Introduction to Probability and Statistics
(EE 354 / CE 361 / Math 310)

Dr. Umer Tariq
Assistant Professor,
Dhanani School of Science & Engineering,
Habib University

Outline: Unit 5

- Continuous Random Variables
- Probability Density Function (PDF) ⟷ PMF

  1) Uniform 2) Exponential
- Cumulative Distribution Function (CDF)
- 3) Normal (Gaussian) Random Variable
- Joint PDF of Multiple Continuous Random Variables
- Conditioning a Continuous Random Variable
- Independence of Continuous Random Variables

Random Variables: The Formalism

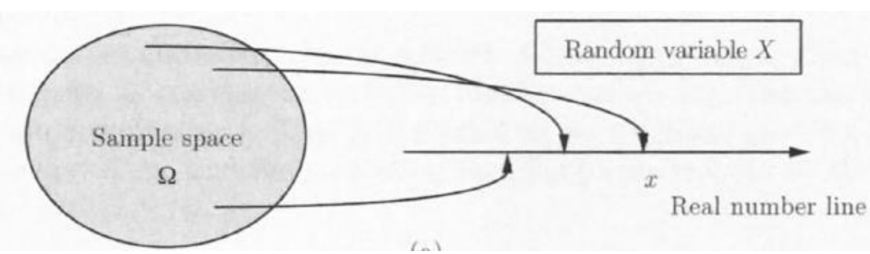
- Random Variable
  - Mathematically, random variable is a function from the sample space to real numbers
  - $X\colon \Omega \to \mathbb{R}$

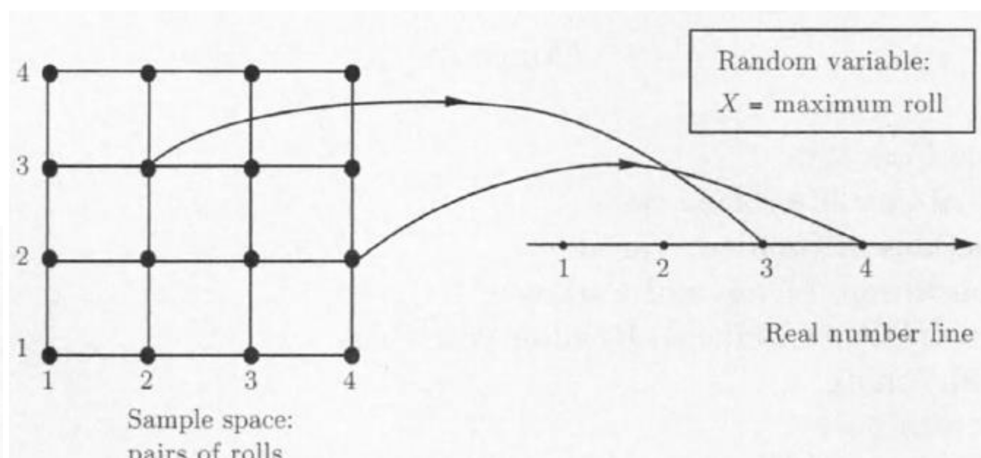
Notation: random variable X numerical value x

- Discrete Random Variables

  A random variable is called **discrete** if its range (the set of values that it can take) is either finite or countably infinite.

- Continuous Random Variables
  - A random variable is called **continuous** if its range (the set of values that it can take) is uncountably infinite.





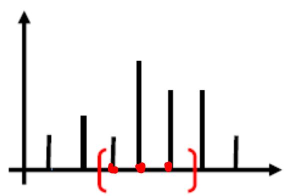
Continuous Random Variables: Examples

- Consider the velocity of a vehicle traveling along the highway
  – If the velocity is measured by a digital speedometer, we may view the speedometer's reading as a discrete random variable
  – If we wish to "model" the exact velocity, a continuous random variable is required
- Other Examples
  – Weight of a randomly selected student in the class
  – Height of a randomly selected student in the class

Continuous RVs and Probability Density Functions (PDFs)

PMF of a Discrete RV

$p_X(x)$



$P(2.5 \le X \le 5.5) = p_X(3) + p_X(4) + p_X(5)$

$$P(a \le X \le b) = \sum_{x:a\le x\le b} p_X(x)$$

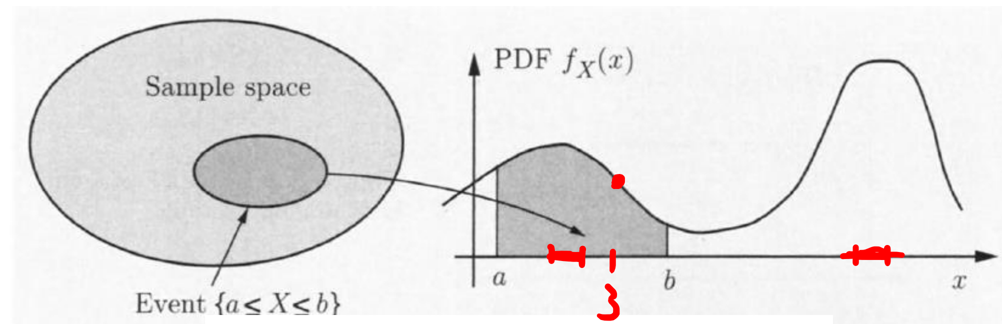
- Properties:

$p_X(x) \ge 0$

$\sum_x p_X(x) = 1$

$P(X=3) = p_X(3)$

PDF of a Continuous RV



$$P(a \le X \le b) = \int_a^b f_X(x)\,dx$$

$P(-\infty \le X \le \infty) = \int_{-\infty}^{\infty} f_X(x)\,dx = 1$

- Properties:

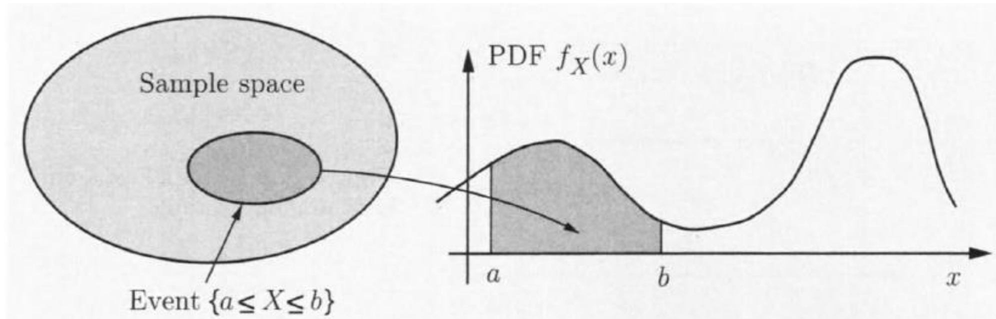
$\boxed{f_X(x) \ge 0}$

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

$P(X=3) \overset{?}{=} f_X(3)$

$P(X=3) = 0$

Continuous RVs and Probability Density Functions (PDFs)

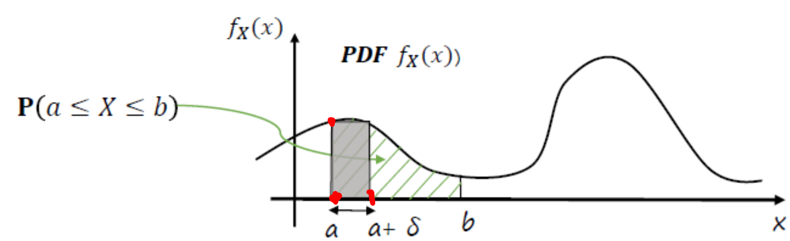


$$\mathbf{P}(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

- Properties:

$$f_X(x) \geq 0 \qquad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

Note:



$$\boxed{\mathbf{P}(\mathrm{a} \leq \mathrm{X} \leq a+\delta) \;\approx\; f_X(a)\delta}$$

$$\boxed{\mathbf{P}(X = a) = 0}$$

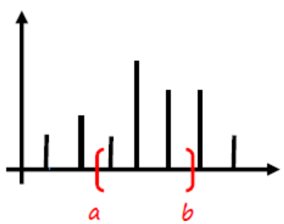
$$\frac{P(a \leq X \leq a+\delta)}{\delta} \simeq f_x(a) \quad \Leftarrow \quad P(a \leq X \leq a+\delta) \simeq f_x(a)\delta$$

$f_x(a)$

δ

Continuous RVs and Probability Density Functions (PDFs)

PMF of a Discrete RV



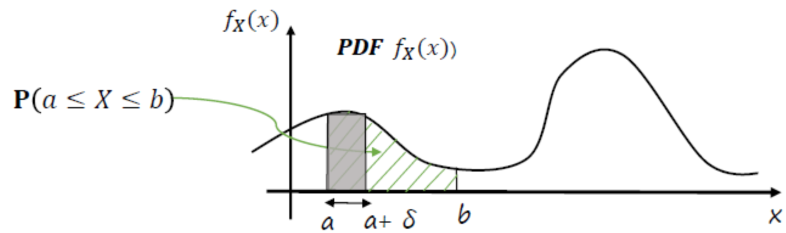
$$\mathbf{P}(\mathrm{a} \le \mathrm{X} \le \mathrm{b}) = \sum_{x: a \le x \le b} p_X(x)$$

- Properties:

$p_X(x) \ge 0$

$\sum_x p_X(x) = 1$

PDF of a Continuous RV



$$\mathbf{P}(a \le X \le b) = \int_a^b f_X(x)\, dx$$

- Properties:

$\boxed{f_X(x) \ge 0}$

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

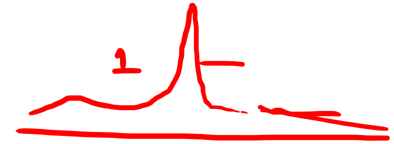
Note:

$\underbrace{\mathbf{P}(\mathrm{a} \le \mathrm{X} \le a + \delta)}_{\delta} \approx f_X(a)\cancel{\delta}$

$\mathbf{P}(X = a) = 0$

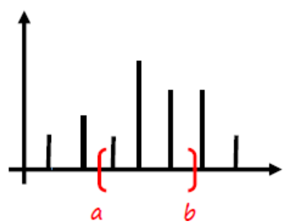
Interpretation of PDF:

- Unlike PMF, the PDF value at a point a does not represent the probability of that point.
- Rather, it represents the "probability per unit length" for a very small interval around a.



Continuous RVs and Probability Density Functions (PDFs)

PMF of a Discrete RV



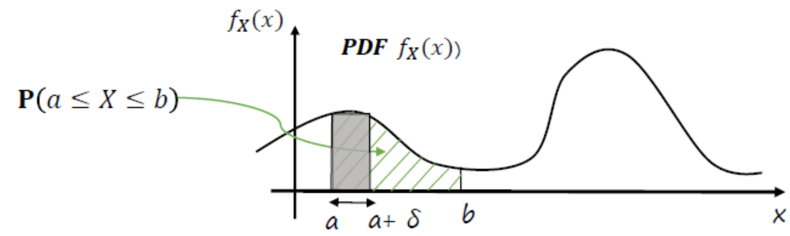
$$\mathbf{P}(a \le X \le b) = \sum_{x: a \le x \le b} p_X(x)$$

- Properties:

$p_X(x) \ge 0$

$\sum_x p_X(x) = 1$

PDF of a Continuous RV



$$\mathbf{P}(a \le X \le b) = \int_a^b f_X(x)\, dx$$

- Properties:

$\boxed{f_X(x) \ge 0}$

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

Note:

$\mathbf{P}(a \le X \le a + \delta) \approx f_X(a)\delta$

$\mathbf{P}(X = a) = 0$

Interpretation of PDF:

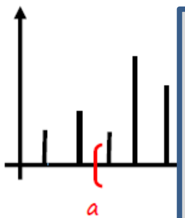
- Unlike PMF, the PDF value at a point a does not represent the probability of that point.
- Rather, it represents the probability per unit length for a very small interval around a.

Consequence

- Unlike PMF, the value of PDF at a point does not have to be less than one. as long as the properties of PDF are satisfied.

Continuous RVs and Probability Density Functions (PDFs)

PMF of a Discrete RV



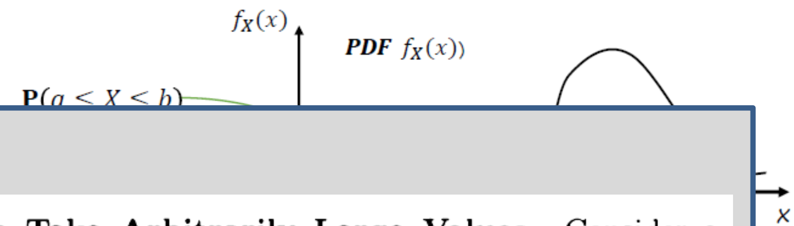
$$\mathbf{P}(a \le X \le b)$$

- Properties:

$$p_X(x) \ge 0$$

x

PDF of a Continuous RV



$)dx = 1$

Example 3.3. A PDF Can Take Arbitrarily Large Values. Consider a random variable X with PDF

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & \text{if } 0 < x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Even though $f_X(x)$ becomes infinitely large as x approaches zero, this is still a valid PDF, because

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 \frac{1}{2\sqrt{x}}\,dx = \sqrt{x}\Big|_0^1 = 1.$$

of PDF:

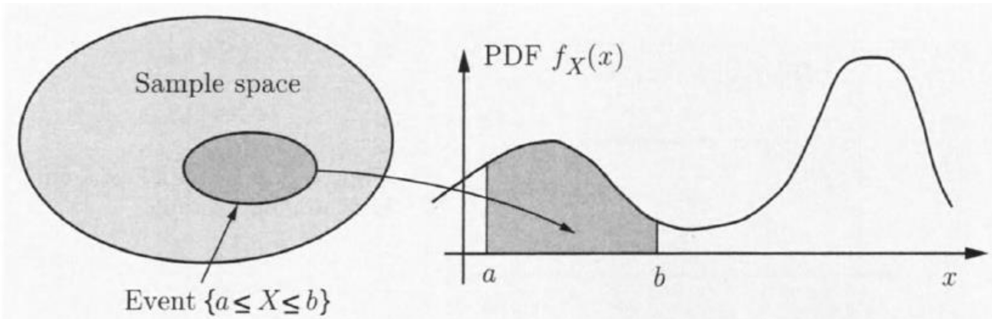
, the PDF value at
oes not represent
ility of that point.

- Rather, it represents the probability per unit length for a very small interval around a.

