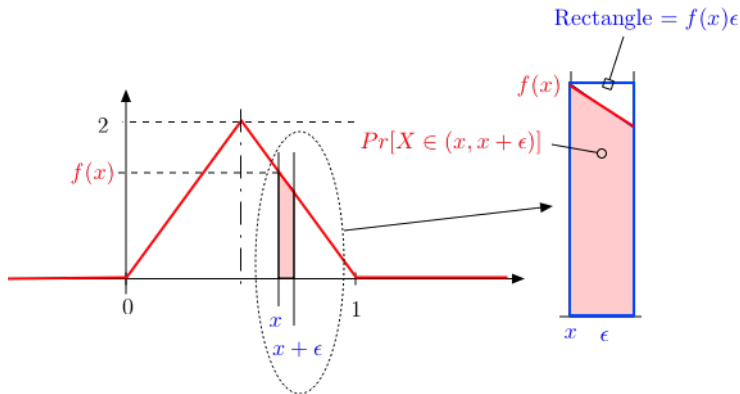
$F(x)$  is **cumulative distribution function (cdf)** of  $X$

$f(x)$  is the **probability density function (pdf)** of  $X$ .

When  $F$  and  $f$  correspond RV  $X$ :  $F_X(x)$  and  $f_X(x)$ .

$$Pr[X \in (x, x + \epsilon)]$$

An illustration of  $Pr[X \in (x, x + \epsilon)] \approx f_X(x)\epsilon$ :



Thus, the pdf is the 'local probability by unit length.'

It is the 'probability density.'

# Discrete Approximation

Fix  $\varepsilon \ll 1$  and let  $Y = n\varepsilon$  if  $X \in (n\varepsilon, (n+1)\varepsilon]$ .

Thus,  $Pr[Y = n\varepsilon] = F_X((n+1)\varepsilon) - F_X(n\varepsilon)$ .

Note that  $|X - Y| \leq \varepsilon$  and  $Y$  is a discrete random variable.

Also, if  $f_X(x) = \frac{d}{dx} F_X(x)$ , then  $F_X(x + \varepsilon) - F_X(x) \approx f_X(x)\varepsilon$ .

Hence,  $Pr[Y = n\varepsilon] \approx f_X(n\varepsilon)\varepsilon$ .

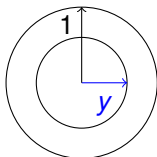
Thus, we can think of  $X$  of being almost discrete with

$Pr[X = n\varepsilon] \approx f_X(n\varepsilon)\varepsilon$ .



## Example: CDF, pre-poll

Example: hitting random location on gas tank.  
Random location on circle.

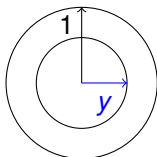


What is probability of being within  $y$  of the center, for non-negative  $y \leq 1$ ?

- (A) 1.
- (B) 0.
- (C)  $\int_0^y (2\pi y) dy$
- (D)  $y^2$ .
- (D) Next slide.

## Example: CDF

Example: hitting random location on gas tank.  
Random location on circle.



Random Variable:  $Y$  distance from center.  
Probability within  $y$  of center:

$$\begin{aligned}Pr[Y \leq y] &= \frac{\text{area of small circle}}{\text{area of dartboard}} \\&= \frac{\pi y^2}{\pi} = y^2.\end{aligned}$$

Hence,

$$F_Y(y) = Pr[Y \leq y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1 \end{cases}$$

## Calculation of event with dartboard..

Probability between .5 and .6 of center?

Recall CDF.

$$F_Y(y) = Pr[Y \leq y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1 \end{cases}$$

$$\begin{aligned} Pr[0.5 < Y \leq 0.6] &= Pr[Y \leq 0.6] - Pr[Y \leq 0.5] \\ &= F_Y(0.6) - F_Y(0.5) \\ &= .36 - .25 \\ &= .11 \end{aligned}$$

Example: “Dart” board.

Recall that

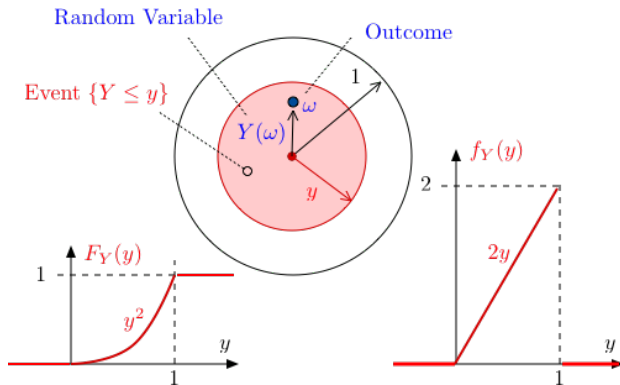
$$F_Y(y) = \Pr[Y \leq y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1 \end{cases}$$

$$f_Y(y) = F'_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ 2y & \text{for } 0 \leq y \leq 1 \\ 0 & \text{for } y > 1 \end{cases}$$

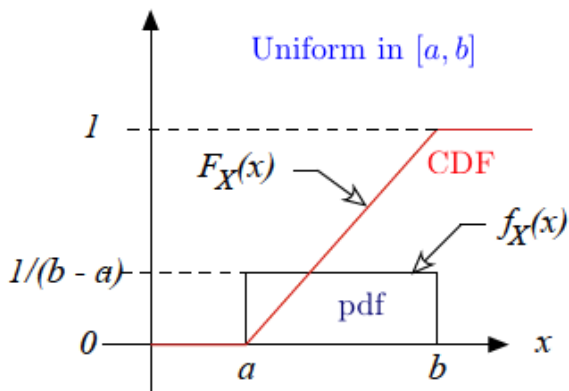
The cumulative distribution function (cdf) and probability distribution function (pdf) give full information.

Use whichever is convenient.

# Target



$U[a, b]$



## Exponential derivation: Poll.

$$\Pr[X = i] = (1 - p)^{i-1} p.$$

Let  $p = \lambda/n$ . and  $Y = X/n$ .

What is true?

(A)  $X \sim G(p)$

(B)  $\Pr[X > i] = (1 - p)^i$ . (C)  $\Pr[Y > i/n] = (1 - \lambda/n)^i$ .

(D)  $\Pr[Y > y] = (1 - \lambda/n)^{ny}$ .

(E)  $\lim_{n \rightarrow \infty} (1 - \lambda/n)^{ny} = e^{-\lambda y}$ .

(A) True by definition.

(B)  $\Pr[X > i] = (1 - p)^i$  at least  $i$  coin flips fail.

(C) True, definition of  $Y$

(D) True,  $y = i/n$  means  $i = ny$ .

(E)  $(1 - \lambda/n)^{ny} = ((1 - \lambda/n)^{n/\lambda})^{\lambda y}$   
and  $\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^{n/\lambda} = e$ .

The limit as  $n \rightarrow \infty$  of  $Y$  has  $\Pr[Y > y] = e^{-\lambda y}$ .

$\Pr[Y > y]$  is defined as “Survival function.”

## Expo( $\lambda$ )

“Limit of geometric.”

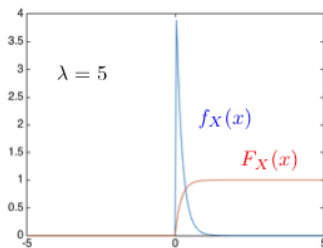
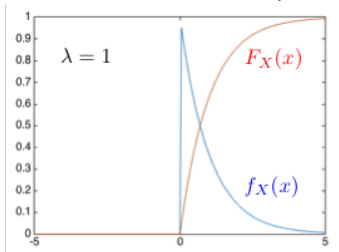
From last slide:  $S(t) = \Pr[X > t] = e^{-\lambda t}$  for  $t > 0$ .

Note:  $f_X(x) = F'(t) = 1 - S(t) = -S'(t)$ .