Consequence

- **Unlike PMF, the value of PDF at a point does not have to be less than one. as long as the properties of PDF are satisfied.**

Continuous RVs and Probability Density Functions (PDFs)



$P(a<X<b) = ?$

$$\mathbf{P}(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

- Properties:

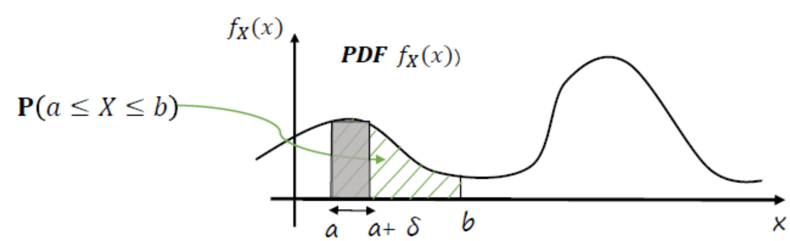
$$f_X(x) \geq 0 \qquad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

Note:

$$\mathbf{P}(a \leq X \leq a+\delta) \approx f_X(a)\delta$$

$$\mathbf{P}(X = a) = 0$$

$$\mathbf{P}(a \leq X \leq b) = \mathbf{P}(a < X < b)$$



$\Rightarrow P(a \leq X \leq b) = \cancel{P(X = a)} + \cancel{P(X = b)} + P(a < X < b)$

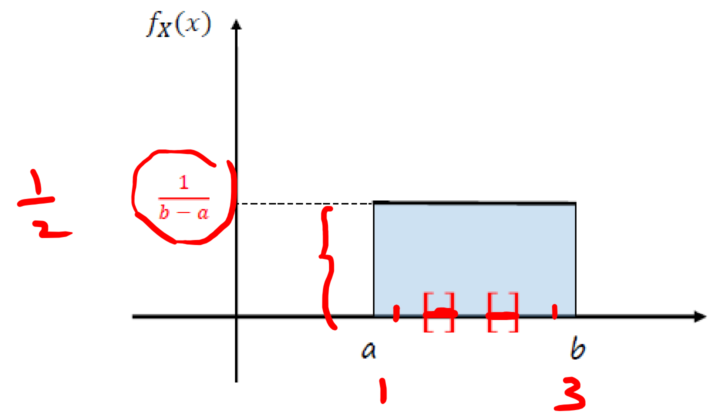
DRV — Bernoulli, Binomial, Geometric, Uniform, Poisson

CRV — Uniform, Exponential, Normal / Gaussian

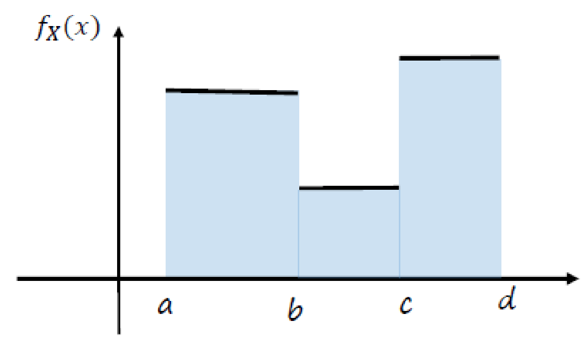
1) PDF Example: Continuous Uniform RV

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & \text{if } a \le x \le b, \\ 0, & \text{otherwise,} \end{cases}$$

$$1 = \int_{-\infty}^{\infty} f_X(x)\, dx = \int_a^b \frac{1}{b-a}\, dx.$$



- Generalization: piecewise constant PDF



PDF Example: Continuous Uniform RV

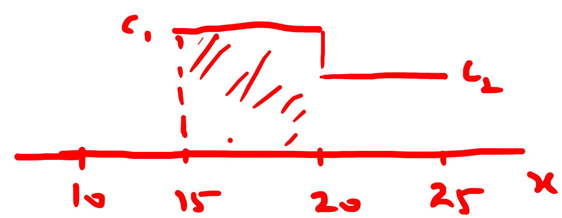
Example 3.2. Piecewise Constant PDF. Alvin's driving time to work is between 15 and 20 minutes if the day is sunny, and between 20 and 25 minutes if the day is rainy, with all times being equally likely in each case. Assume that a day is sunny with probability 2/3 and rainy with probability 1/3. What is the PDF of the driving time, viewed as a random variable X?

$P(\text{Sunny Day}) = \frac{2}{3}$ $P(\text{Rainy Day}) = \frac{1}{3}$

PDF of X

$f_X(x)$

$P(15 \le X \le 20) = \frac{2}{3}$ $P(20 \le X \le 25) = \frac{1}{3}$



$5c_1 = \frac{2}{3}$ $5c_2 = \frac{1}{3}$

$\boxed{c_1 = \frac{2}{15}}$ $\boxed{c_2 = 1/15}$

PDF Example: Continuous Uniform RV

Example 3.2. Piecewise Constant PDF. Alvin's driving time to work is between 15 and 20 minutes if the day is sunny, and between 20 and 25 minutes if the day is rainy, with all times being equally likely in each case. Assume that a day is sunny with probability 2/3 and rainy with probability 1/3. What is the PDF of the driving time, viewed as a random variable X?

We interpret the statement that "all times are equally likely" in the sunny and the rainy cases, to mean that the PDF of X is constant in each of the intervals $[15, 20]$ and $[20, 25]$. Furthermore, since these two intervals contain all possible driving times, the PDF should be zero everywhere else:

$$f_X(x) = \begin{cases} c_1, & \text{if } 15 \le x < 20, \\ c_2, & \text{if } 20 \le x \le 25, \\ 0, & \text{otherwise}, \end{cases}$$

where c_1 and c_2 are some constants. We can determine these constants by using the given probabilities of a sunny and of a rainy day:

$$\frac{2}{3} = \mathbf{P}(\text{sunny day}) = \int_{15}^{20} f_X(x)\,dx = \int_{15}^{20} c_1\,dx = 5c_1,$$

$$\frac{1}{3} = \mathbf{P}(\text{rainy day}) = \int_{20}^{25} f_X(x)\,dx = \int_{20}^{25} c_2\,dx = 5c_2,$$

so that

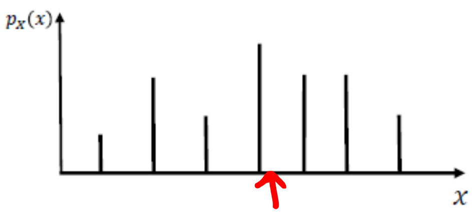
$$c_1 = \frac{2}{15}, \qquad c_2 = \frac{1}{15}.$$

PMF — Expectation/Mean, Variance, S.D.

PDF — Mean, Variance, S.D.

Expectation and Its Properties

Discrete RV



$$\mathbf{E}[X] = \sum_x x p_X(x).$$

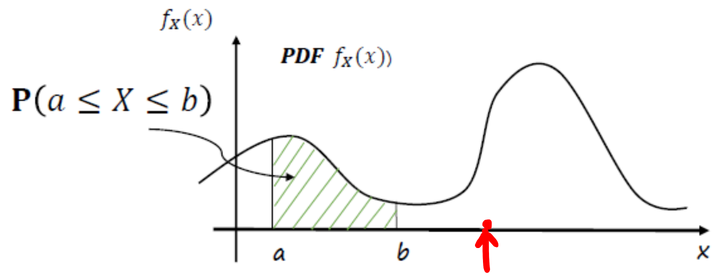
- Properties:
  - If $X \geq 0$, then $\mathbf{E}[\mathrm{X}] \geq 0$
  - If $a \leq X \leq b$, then $a \leq \mathbf{E}[\mathrm{X}] \leq b$
- Expected value rule:

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x).$$

- Linearity

$$\boldsymbol{E}[aX + b] = a\boldsymbol{E}[X] + b$$

Continuous RV



$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx.$$

Assume, $\int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty$

- Properties:
  - If $X \geq 0$, then $\mathbf{E}[\mathrm{X}] \geq 0$
  - If $a \leq X \leq b$, then $a \leq \mathbf{E}[\mathrm{X}] \leq b$
- Expected value rule:

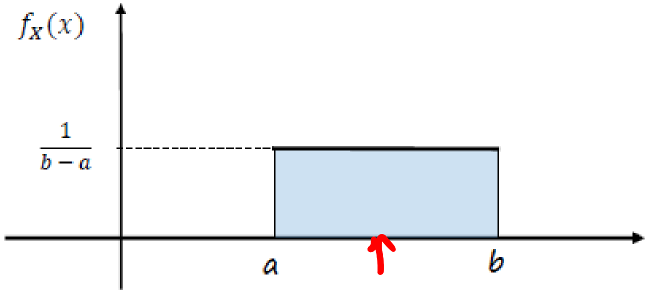
$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

- Linearity

$$\boldsymbol{E}[aX + b] = a\boldsymbol{E}[X] + b$$

Expectation: Interpretation

- The following interpretations of expectation/mean still remain valid in the case of continuous RV
  1. Center of gravity of the PDF
  2. Average of X in large number of repetitions of the experiment



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

$$= \int_a^b x \frac{1}{b-a} dx = \boxed{\frac{a+b}{2}}$$

Variance and Its Properties

Discrete RV

$$\boxed{\mathrm{var}(X) = \mathbf{E}\left[\left(X - \mathbf{E}[X]\right)^2\right],}$$

- Calculation using the expected value rule:

$$\boxed{\mathrm{var}(X) = \sum_x \left(x - \mathbf{E}[X]\right)^2 p_X(x).}$$

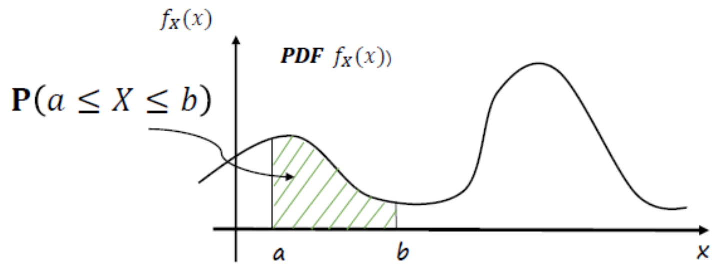
- Properties:

$$\mathrm{var}(X) \geq 0$$

$$\mathrm{var}(X) = \mathbf{E}[X^2] - \left(\mathbf{E}[X]\right)^2.$$

$$var(aX + b) = a^2\, var(X)$$

Continuous RV



$$\boxed{\mathrm{var}(X) = \mathbf{E}\left[\left(X - \mathbf{E}[X]\right)^2\right],}$$

- Calculation using the expected value rule:

$$\boxed{\mathrm{var}(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mathbf{E}[X])^2 f_X(x)\, dx}$$

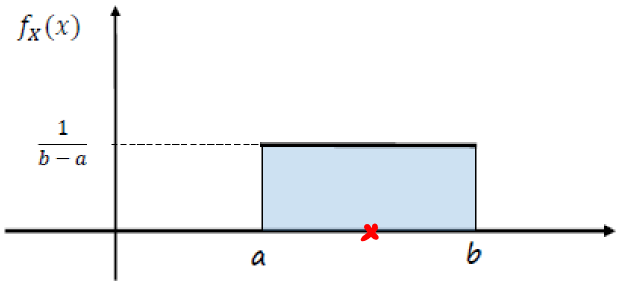
- Properties:

$$\mathrm{var}(X) \geq 0$$

$$\mathrm{var}(X) = \mathbf{E}[X^2] - \left(\mathbf{E}[X]\right)^2.$$

$$var(aX + b) = a^2\, var(X)$$

Uniform Random Variable



- Parameters
  - a, b

- Expectation

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\ dx$$
$$= \int_a^b x \frac{1}{b-a}\, dx = \boxed{\frac{a+b}{2}}$$

- Variance

$$E[X^2] = \int_a^b x^2 \frac{1}{b-a}\, dx = \frac{1}{b-a}\left(\frac{b^3}{3} - \frac{a^3}{3}\right)$$
$$\text{var}(X) = E[X^2] - (E[X])^2 = \boxed{\frac{(b-a)^2}{12}}$$

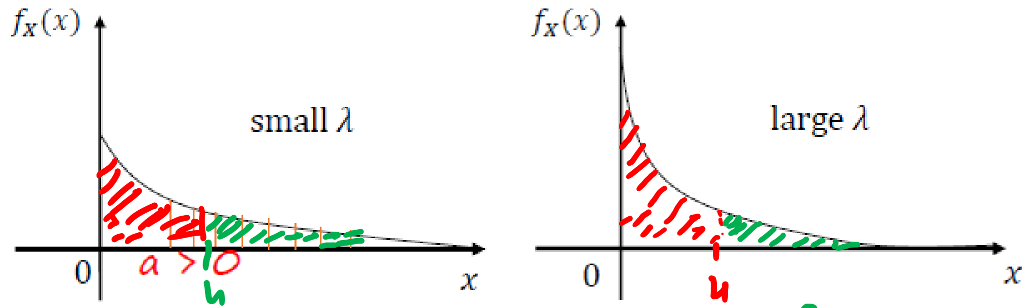
- Standard Deviation

$$\sigma = \frac{b-a}{\sqrt{12}}$$

$P(X \leq 4)$ $P(X > 4)$

2) Exponential Random Variable

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



- Parameters
  - λ

$$\boxed{P(X \geq a)} = \int_a^\infty \underbrace{\lambda e^{-\lambda x}}_{f_X(x)} dx$$

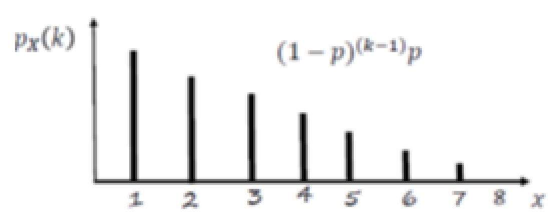
$$= \lambda \left(\frac{-1}{\lambda}\right) e^{-\lambda x} \Big|_a^\infty$$

$$= -e^{-\lambda \infty} + e^{-\lambda a} = \boxed{e^{-\lambda a}}$$

$$P(X \geq 4) = e^{-\lambda 4}$$

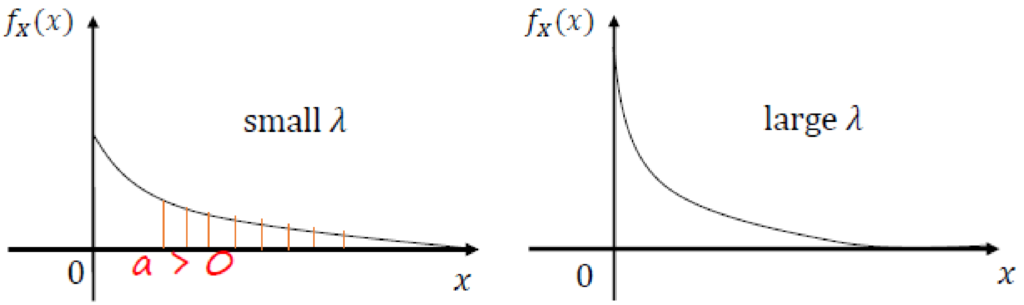
- Model of:
  - "Amount of time until an incident of interest takes place"
  - Example: Amount of time till a message arrives at an idle computer

- Discrete Counterpart
  - Geometric random variable



Exponential Random Variable

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



- Parameters
  - λ

- Expectation

$$\boldsymbol{E}[X] = \int_0^\infty x\,\lambda\, e^{-\lambda x}\, dx = \boxed{\frac{1}{\lambda}}$$

- Variance

$$\boldsymbol{E}[X^2] = \int_0^\infty x^2\,\lambda\, e^{-\lambda x}\, dx = \frac{2}{\lambda^2}$$

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

Example: Exponential RV

Example 3.5. The time until a small meteorite first lands anywhere in the Sahara desert is modeled as an exponential random variable with a mean of 10 days. The time is currently midnight. Q) What is the probability that a meteorite first lands some time between 6 a.m. and 6 p.m. of the first day?

$f_X(x) = \lambda e^{-\lambda x}$

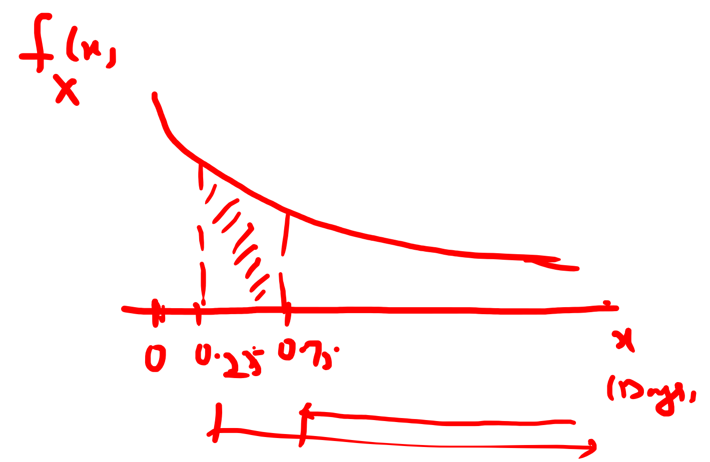
$\lambda = ?$

Mean = 10 days

$\frac{1}{\lambda} = 10 \Rightarrow \lambda = \frac{1}{10}$

$P(0.25 \le X \le 0.75) = ?$

$= \int_{0.25}^{0.75} f_X(x)\, dx$



$P(X \ge 0.25) - P(X \ge 0.75)$

$e^{-\lambda 0.25} - e^{-\lambda 0.75}$

Example: Exponential RV

Example 3.5. The time until a small meteorite first lands anywhere in the Sahara desert is modeled as an exponential random variable with a mean of 10 days. The time is currently midnight. What is the probability that a meteorite first lands some time between 6 a.m. and 6 p.m. of the first day?

Let X be the time elapsed until the event of interest, measured in days. Then, X is exponential, with mean $1/\lambda = 10$, which yields $\lambda = 1/10$. The desired probability is

$$\mathbf{P}(1/4 \le X \le 3/4) = \mathbf{P}(X \ge 1/4) - \mathbf{P}(X > 3/4) = e^{-1/40} - e^{-3/40} = 0.0476.$$

where we have used the formula $\mathbf{P}(X \ge a) = \mathbf{P}(X > a) = e^{-\lambda a}$.

PMF, PDF, CDF

Discrete Continuous

$\underbrace{f_X(x) \quad f_X(x)}_{F_X(x)}$

Cumulative Distribution Functions (CDFs)

- We have been dealing with discrete and continuous random variables in a somewhat different manner, using PMF and PDF, respectively.

- It would be desirable to describe all kinds of random variables with a single mathematical concept.
  - ***Cumulative Distribution Functions (CDFs)***

Cumulative Distribution Functions (CDFs)

- We have been dealing with discrete and continuous random variables in a somewhat different manner, using PMF and PDF, respectively.

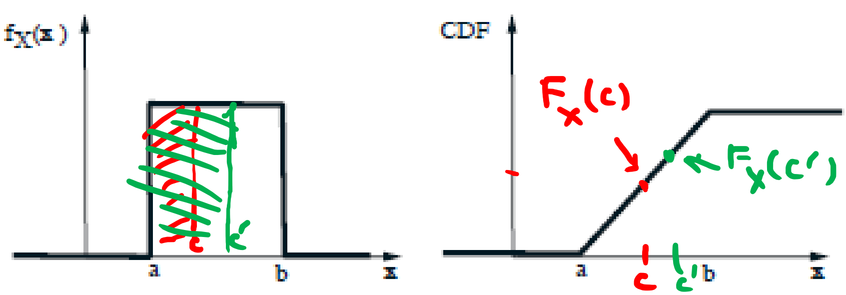
- It would be desirable to describe all kinds of random variables with a single mathematical concept.
  - ***<u>Cumulative Distribution Functions (CDFs)</u>***

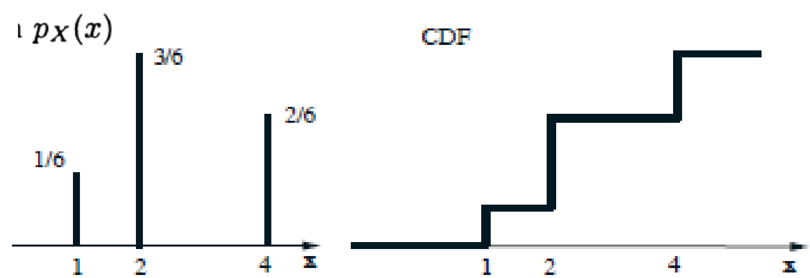
The CDF of a random variable X is denoted by F_X and provides the probability $\mathbf{P}(X \leq x)$. In particular, for every x we have

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

$F_X(c') = P(X \leq c')$

$$\boxed{F_X(x) = \mathbf{P}(X \leq x)} = \begin{cases} \displaystyle\sum_{k \leq x} p_X(k), & \text{if } X \text{ is discrete, } \checkmark \\ \displaystyle\int_{-\infty}^{x} f_X(t)\,dt, & \text{if } X \text{ is continuous. } \checkmark \end{cases}$$





Cumulative Distribution Functions (CDFs)

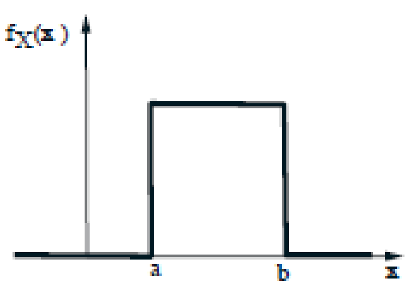
- We have been dealing with discrete and continuous random variables in a somewhat different manner, using PMF and PDF, respectively.

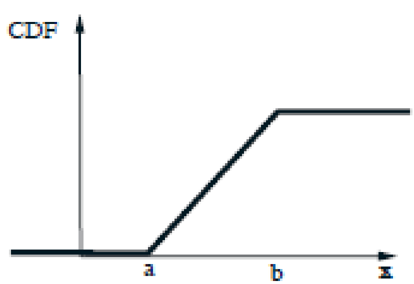
- Loosely speaking, the CDF $F_X(x)$ "accumulates" probability "up to" the value x.

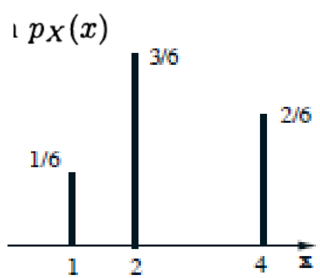
  – ***Cumulative Distribution Functions (CDFs)***

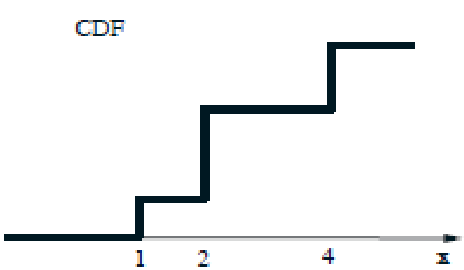
The CDF of a random variable X is denoted by F_X and provides the probability $\mathbf{P}(X \leq x)$. In particular, for every x we have

$$F_X(x) = \mathbf{P}(X \leq x) = \begin{cases} \displaystyle\sum_{k \leq x} p_X(k), & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{x} f_X(t)\,dt, & \text{if } X \text{ is continuous.} \end{cases}$$



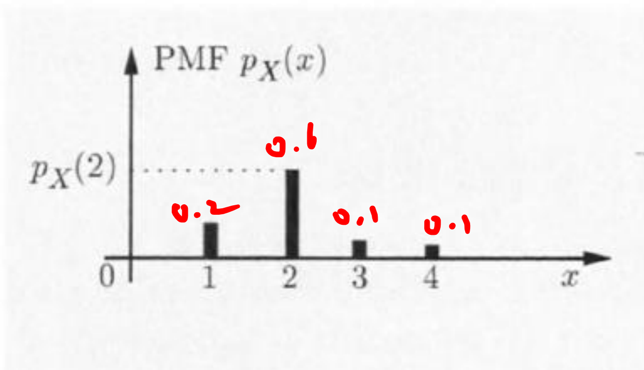






$F_X(c) = P(X \le c)$

CDFs: Discrete RVs Examples



CDF = ?

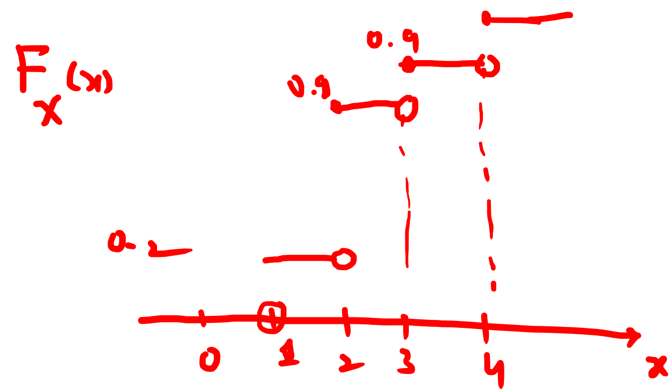
$F_X(x) = ?$

$F_X(0) = P(X \le 0) = 0$

$F_X(0.99) = P(X \le 0.99) = 0$

$F_X(1) = P(X \le 1) = 0.2$

$F_X(2) = P(X \le 2) = 0.8$



CDFs: Discrete RVs Examples

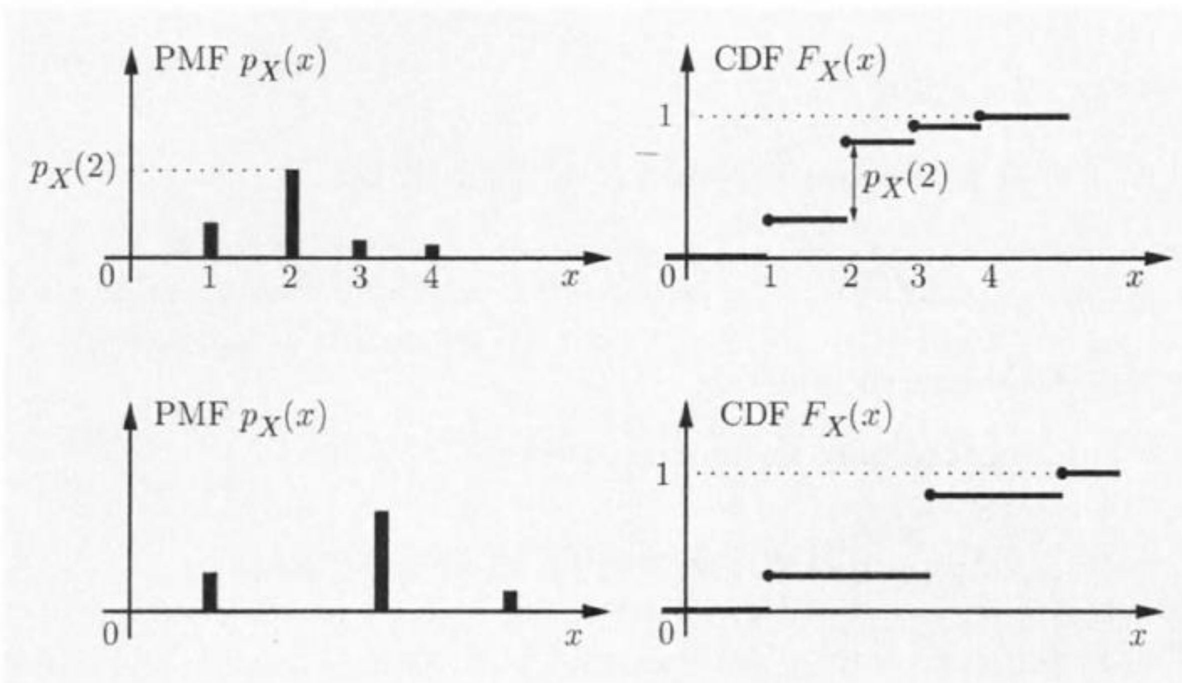


Figure 3.6: CDFs of some discrete random variables. The CDF is related to the PMF through the formula

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

and has a staircase form, with jumps occurring at the values of positive probability mass. Note that at the points where a jump occurs, the value of F_X is the larger of the two corresponding values (i.e., F_X is continuous from the right).

CDFs: Continuous RVs Examples

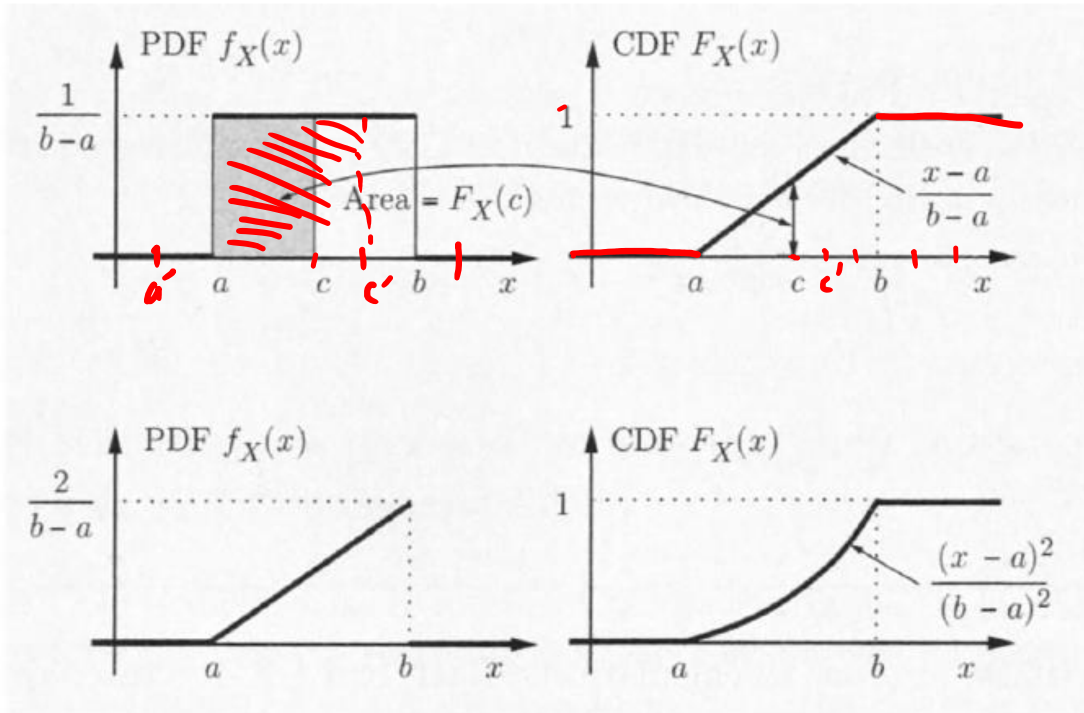
$F_X(c) = P(X \le c)$

$= \int_{-\infty}^{c} f_X(x)\, dx$

$= \int_{-\infty}^{c} \frac{1}{b-a} \cdot dx$

$= \frac{c-a}{b-a}$

$\int_{-\infty}^{c} (mx+c) dx$



$y = mx + c$

Figure 3.7: CDFs of some continuous random variables. The CDF is related to the PDF through the formula

$$F_X(x) = \mathbf{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

Thus, the PDF f_X can be obtained from the CDF by differentiation:

$$f_X(x) = \frac{dF_X}{dx}(x).$$

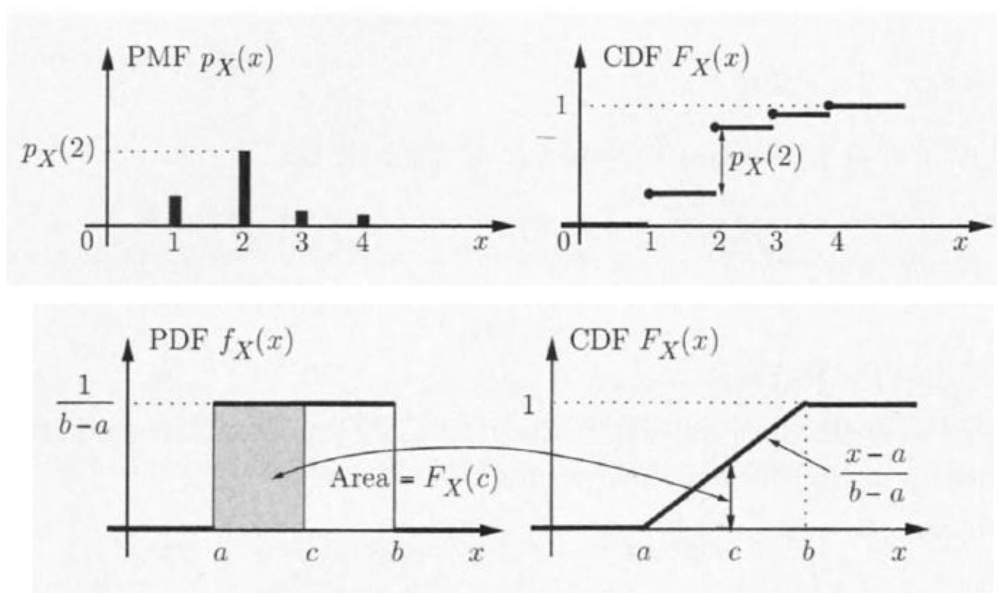
For a continuous random variable, the CDF has no jumps, i.e., it is continuous.