The exponential distribution with parameter  $\lambda > 0$  is defined by

$$f_X(x) = \lambda e^{-\lambda x} 1_{\{x \geq 0\}}$$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geq 0. \end{cases}$$





# Continuous Random Variables

Continuous random variable  $X$ , specified by

1.  $F_X(x) = \Pr[X \leq x]$  for all  $x$ .

**Cumulative Distribution Function (cdf).**

$$\Pr[a < X \leq b] = F_X(b) - F_X(a)$$

1.1  $0 \leq F_X(x) \leq 1$  for all  $x \in \mathfrak{R}$ .

1.2  $F_X(x) \leq F_X(y)$  if  $x \leq y$ .

2. Or  $f_X(x)$ , where  $F_X(x) = \int_{-\infty}^x f_X(u) du$  or  $f_X(x) = \frac{d(F_X(x))}{dx}$ .

**Probability Density Function (pdf).**

$$\Pr[a < X \leq b] = \int_a^b f_X(x) dx = F_X(b) - F_X(a)$$

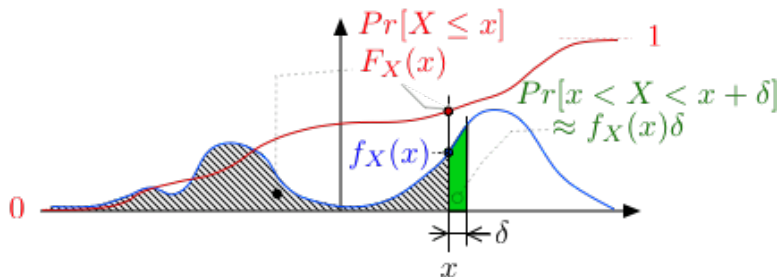
2.1  $f_X(x) \geq 0$  for all  $x \in \mathfrak{R}$ .

2.2  $\int_{-\infty}^{\infty} f_X(x) dx = 1$ .

Recall that  $\Pr[X \in (x, x + \delta)] \approx f_X(x)\delta$ .

$X$  “takes” value  $n\delta$ , for  $n \in \mathbb{Z}$ , with  $\Pr[X = n\delta] = f_X(n\delta)\delta$

## A Picture



The pdf  $f_X(x)$  is a nonnegative function that integrates to 1.

The cdf  $F_X(x)$  is the integral of  $f_X$ .

$$Pr[x < X < x + \delta] \approx f_X(x)\delta$$

$$Pr[X \leq x] = F_X(x) = \int_{-\infty}^x f_X(u) du$$

# Multiple Continuous Random Variables

One defines a pair  $(X, Y)$  of continuous RVs by specifying  $f_{X,Y}(x, y)$  for  $x, y \in \Re$  where

$$f_{X,Y}(x, y) dx dy = Pr[X \in (x, x + dx), Y \in (y, y + dy)].$$

The function  $f_{X,Y}(x, y)$  is called the **joint pdf** of  $X$  and  $Y$ .

**Example:** Choose a point  $(X, Y)$  uniformly in the set  $A \subset \Re^2$ . Then

$$f_{X,Y}(x, y) = \frac{1}{|A|} \mathbf{1}\{(x, y) \in A\}$$

where  $|A|$  is the area of  $A$ .

**Interpretation.** Think of  $(X, Y)$  as being discrete on a grid with mesh size  $\varepsilon$  and  $Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2$ .

Recall Marginal Distribution:

$$Pr[X = x] = \sum_y Pr[X = x, Y = y].$$

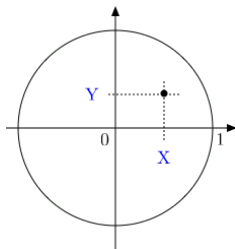
Similarly:

$$f_X(x) = \int f_{X,Y}(x, y) dy.$$

Sum “goes to” integral.

## Example of Continuous $(X, Y)$

Pick a point  $(X, Y)$  uniformly in the unit circle.



Thus,  $f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$ .