Properties of CDFs

- F_X is monotonically nondecreasing:

$$\text{if } x \leq y, \text{ then } F_X(x) \leq F_X(y).$$

- $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$.
- If X is discrete, then $F_X(x)$ is a piecewise constant function of x.
- If X is continuous, then $F_X(x)$ is a continuous function of x.



Properties of CDFs

- If X is discrete and takes integer values, the PMF and the CDF can be obtained from each other by summing or differencing:

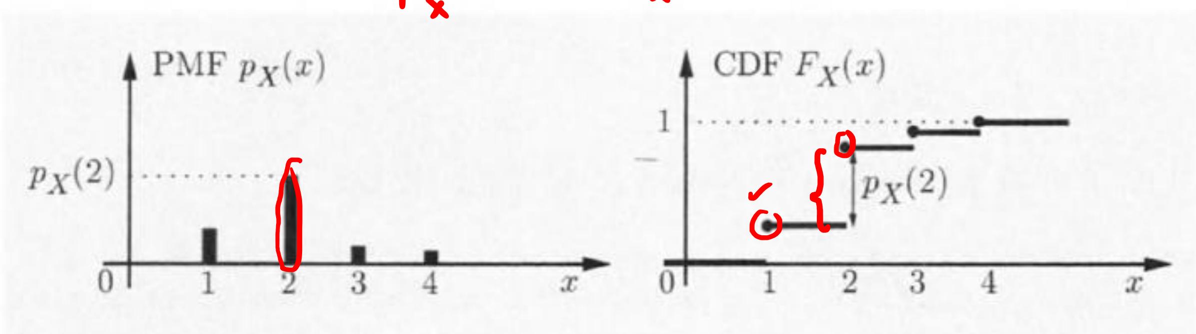
$$F_X(k) = \sum_{i=-\infty}^{k} p_X(i),$$

$$\boxed{p_X(k)} = \mathbf{P}(X \le k) - \mathbf{P}(X \le k-1) = \boxed{F_X(k) - F_X(k-1),}$$

for all integers k.

$p_X(2) = F_X(2) - F_X(1)$

$p_X(3) = F_X(3) - F_X(2)$

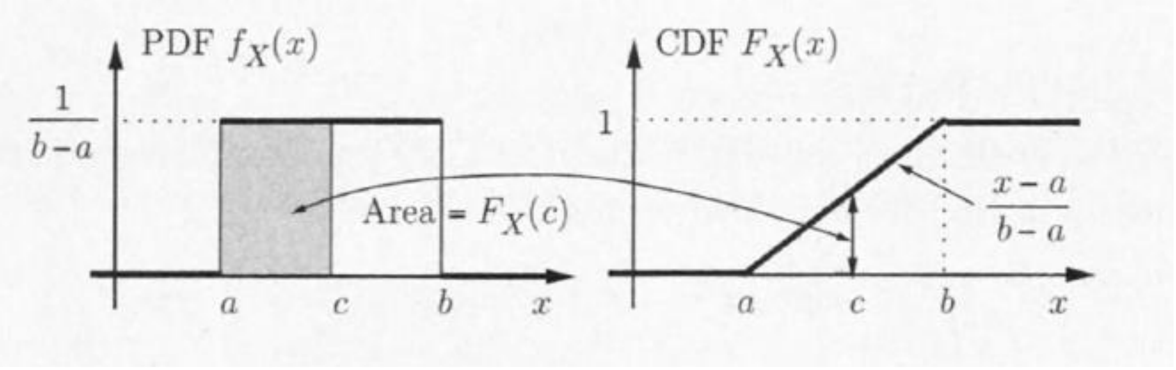


Properties of CDFs

- If X is continuous, the PDF and the CDF can be obtained from each other by integration or differentiation:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \qquad \boxed{f_X(x) = \frac{dF_X}{dx}(x).}$$

(The second equality is valid for those x at which the PDF is continuous.)



$\frac{4}{10}$ $P(A \cap B \cap C) = P(A)P(B)P(C)$

PMF of X_1:-

Example: PDF through CDF

Example 3.6. The Maximum of Several Random Variables. You are allowed to take a certain test three times, and your final score will be the maximum of the test scores. Thus,

$$X = \max\{X_1, X_2, X_3\},$$

where X_1, X_2, X_3 are the three test scores and X is the final score. Assume that your score in each test takes one of the values from 1 to 10 with equal probability 1/10, independently of the scores in other tests. What is the PMF p_X of the final score?

$F_X(4) = P(X \le 4)$

$= P(X_1 \le 4, X_2 \le 4, X_3 \le 4)$

$\frac{4}{10} \cdot \frac{4}{10} \cdot \frac{4}{10}$

1) CALCULATE CDF of X:-

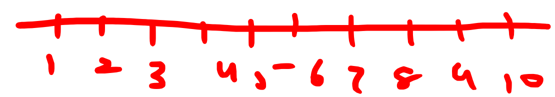
$F_X(k) = P(X \le k)$

$= P(X_1 \le k, X_2 \le k, X_3 \le k)$

$= P(X_1 \le k)\, P(X_2 \le k) \cdot P(X_3 \le k)$

$F_X(k) = \frac{k}{10} \cdot \frac{k}{10} \cdot \frac{k}{10} = \left(\frac{k}{10}\right)^3$

PMF of X:-



2) CDF of X $\longrightarrow$ PMF of X

$p_X(k) = F_X(k) - F_X(k-1)$

$= \left(\frac{k}{10}\right)^3 - \left(\frac{k-1}{10}\right)^3$

Example: PDF through CDF

Example 3.6. The Maximum of Several Random Variables. You are allowed to take a certain test three times. and your final score will be the maximum of the test scores. Thus,

$$X = \max\{X_1, X_2, X_3\},$$

where $X_1, X_2. X_3$ are the three test scores and X is the final score. Assume that your score in each test takes one of the values from 1 to 10 with equal probability 1/10, independently of the scores in other tests. What is the PMF p_X of the final score?

We calculate the PMF indirectly. We first compute the CDF F_X and then obtain the PMF as

$$p_X(k) = F_X(k) - F_X(k-1), \qquad k = 1, \ldots, 10.$$

We have

$$\begin{aligned} F_X(k) &= \mathbf{P}(X \le k) \\ &= \mathbf{P}(X_1 \le k,\ X_2 \le k,\ X_3 \le k) \\ &= \mathbf{P}(X_1 \le k)\,\mathbf{P}(X_2 \le k)\,\mathbf{P}(X_3 \le k) \\ &= \left(\frac{k}{10}\right)^3. \end{aligned}$$

where the third equality follows from the independence of the events $\{X_1 \le k\}$, $\{X_2 \le k\}$, $\{X_3 \le k\}$. Thus, the PMF is given by

$$p_X(k) = \left(\frac{k}{10}\right)^3 - \left(\frac{k-1}{10}\right)^3, \qquad k = 1, \ldots, 10.$$

3)

Normal (Gaussian) Random Variable

<u>The most important type of random variable (and associated PDF, CDF) in probability theory and its applications</u>.

- Intuitive Reason of Importance
  - Normal random variables model an additive effect of many independent factors in a variety of engineering and physical contexts

- Mathematical Reason of Importance "i.i.d"
  - Sum of a large number of "independent and identically distributed" (not necessarily normal) random variables has an approximately normal CDF, regardless of the CDF of the individual random variables.

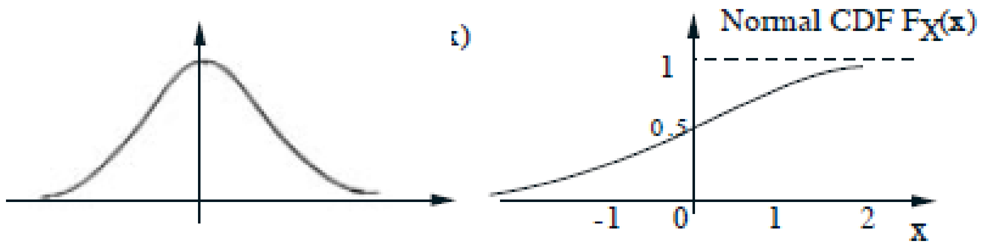
$$Y = X_1 + X_2 + X_3 + \cdots\cdots X_N$$

$Y \sim$ Normal Gaussian   N is large.   (Central Limit Theorem)

Normal (Gaussian) Random Variable

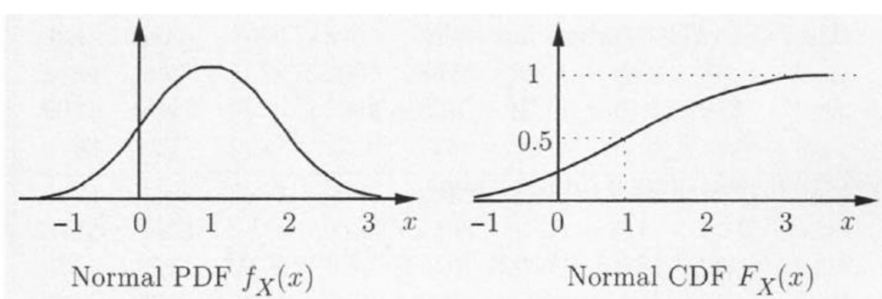
- Standard Normal RV

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



- General Normal Random Variable

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$



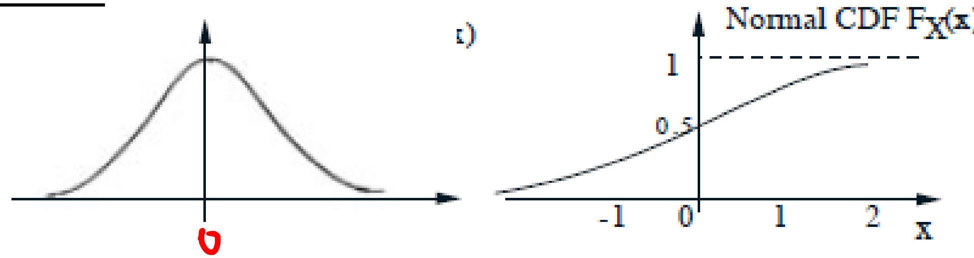
Normal PDF $f_X(x)$ Normal CDF $F_X(x)$

Normal (Gaussian) Random Variable

- Standard Normal RV

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

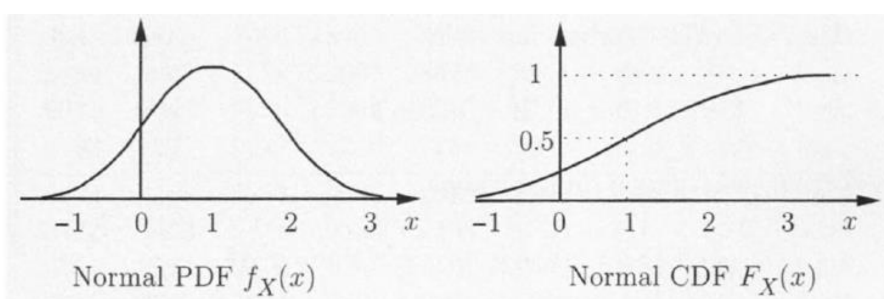
"Bell Curve"



- General Normal Random Variable

μ
σ^2

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$



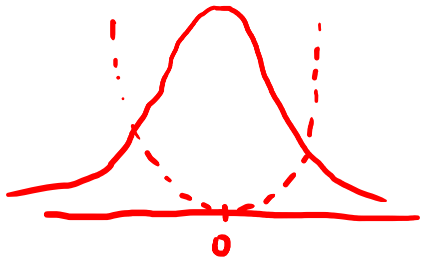
Normal PDF $f_X(x)$ Normal CDF $F_X(x)$

Normal (Gaussian) Random Variable

- Standard Normal RV

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$e^{-\frac{x^2}{2}}$



Mean = μ

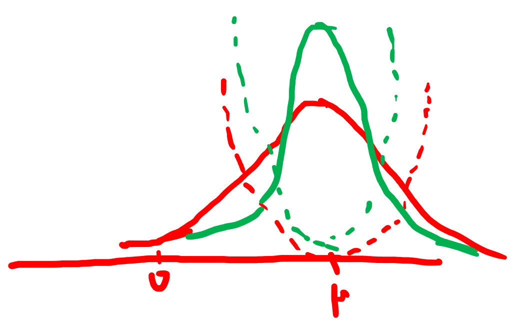
Variance: σ^2

$\frac{(x-3)^2}{2}$

- General Normal Random Variable

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

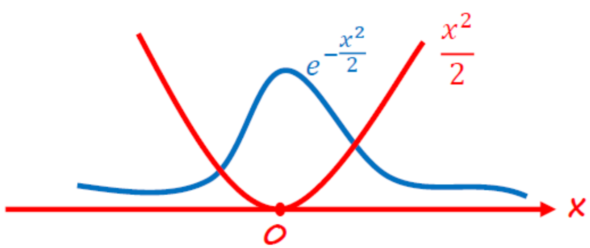


Lower σ^2 → Narrower Bell Curve

Normal (Gaussian) Random Variable

- Standard Normal RV

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

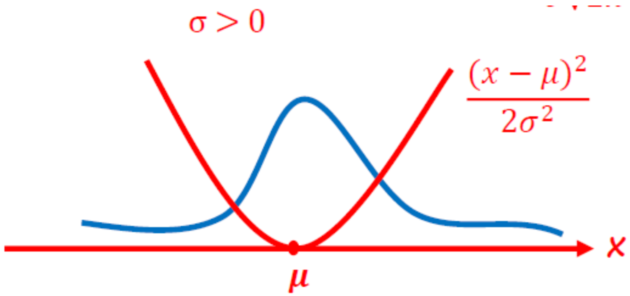


Calculus:

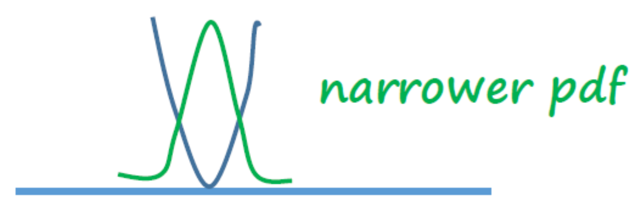
$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi}$$

- General Normal Random Variable

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



σ small



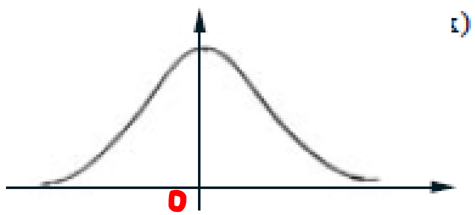
Normal (Gaussian) Random Variable

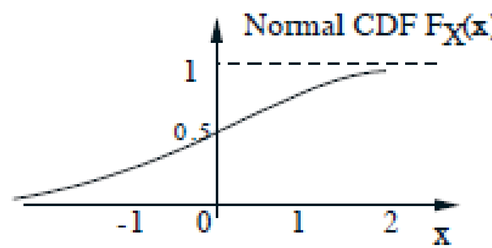
1) • Standard Normal RV $\boxed{N(\overset{\text{Mean}}{0}, \overset{\text{Variance}}{1})}$

$X \sim N(0,1)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

- E[X] = 0
- Var(X) = 1



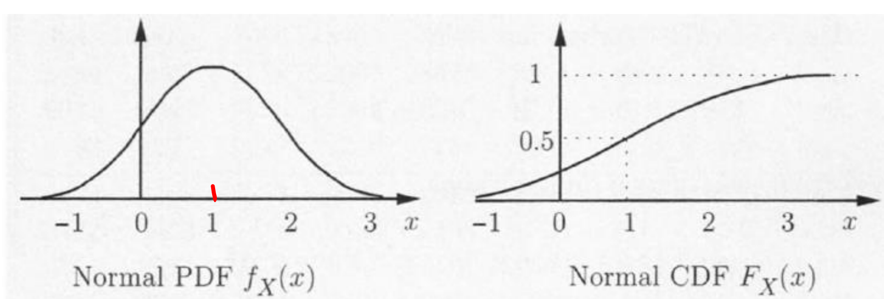


- General Normal Random Variable $\boxed{N(\mu, \sigma^2)}$

$X \sim N(4,5)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

- E[X] = μ
- Var(X) = σ^2



$E[Y] = aE[X] + b$

"Linear Functions" of a Normal Random Variable

$$X \sim N(\mu, \sigma^2)$$

Let $Y = aX + b$

$$\mu_Y = E[Y] = aE[X] + b = a\mu + b$$

$$\text{Var}(Y) = a^2 \text{Var}(X) = a^2\sigma^2$$

PDF of Y:- Y is also a normal RV.

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

Linear Functions of a Normal Random Variable

$$X \sim N(\mu, \sigma^2)$$

Let $Y = aX + b$

E[Y] = $a\mu + b$

Var[Y] = $a^2\sigma^2$

$$\boxed{Y \sim N(a\mu + b, a^2\sigma^2)}$$

Linear Functions of a Normal Random Variable

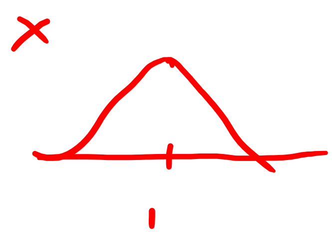
Normality is Preserved by Linear Transformations

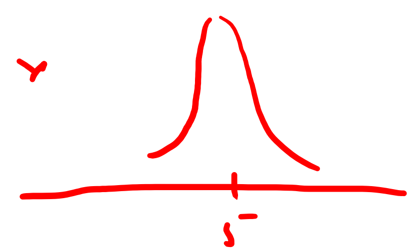
If X is a normal random variable with mean μ and variance σ^2, and if $a \neq 0$, b are scalars, then the random variable

$$\underset{N}{Y} = a\underset{N}{X} + b$$

is also normal, with mean and variance

$$\mathbf{E}[Y] = a\mu + b, \qquad \mathrm{var}(Y) = a^2\sigma^2.$$





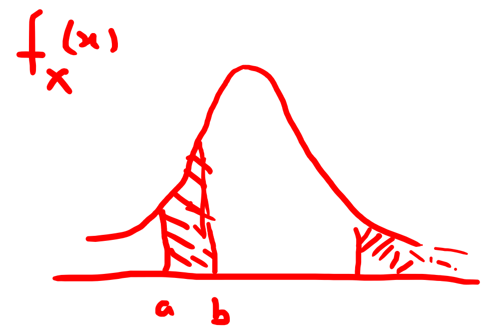
Let's Try to "Use" the Normal Random Variable

Let's say $X \sim \underline{N}(\mu_x, \sigma^2)$

1) find $P(X \geq 10)$?

$$= \int_{10}^{\infty} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} dx$$

No closed form available



$P(a \leq X \leq b)$

$P(X \geq c)$

Let’s Try to Use the Normal Random Variable

Let’s say $X \sim N(\mu_x, \sigma^2)$

1) find $P(X \geq 10)$?

$$= \int_{10}^{\infty} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} dx$$

No closed form available

2) How about $P(X \leq x)$?

$$= \int_{-\infty}^{x} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} dx$$

No closed form CDF

What should we do?

Let's Try to Use the Normal Random Variable

Let's say $X \sim N(\mu_x, \sigma^2)$

How about $P(X \leq x)$?

$$= \int_{-\infty}^{x} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} dx$$

No closed form CDF

What should we do?

$P(Y \leq -2) = P(Y \geq 2) = 1 - P(Y \leq 2) = 1 - 0.9772 =$

Let's Try to Use the Normal Random Variable

Let's say $X \sim N(\mu_x, \sigma^2)$

How about $P(X \leq x)$?

$$= \int_{-\infty}^{x} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} dx$$

No closed form CDF

What should we do?

The standard normal table.

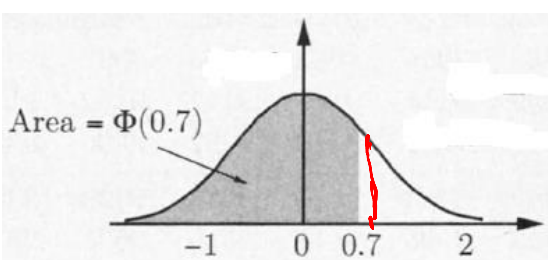
The entries in this table provide the numerical values of $Y \sim N(0,1)$

$$\Phi(y) = \mathbf{P}(Y \leq y)$$

$P(Y \leq 0.70) = 0.7580$

$P(Y \leq 0.72) = 0.7642$

where Y is a standard normal random variable, for y between 0 and 3.49



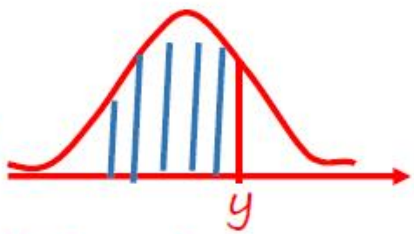
	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

How to Use the Standard Normal Table

Standard Normal Tables

- No closed form available for CDF
- But we have tables, for the standard normal

➡ $Y = N(0, 1)$

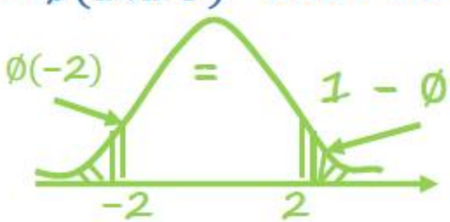


➡ $\phi(y) = F_Y(y) = P(Y \leq y)$

➡ $\phi(0) = P(Y \leq 0) = 0.5$

$P(Y \leq 1.16)$

➡ $\phi(1.16) = 0.8770$ $\phi(2.9) = 0.9981$



$1 - \phi(2) = 1 - 0.9772$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

How to Use the Standard Normal Table

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

"Standardizing" a Normal RV

$Y \sim N(0,1)$

$P(Y \leq 3)$

$X \sim N(\mu_x, \sigma^2)$ $\quad$ $X \sim N(4,6)$

$P(X \leq x)$ $\quad$ $P(X \leq 3)$

- Cannot use the table directly.

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

$\rightarrow E[Y] = aE[X] + b$ $\quad$ $Var(Y) = a^2 Var(X)$

$Y = aX + b$

“Standardizing” a Normal RV

$$\boxed{X \sim N(\mu_x, \sigma^2)}$$

$X \sim N(4, 6)$

$P(X \leq x)$ $\quad$ $P(X \leq 3)$

- Cannot use the table directly.

$P(X \leq x) = P(X - \mu_X \leq x - \mu_X)$

$= P\left(\frac{X - \mu_X}{\sigma_X} \leq \frac{x - \mu_X}{\sigma_X}\right)$

$P(X \leq x) = P\left(Y \leq \frac{x - \mu_X}{\sigma_X}\right)$

Where $\boxed{Y = \frac{X - \mu_X}{\sigma_X}}$

$= \frac{1}{\sigma_X} X - \frac{\mu_X}{\sigma_X}$

$Y = aX + b$

$a = \frac{1}{\sigma_X}$ $\quad$ $b = -\frac{\mu_X}{\sigma_X}$

$E[Y] = E\left[\frac{1}{\sigma_X} X - \frac{\mu_X}{\sigma_X}\right]$

$= \frac{1}{\sigma_X}\mu_X - \frac{\mu_X}{\sigma_X} = \boxed{0}$

$Var(Y) = \frac{1}{\sigma_X^2}\sigma_X^2 = \boxed{1}$

$Y = \frac{X - \text{Mean}}{\text{S.D}}$

$$\boxed{Y \sim N(0, 1)}$$

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

"Standardizing" a Normal RV

$X \sim N(\mu_x, \sigma^2)$

$X \sim N(6, 4)$

Varance = 4

S.D. = 2

$P(2 \le X \le 8)$

$= P\left(\frac{2 - \text{Mean}}{S.D} \le \frac{X - \text{Mean}}{S.D} \le \frac{8 - \text{Mean}}{S.D}\right)$

$= P\left(\frac{2-6}{2} \le Y \le \frac{8-6}{2}\right)$

$= P(-2 \le Y \le 1)$

$= P(Y \le 1) - P(Y \le -2)$

$= P(Y \le 1) - P(Y \ge 2)$

$= P(Y \le 1) - [1 - P(Y \le 2)]$

$= 0.8413 - [1 - 0.9772]$

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

“Standardizing” a Normal RV

Calculating normal probabilities

- Express an event of interest in terms of standard normal

$\rightarrow$ $X \sim (6, 4) \qquad \sigma = 2$

st, normal

$\rightarrow$ $\dfrac{2-6}{2} \le \dfrac{X-6}{2} \le \dfrac{8-6}{2}$

$\rightarrow$ $P(2 \le X \le 8) = P(-2 \le Y \le 1)$

$= P(Y \le 1) - P(Y \le -2)$

$= P(Y \le 1) - (1 - P(Y \le 2))$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

CDF Calculation for a Normal RV

CDF Calculation for a Normal Random Variable

For a General normal random variable X with mean μ and variance σ^2, we use a two-step procedure.

(a) "Standardize" X, i.e., subtract μ and divide by σ to obtain a standard normal random variable Y.

(b) Read the CDF value from the standard normal table:

$$\mathbf{P}(X \leq x) = \mathbf{P}\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \mathbf{P}\left(Y \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

CDF Calculation for a Normal RV

Example 3.7. Using the Normal Table. The annual snowfall at a particular geographic location is modeled as a normal random variable with a mean of $\mu = 60$ inches and a standard deviation of $\sigma = 20$. What is the probability that this year's snowfall will be at least 80 inches?

Annual Snowfall : $X \sim N(60, 20^2)$

$P(X \geq 80) = ?$

$$P(X \geq 80) = P\left(\frac{X - \text{Mean}}{\text{S.D}} \geq \frac{80 - \text{Mean}}{\text{S.D}}\right)$$

$$= P\left(Y \geq \frac{80 - 60}{20}\right)$$

$$= P(Y \geq 1)$$

$$= 1 - P(Y \leq 1)$$

$$= 1 - 0.8413 = 0.1587$$

CDF Calculation for a Normal RV

Example 3.7. Using the Normal Table. The annual snowfall at a particular geographic location is modeled as a normal random variable with a mean of $\mu = 60$ inches and a standard deviation of $\sigma = 20$. What is the probability that this year's snowfall will be at least 80 inches?

Let X be the snow accumulation, viewed as a normal random variable, and let

$$Y = \frac{X - \mu}{\sigma} = \frac{X - 60}{20},$$

be the corresponding standard normal random variable. We have

$$\mathbf{P}(X \geq 80) = \mathbf{P}\left(\frac{X - 60}{20} \geq \frac{80 - 60}{20}\right) = \mathbf{P}\left(Y \geq \frac{80 - 60}{20}\right) = \mathbf{P}(Y \geq 1) = 1 - \Phi(1),$$

where Φ is the CDF of the standard normal. We read the value $\Phi(1)$ from the table:

$$\Phi(1) = 0.8413,$$

so that

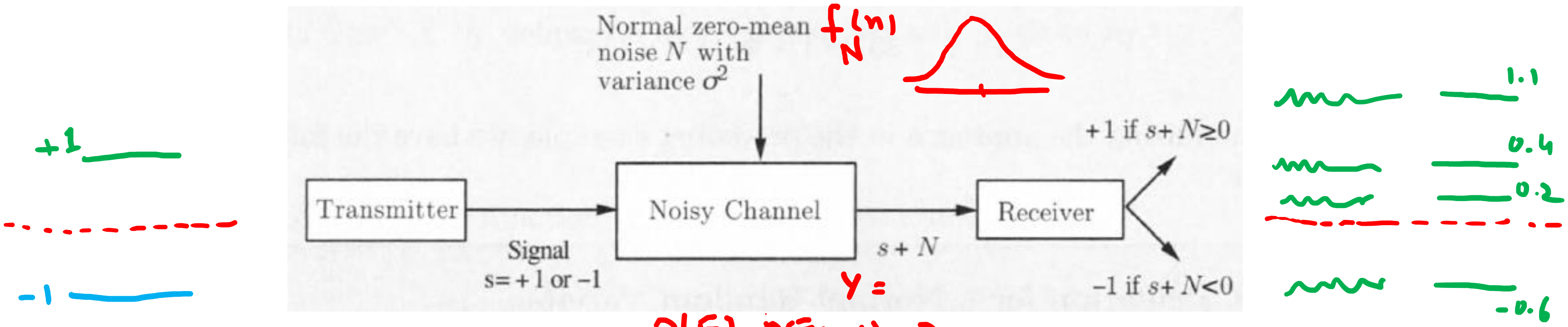
$$\mathbf{P}(X \geq 80) = 1 - \Phi(1) = 0.1587.$$

$P(N \le -1) = P(N \ge 1) = 1 - P(N \le 1)$

$P(N \ge 1) = 1 - P(N \le 1)$

Modeling Applications of a Normal RV: Example

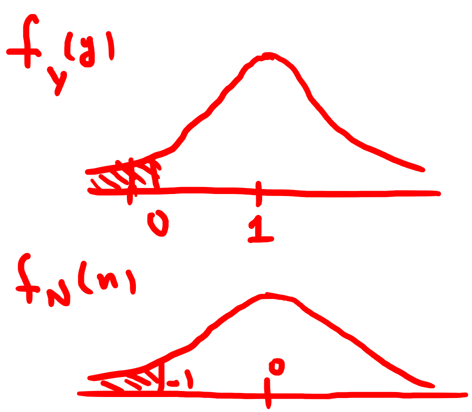
Example 3.8. Signal Detection. A binary message is transmitted as a signal s, which is either -1 or $+1$. The communication channel corrupts the transmission with additive normal noise with mean $\mu = 0$ and variance $\sigma^2 = 1$. The receiver concludes that the signal -1 (or $+1$) was transmitted if the value received is < 0 (or ≥ 0, respectively); see Fig. 3.11. What is the probability of error?