Consequently,

$$Pr[X > 0, Y > 0] = \frac{1}{4}$$

$$Pr[X < 0, Y > 0] = \frac{1}{4}$$

$$Pr[X^2 + Y^2 \leq r^2] = \frac{\pi r^2}{\pi} = r^2$$

$$Pr[X > Y] = \frac{1}{2}.$$

# Independent Continuous Random Variables

**Definition:** Continuous RVs  $X$  and  $Y$  independent if and only if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$

**Theorem:** Continuous RVs  $X$  and  $Y$  independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

Note:  $f_X(x)$  ( $f_Y(y)$ ) is (marginal) distribution of  $X$  ( $Y$ ).

**Proof:** Intervals:  $A = [x, x + dx]$ ,  $B = [y, y + dy]$ .

$$\begin{aligned} Pr[X \in A, Y \in B] &= Pr[X \in A] \times Pr[Y \in B] \\ &\approx f_X(x) dx \times f_Y(y) dy \\ &= f_X(x)f_Y(y) dx dy. \end{aligned}$$

Thus,  $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ .



# Mutual Independence.

**Definition:** Continuous RVs  $X_1, \dots, X_n$  are mutually independent if

$$Pr[X_1 \in A_1, \dots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \dots, A_n.$$

**Theorem:** Continuous RVs  $X_1, \dots, X_n$  are mutually independent if and only if

$$f_{\mathbf{X}}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

**Proof:** As in the discrete case.

## Conditional density.

Conditional Density:  $f_{X|Y}(x, y)$ .

Conditional Probability:  $Pr[X \in A | Y \in B] = \frac{Pr[X \in A, Y \in B]}{Pr[Y \in B]}$

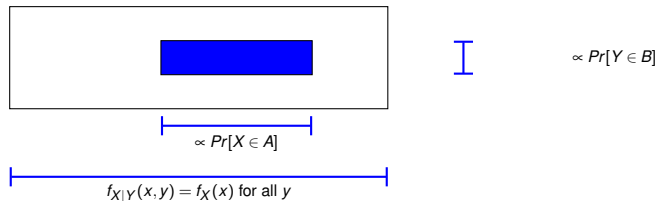
$$Pr[X \in [x, x + dx] | Y \in [y, y + dy]] = \frac{f_{X,Y}(x,y) dx dy}{f_Y dy}$$

$$f_{X|Y}(x, y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy}$$

Corollary: For independent random variables,  $f_{X|Y}(x, y) = f_X(x)$ .

# Independent Random Variables?

Uniform on a rectangle? Independent?



Also:  $Pr[X \in A, Y \in B] \propto \text{Area of rectangle} \propto Pr[X \in A] \times Pr[Y \in B]$ .

Independent!

Uniform on a circle? Independent?





# Summary

## Continuous Probability 1

1. **pdf:**  $Pr[X \in (x, x + \delta)] = f_X(x)\delta$ .
2. **CDF:**  $Pr[X \leq x] = F_X(x) = \int_{-\infty}^x f_X(y)dy$ .
3.  **$U[a, b]$ :**  $f_X(x) = \frac{1}{b-a}1\{a \leq x \leq b\}$ ;  $F_X(x) = \frac{x-a}{b-a}$  for  $a \leq x \leq b$ .
4.  **$Expo(\lambda)$ :**  
 $f_X(x) = \lambda \exp\{-\lambda x\}1\{x \geq 0\}$ ;  $F_X(x) = 1 - \exp\{-\lambda x\}$  for  $x \geq 0$ .
5. **Target:**  $f_X(x) = 2x1\{0 \leq x \leq 1\}$ ;  $F_X(x) = x^2$  for  $0 \leq x \leq 1$ .
6. **Joint pdf:**  $Pr[X \in (x, x + \delta), Y \in (y, y + \delta)] = f_{X,Y}(x, y)\delta^2$ .
  - 6.1 Conditional Distribution:  $f_{X|Y}(x, y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ .
  - 6.2 Independence:  $f_{X|Y}(x, y) = f_X(x)$

# Summary

## Continuous Probability

- ▶ Continuous RVs are similar to discrete RVs (break into intervals.)
- ▶ Think that  $X \approx x$  with probability  $f_X(x)\varepsilon$
- ▶ Sums become integrals, ....