$P(E) = P(\text{Error}) = ?$

$s = +1$

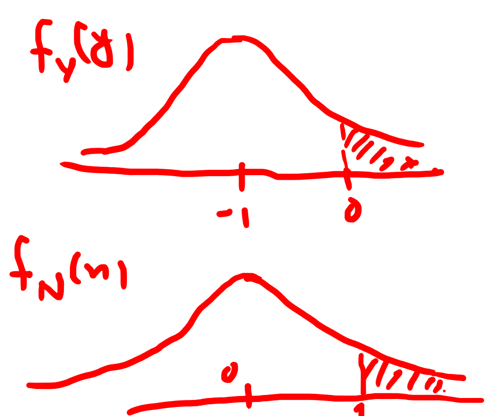
$Y = s + N = 1 + N$



$P(\text{Error})$
$= P(Y \le 0)$
$= P(N \le -1)$
$= 0.1587$

$s = -1$

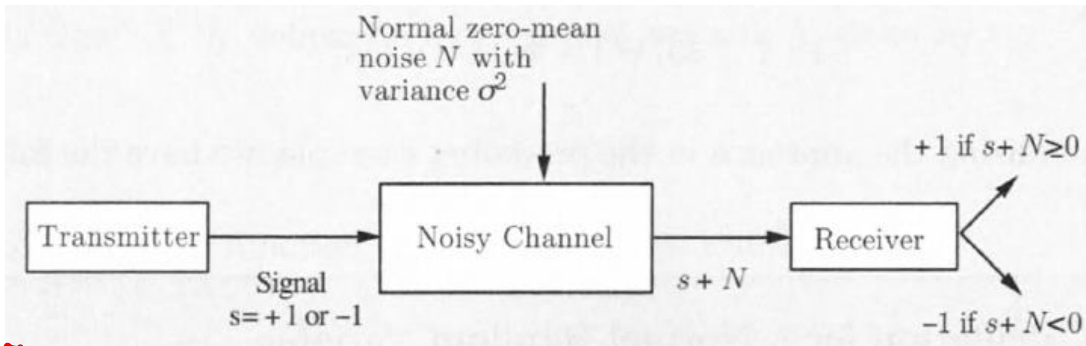
$Y = s + N = -1 + N$



$P(\text{Error})$
$= P(Y \ge 0)$
$= P(N \ge 1)$
$= 0.1587$

Modeling Applications of a Normal RV: Example

Example 3.8. Signal Detection. A binary message is transmitted as a signal s, which is either -1 or $+1$. The communication channel corrupts the transmission with additive normal noise with mean $\mu = 0$ and variance σ^2. The receiver concludes that the signal -1 (or $+1$) was transmitted if the value received is < 0 (or ≥ 0, respectively); see Fig. 3.11. What is the probability of error?



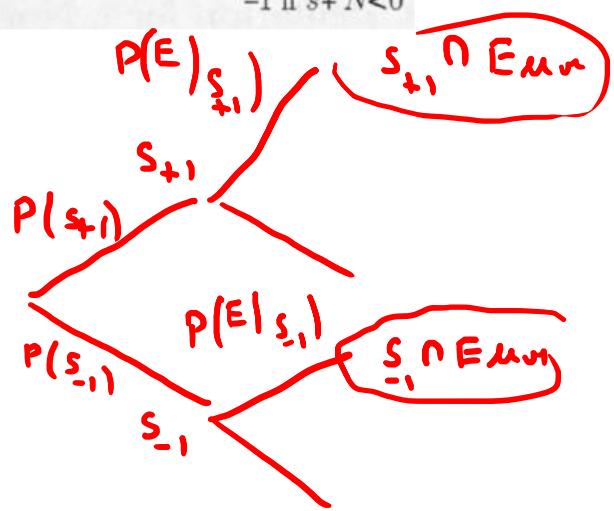
Total Prob Thm

$$P(E_{rr}) = P(s_1 \cap E_{rr}) + P(s_{-1} \cap E_{rr})$$

$$= P(s_{+1})\, P(E \mid s_{+1}) + P(s_{-1})\, P(E \mid s_{-1})$$

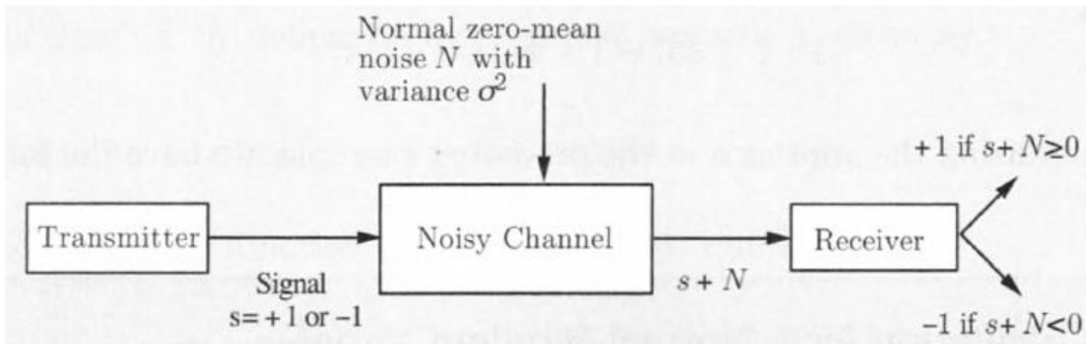
$$P(E_{rr}) = 0.5\,(0.1587) + 0.5\,(0.1587)$$

$$=$$



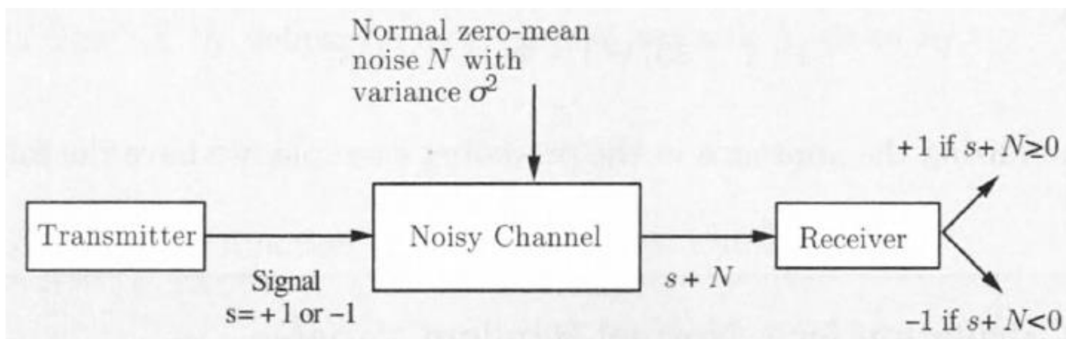
Modeling Applications of a Normal RV: Example

Example 3.8. Signal Detection. A binary message is transmitted as a signal s, which is either -1 or $+1$. The communication channel corrupts the transmission with additive normal noise with mean $\mu = 0$ and variance σ^2. The receiver concludes that the signal -1 (or $+1$) was transmitted if the value received is < 0 (or ≥ 0, respectively); see Fig. 3.11. What is the probability of error?



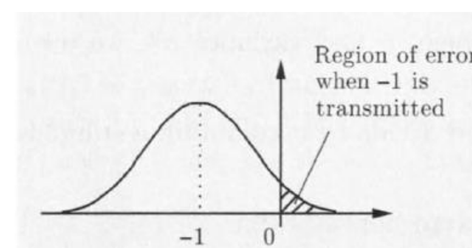
Modeling Applications of a Normal RV: Example

Example 3.8. Signal Detection. A binary message is transmitted as a signal s, which is either -1 or $+1$. The communication channel corrupts the transmission with additive normal noise with mean $\mu = 0$ and variance σ^2. The receiver concludes that the signal -1 (or $+1$) was transmitted if the value received is < 0 (or ≥ 0, respectively); see Fig. 3.11. What is the probability of error?

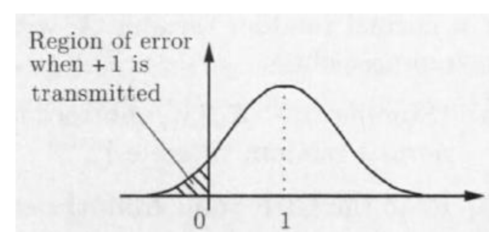


An error occurs whenever -1 is transmitted and the noise N is at least 1 so that $s+N = -1+N \geq 0$, or whenever $+1$ is transmitted and the noise N is smaller than -1 so that $s + N = 1 + N < 0$. In the former case, the probability of error is

$$\mathbf{P}(N \geq 1) = 1 - \mathbf{P}(N < 1) = 1 - \mathbf{P}\left(\frac{N-\mu}{\sigma} < \frac{1-\mu}{\sigma}\right)$$
$$= 1 - \Phi\left(\frac{1-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right).$$



In the latter case, the probability of error is the same, by symmetry. The value of $\Phi(1/\sigma)$ can be obtained from the normal table. For $\sigma = 1$, we have $\Phi(1/\sigma) = \Phi(1) = 0.8413$, and the probability of error is 0.1587.



Big Picture

Unit 4 | Unit 5

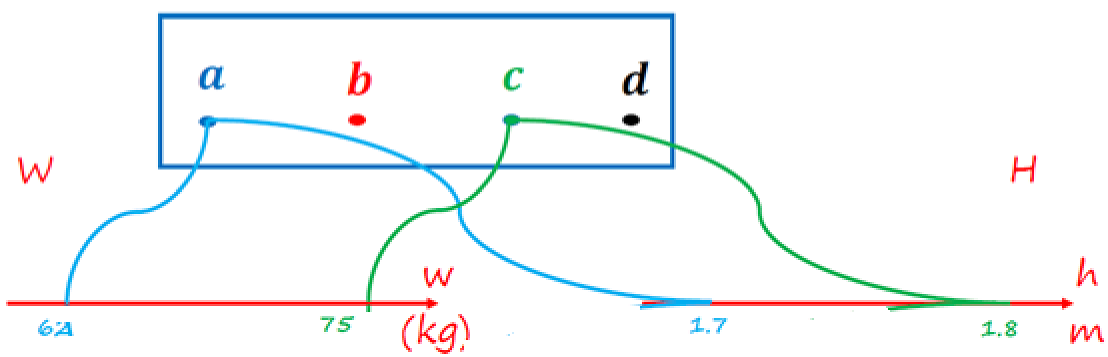
	Discrete RV		Continuous RV	
PMF	$p_X(x)$	(CDF) $F_X(x)$	$f_X(x)$	PDF
	$\sum_x x p_X(x)$	$\mathbf{E}[X]$	$\int x f_X(x)\,dx$	
		$\text{var}(X)$		
Joint PMF	$p_{X,Y}(x,y)$		$f_{X,Y}(x,y)$	Joint PDF
Conditional PMF	$p_{X\|A}(x)$		$f_{X\|A}(x)$	Conditional PDF
	$p_{X\|Y}(x \mid y)$		$f_{X\|Y}(x \mid y)$	

Independence of RVs

Remember: Multiple Random Variables on the Same Sample Space

- We can have multiple random variables defined on the same sample space

- Example
  - Weight of a randomly selected student in the class (W)
  - Height of a randomly selected student in the class (H)
  - GPA of a randomly selected student in the class (G)



Remember: Multiple Random Variables and Joint PMF

- Consider two discrete random variables X and Y associated with the same experiment.
  - The probabilities of the values that X and Y can take are captured by the **joint PMF** of X and Y, denoted by $\boldsymbol{\rho}_{X,Y}$
  - In particular, if (x, y) is a pair of possible values of X and Y, the probability mass of (x, y) is the probability of the event $\{X = x, Y = y\}$

Joint PMF $p_{X,Y}(x,y)$ in tabular form

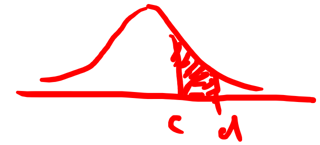
y \ x	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

$$p_{X,Y}(x,y) = \mathbf{P}(X = x, Y = y).$$

$$= \mathbf{P}(\{X = x\} \cap \{Y = y\})$$

$$= \mathbf{P}(X = x \text{ and } Y = y)$$

X:  $f_X(x)$ $P(a \le X \le b)$

Y: $f_Y(y)$  $P(c \le Y \le d)$

Joint PDF of Multiple Continuous Random Variables

Joint PDF $f_{X,Y}(x,y)$

> We say that two continuous random variables associated with the same experiment are **jointly continuous** and can be described in terms of a **joint PDF** $f_{X,Y}$ if $f_{X,Y}$ is a nonnegative function that satisfies
>
> $$\mathbf{P}((X,Y) \in B) = \int\int_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy,$$
>
> for every subset B of the two-dimensional plane. The notation above means that the integration is carried over the set B.

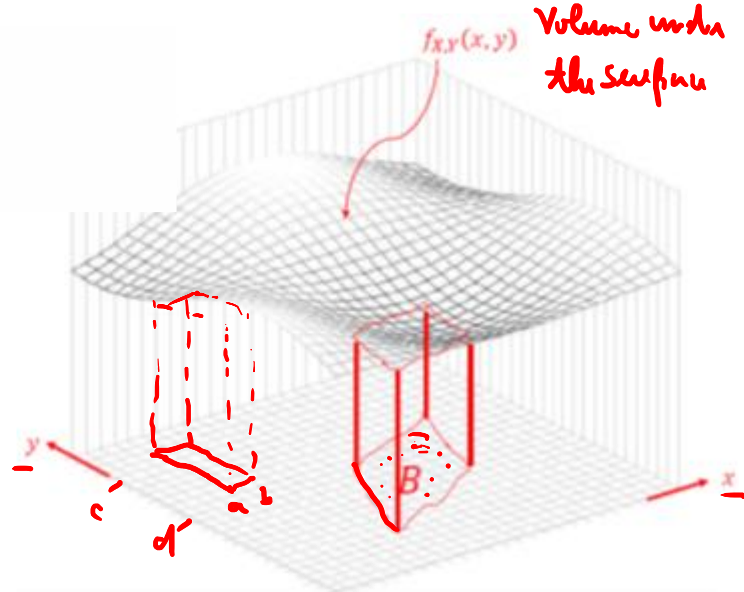
$P((x,y) \in B) =$ Volume under the surface

a rectangle of the form $B = \{(x,y) \mid a \le x \le b,\ c \le y \le d\}$, we have

$$\boxed{\mathbf{P}(a \le X \le b,\ c \le Y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\,dx\,dy.}$$

- Normalization Property

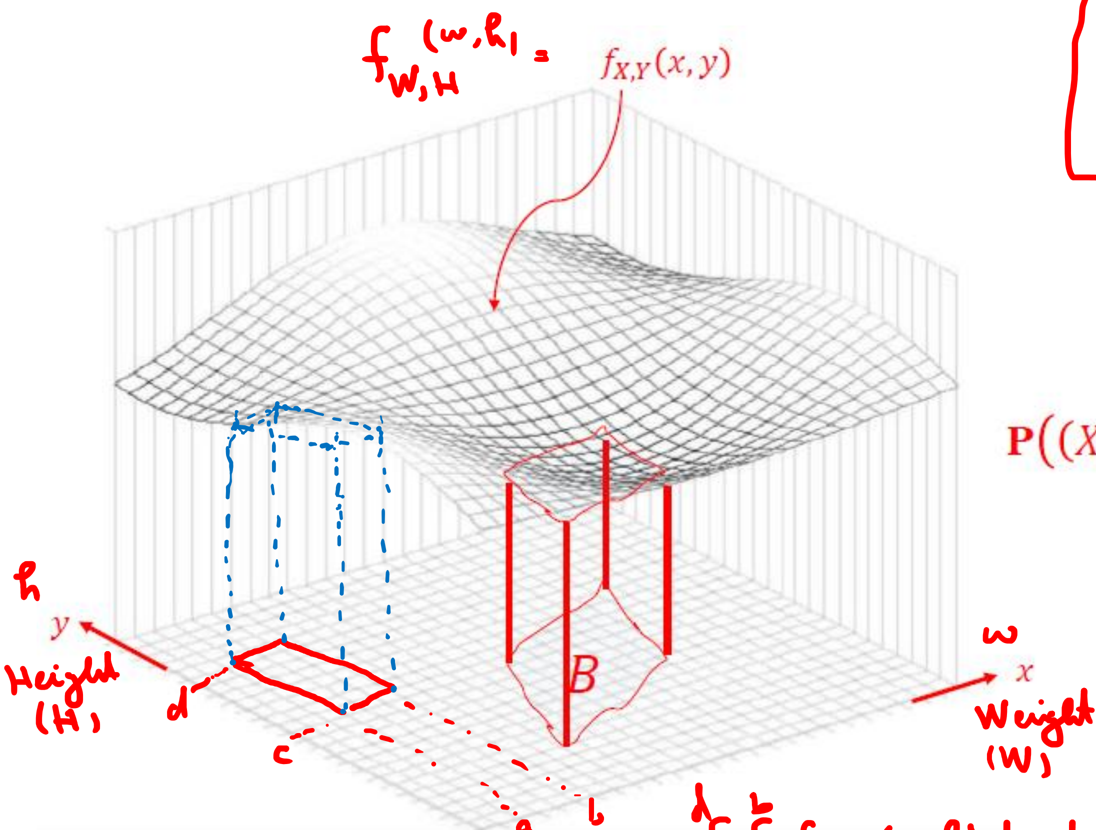
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1.$$



Visualizing Joint PDF of Multiple Continuous RVs

$$P(-\infty \le W \le \infty, -\infty \le H \le \infty) = 1$$

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,H}(w,h)\,dw\,dh = 1}$$

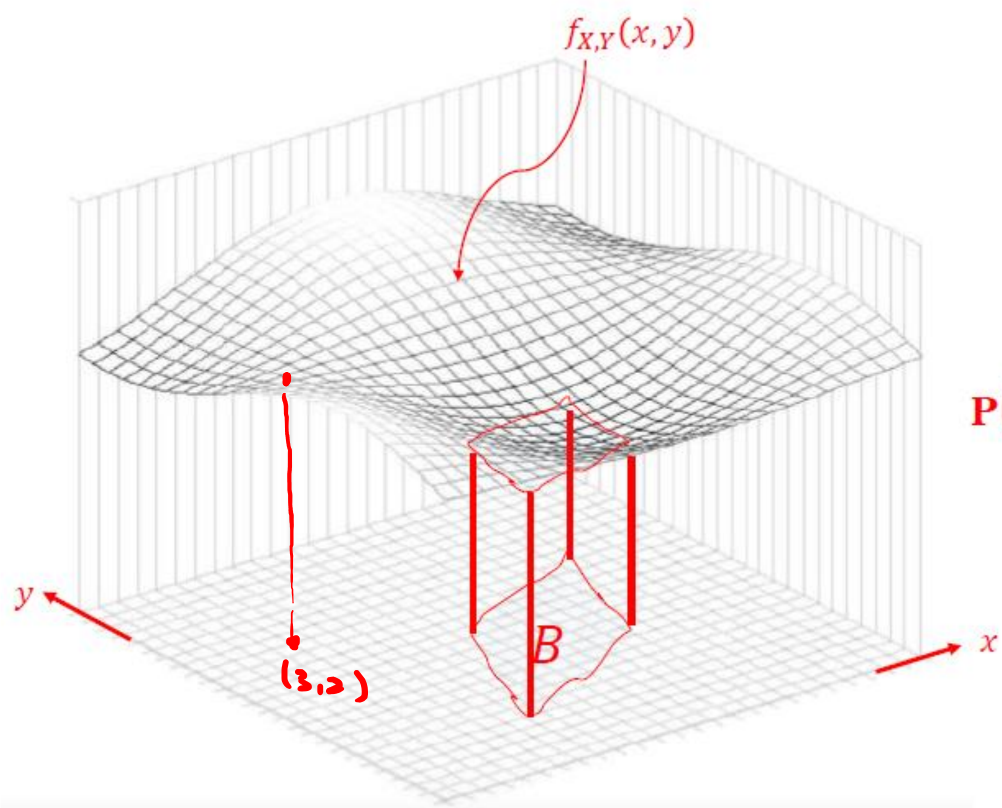


$$\mathbf{P}\big((X,Y) \in B\big) = \iint_{(X,Y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

$$P(a \le W \le b,\ c \le H \le d) = \int_c^d \int_a^b f_{W,H}(w,h)\,dw\,dh$$

$$P(55 \le W \le 70,\ 5 \le H \le 6) = \int_5^6 \int_{55}^{70} f_{W,H}(w,h)\,dw\,dh$$

Visualizing Joint PDF of Multiple Continuous RVs

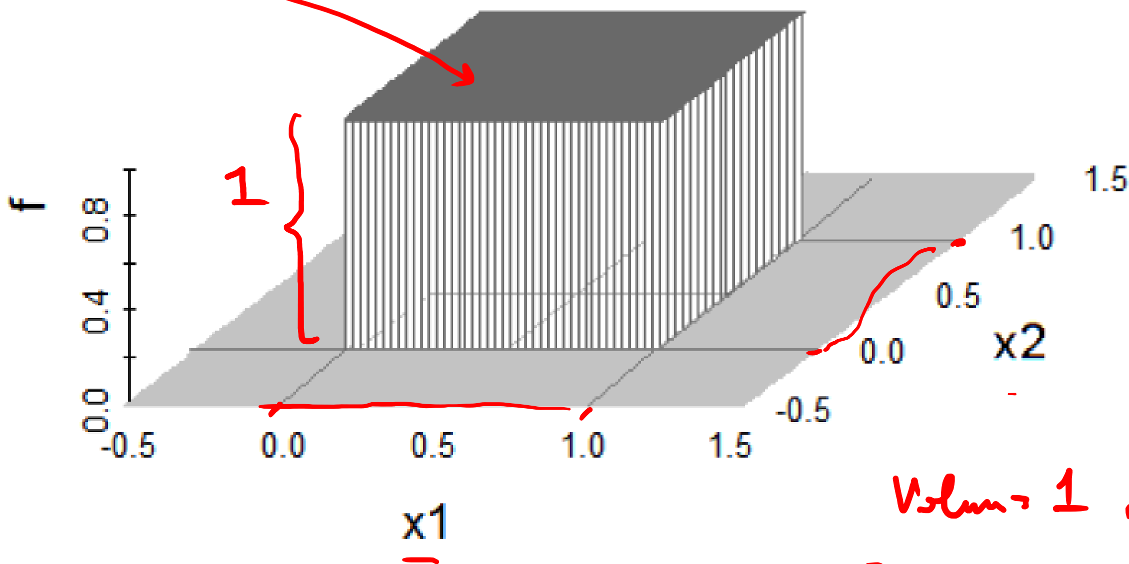


$$\mathbf{P}\big((X,Y) \in B\big) = \iint_{(X,Y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

Visualizing Joint PDF of Multiple Continuous RVs

Uniform

$f_{X_1,X_2}(x_1,x_2)$



bivariate uniform

Volume = 1

Base × Height = 1

1 × 1

$P((x,y) \in A) = ?$

$= P((x,y) \in A \cap S)$

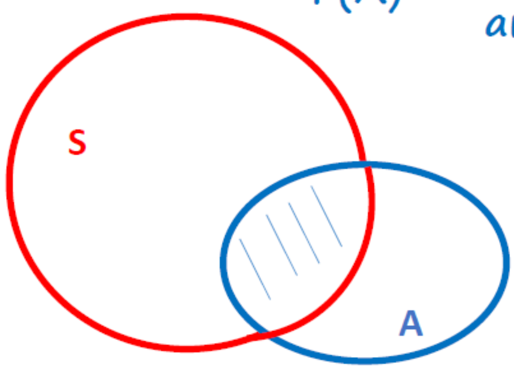
Joint PDF of Multiple Continuous RVs: Examples

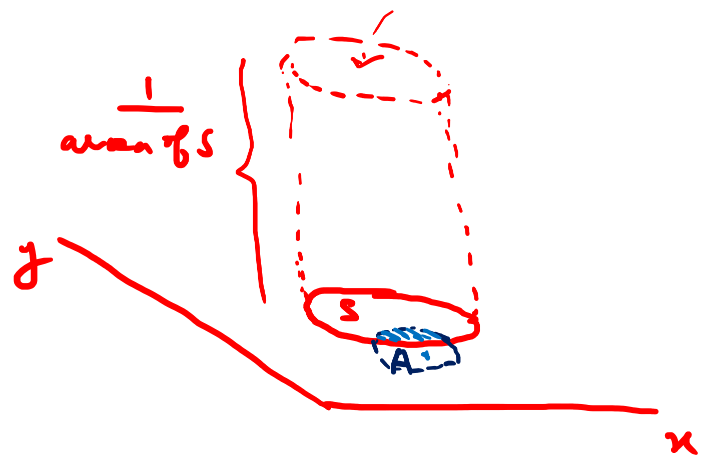
• Joint PDF of X and Y is "Uniform on set S".

"Uniform joint PDF on a set S"

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{area\ of\ S}, & if\ (x,y) \in S, \\ 0, & \text{Otherwise.} \end{cases}$$

$P(x,y \in A) = P(A) = \dfrac{area\ (A \cap S)}{area\ (S)}$





Volume = Base x Height

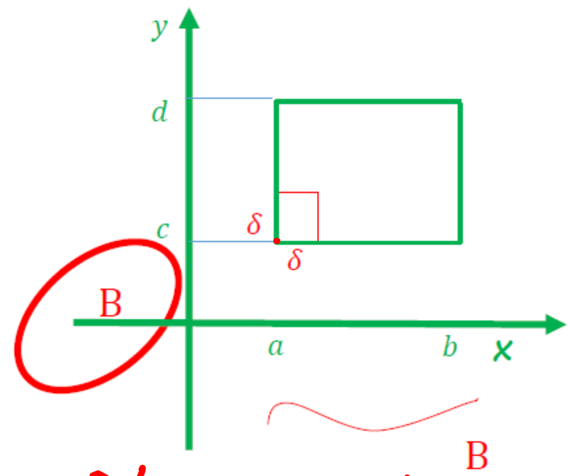
$= \text{Area of } (A \cap S) \times \dfrac{1}{\text{area of } S}$

$P(X=3, Y=2) = 0$

Joint PDF of Multiple Continuous RVs: Interpretation

$$\mathbf{P}\big((X,Y) \in B\big) = \iint_{(X,Y)\in B} f_{X,Y}(x,y)\,dxdy$$

$$\mathbf{P}(a \le X \le b, c \le Y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\,dxdy$$



DRV:

✓ $p_X(3) = P(X=3)$

CRV:-

$f_X(3) \neq P(X=3)$

✓ $f_X(3) = \dfrac{P(3 \le X \le 3+\delta)}{\delta}$

= "Prob per unit length."

DRV:

✓ $p_{X,Y}(3,2) = P(X=3, Y=2)$

CRV:-

$f_{X,Y}(3,2) \neq P(X=3, Y=2)$

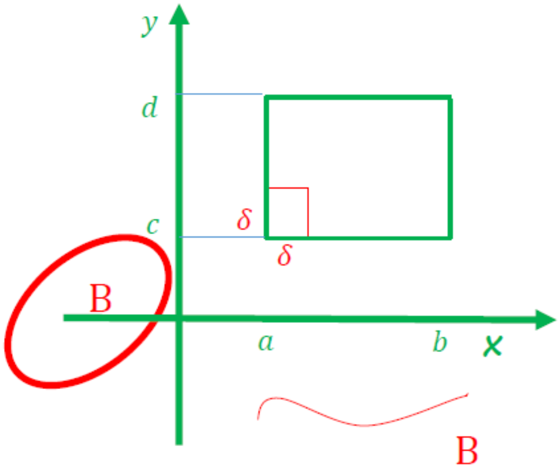
$f_{X,Y}(3,2) = \dfrac{P(3 \le X \le 3+\delta,\ 2 \le Y \le 2+\delta)}{\delta^2}$

= Prob per unit Area

Joint PDF of Multiple Continuous RVs: Interpretation

$$\mathbf{P}\big((X,Y) \in B\big) = \iint_{(X,Y)\in B} f_{X,Y}(x,y)\,dxdy$$

$$\mathbf{P}(a \le X \le b, c \le Y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\,dxdy$$



$$\mathbf{P}(a \le X \le a+\delta, c \le Y \le c+\delta) \approx f_{X,Y}(a,c)\,\delta^2$$

$f_{X,Y}(x,y)$: "probability per unit area" Joint PDF area

Continuous RVs: From Joint PDF to Marginal PDF

Joint PMF $\longrightarrow$ Marginal PMF

$$p_X(x) = \sum_y p_{X,Y}(x,y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y)$$

Joint PDF $\longrightarrow$ Marginal PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx.$$

let A be a subset of the real line and consider the event $\{X \in A\}$. We have

$$\mathbf{P}(X \in A) = \mathbf{P}\big(X \in A \text{ and } Y \in (-\infty,\infty)\big) = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx.$$

Comparing with the formula

$$\mathbf{P}(X \in A) = \int_A f_X(x)\,dx.$$

we see that the **marginal** PDF f_X of X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy.$$

Joint PDF of Multiple Continuous RVs: Examples

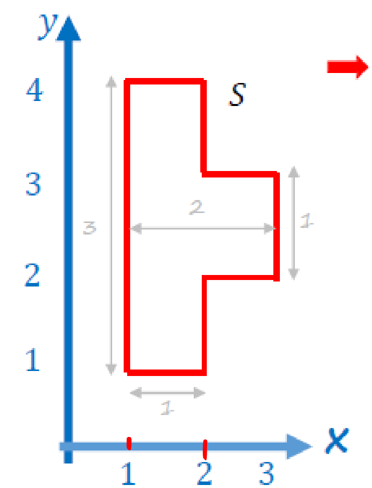
$f_{X,Y}(x,y)$:

Uniform joint PDF on a set S

(a) $f_{X,Y}(4, 4) = ?$ $\quad f_{X,Y}(4,4) = 0$

(b) $f_{X,Y}(2, 2.5) = ?$ $\quad = \frac{1}{4}$

Area of $S = (1\times 3) + (1\times 1) = 4$

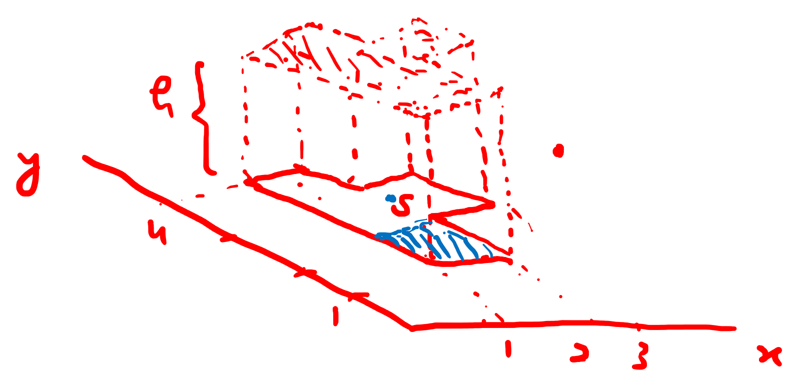


(c) $P(1 < X < 2, 1 < Y < 2) = ?$

$=$ Base $\times$ Height

$(1\times 1) \times \frac{1}{4}$

$= \frac{1}{4}$



Joint PDF of Multiple Continuous RVs: Examples

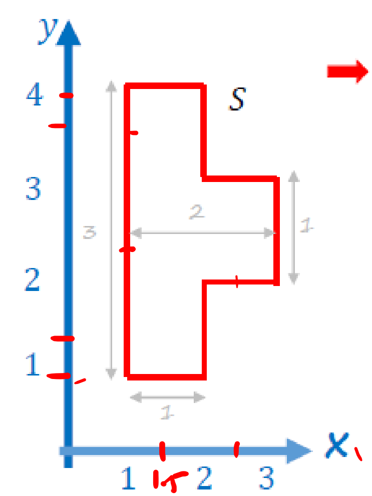
Uniform joint PDF on a set S

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

(d) $f_X(x) = ?$

$$f_X(1.5) = \int_{-\infty}^{\infty} f_{X,Y}(1.5, y)\cdot dy = \int_{1}^{4} \frac{1}{4}\cdot dy = \frac{3}{4}$$

$$f_X(2.5) = \int_{-\infty}^{\infty} f_{X,Y}(2.5, y)\cdot dy = \int_{2}^{3} \frac{1}{4}\cdot dy = \frac{1}{4}$$



$f_X(0.5) = 0$ $\quad f_X(3.5) = 0$

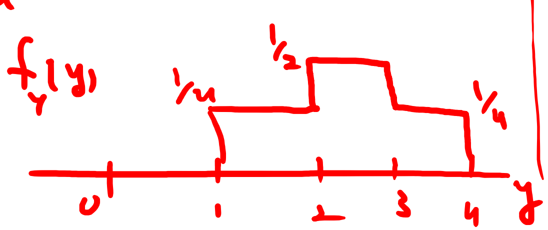
(e) $f_Y(y) = ?$

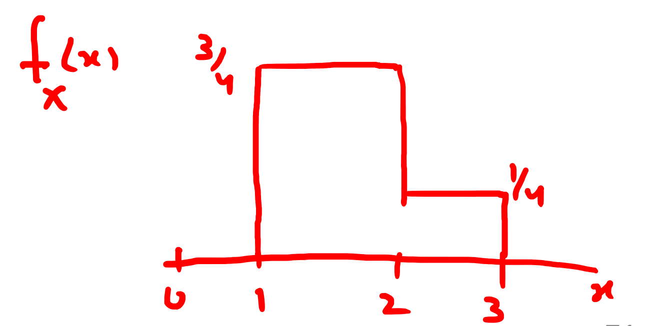
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

$$f_Y(1.5) = \int_{-\infty}^{\infty} f_{X,Y}(x, 1.5)\cdot dx = \int_{1}^{2} \frac{1}{4}\cdot dx = 1/4$$

$$f_Y(2.5) = \int_{1}^{3} \frac{1}{4}\cdot dx = \frac{2}{4} = 1/2$$

$$f_Y(3.5) = \int_{1} \frac{1}{4}\cdot dx = \frac{1}{4}$$

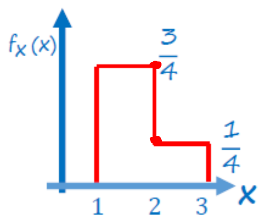


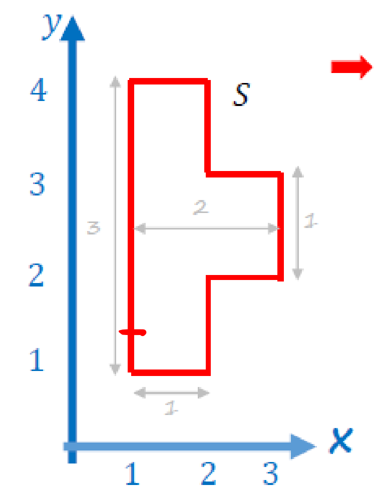


Joint PDF of Multiple Continuous RVs: Examples

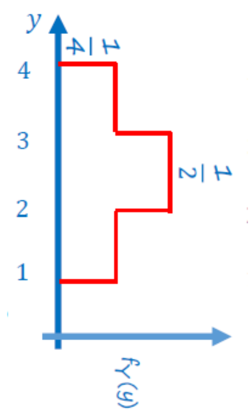
Uniform joint PDF on a set S

(d) $f_X(x) = ?$





(e) $f_Y(y) = ?$



Functions of Multiple Continuous RVs

Joint PDF : $f_{X,Y}(x,y)$

➡ $Z = g(X,Y)$ $Z = 2X + 3Y$

Expected value rule:

➡ $E[g(X,Y)] = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$ $E[g(X,Y)] = \int\int g(x,y) f_{X,Y}(x,y)\,dx\,dy$

μ_Z

Linearity of expectations

➡ $E[aX + b] = aE[X] + b$

➡ $E[X + Y] = E[X] + E[Y]$

$$\boldsymbol{E[X_1 + \ldots + X_n] = E[X_1] + \ldots + E[X_n]}$$

CDF:. $F_X(x) = P(X \leq x)$

$F_X(1.5) = P(X \leq 1.5^-)$

$F_{X,Y}(x,y)$

Joint CDF

If X and Y are two random variables associated with the same experiment, we define their joint CDF by

$$F_{X,Y}(x,y) = \mathbf{P}(X \leq x,\ Y \leq y).$$

$f_{X,Y}(x,y)$

$F_{X,Y}(2,3) = \boxed{P(X \leq 2, Y \leq 3)}$

$f_{X,Y}(x,y)$

As in the case of a single random variable, the advantage of working with the CDF is that it applies equally well to discrete and continuous random variables.

In particular, if X and Y are described by a joint PDF $f_{X,Y}$, then

$$F_{X,Y}(x,y) = \mathbf{P}(X \leq x,\ Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\, dt\, ds.$$

Conversely, the PDF can be recovered from the CDF by differentiating:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y).$$

$F_{X,Y}(0.2, 0.3) = 0.2 \times 0.3$

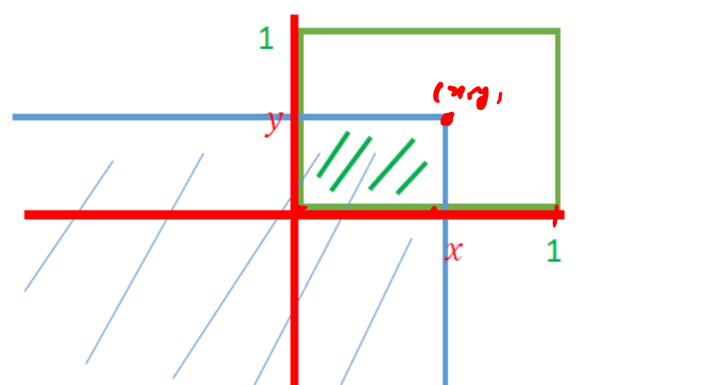
=

$F_{X,Y}(0.4, 0.5) = 0.4 \times 0.5$

Example:

$f_{X,Y}(x,y) = 1$, for $0 < (x, y) < 1$

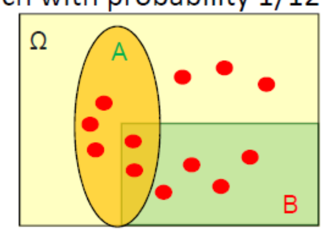
$$F_{X,Y}(x,y) = \mathbf{P}(X \leq x,\ Y \leq y) = xy, \qquad \text{for } 0 \leq x, y \leq 1.$$



Remember: Conditional Probability

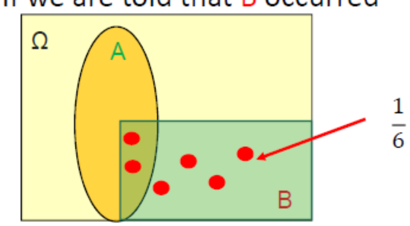
- Probability Law:
  - Encodes our "belief" about the likelihood of an event in a random experiment.
- Conditional Probability Law:
  - Encodes our "updated belief" about the likelihood of an event **IF** we have "additional partial information" about the results of the random experiment.

Assume 12 equally likely outcomes, each with probability 1/12



$P(A) = \frac{5}{12}$ $P(B) = \frac{6}{12}$

If we are told that B occurred



$P(A|B) = \frac{2}{6} = \frac{1}{3}$

- $P(A \mid B)$ = probability of A, given that B occurred
  - B is our new universe

- **Definition:** Assuming $P(B) \neq 0$,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$P(A \mid B)$ undefined if $P(B) = 0$

Remember: Conditioning a Discrete Random Variable on an Event

- <u>Conditional PMF</u>
  - The conditional PMF of a random variable X, conditioned on a particular event A with P(A) > 0, is defined by

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A) = \frac{\mathbf{P}(\{X = x\} \cap A)}{\mathbf{P}(A)}.$$

$$\boxed{p_X(x) = P\,(X = x)}$$

  - Similar to the unconditional case

$$\sum_x p_{X|A}(x) = 1,$$

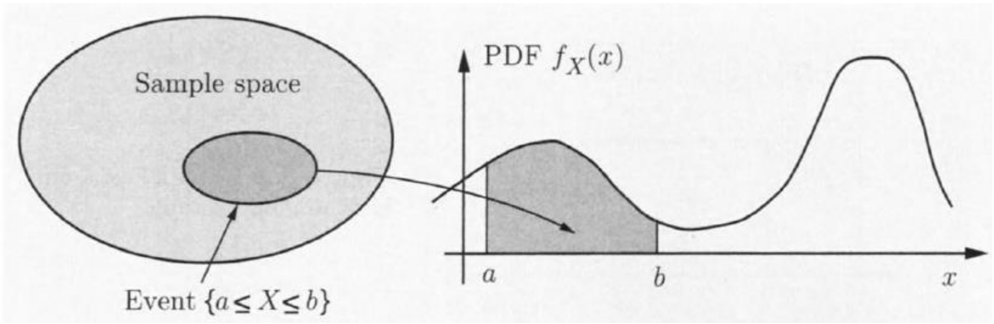
$$\boxed{\sum_x p_X(x) = 1}$$

- <u>Conditional Expectation</u>

$$E[X] = \sum_x x p_X(x) \longrightarrow E[X|A] = \sum_x x p_{X|A}(x)$$

$$E[g(X)] = \sum_x g(x) p_X(x) \longrightarrow E[g(X)|A] = \sum_x g(x) p_{X|A}(x)$$

Remember: Probability Density Functions (PDFs)



$$\mathbf{P}(a \le X \le b) = \int_a^b f_X(x)\,dx$$

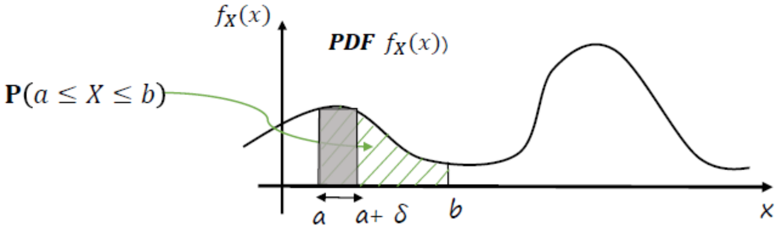
- Properties:

$$f_X(x) \ge 0 \qquad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

Note:

$$\mathbf{P}(a \le X \le a + \delta) \approx f_X(a)\delta$$

$$\mathbf{P}(X = a) = 0$$



H : Height of Student in class

Conditioning a Continuous Random Variable on an Event

- <u>Conditional PDF</u>

> The **conditional PDF** of a continuous random variable X, given an event A with $\mathbf{P}(A) > 0$, is defined as a nonnegative function $f_{X|A}$ that satisfies
>
> $$\mathbf{P}(X \in B \mid A) = \int_B f_{X|A}(x)\,dx,$$
>
> for any subset B of the real line.

  – Similar to the unconditional case

$A = \{Male\}$

$$\int_{-\infty}^{\infty} f_{X|A}(x)\,dx = 1.$$

$A = \{H \geq 72\}$ so that $f_{X|A}$ is a legitimate PDF.

$f_H(h)$

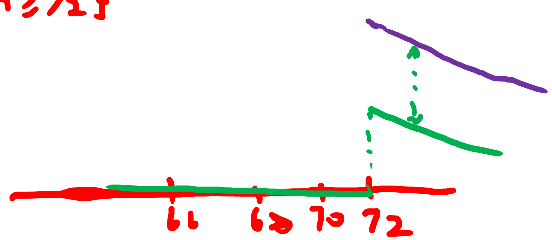
$P(60 \leq H \leq 62)$

$f_{H|Male}(h)$

$E[H|Male]$

$P(60 \leq H \leq 62 \mid Male)$

$f_{H|\{H \geq 72\}}(h)$



Conditioning a Continuous Random Variable on an Event

- Conditional PDF

> The **conditional PDF** of a continuous random variable X, given an event A with $\mathbf{P}(A) > 0$, is defined as a nonnegative function $f_{X|A}$ that satisfies
>
> $$\mathbf{P}(X \in B \mid A) = \int_B f_{X|A}(x)\,dx,$$
>
> for any subset B of the real line.

  – Similar to the unconditional case

$$\int_{-\infty}^{\infty} f_{X|A}(x)\,dx = 1.$$

so that $f_{X|A}$ is a legitimate PDF.

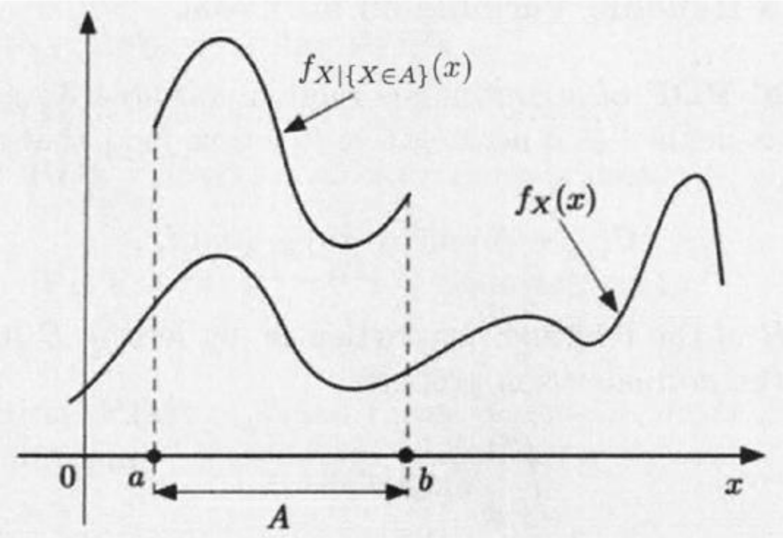
- Conditional Expectation

$$E[X] = \sum_x x p_X(x) \longrightarrow \mathbf{E}[X \mid A] = \int_{-\infty}^{\infty} x f_{X|A}(x)\,dx.$$

Expected Value Rule $$E[g(X)] = \sum_x g(x) p_X(x) \longrightarrow \mathbf{E}\big[g(X) \mid A\big] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)\,dx,$$

Conditioning a Continuous Random Variable on an Event

- Special Case: Conditioning on an event of the form $\{X \epsilon A\}$



Conditioning a Continuous Random Variable on an Event

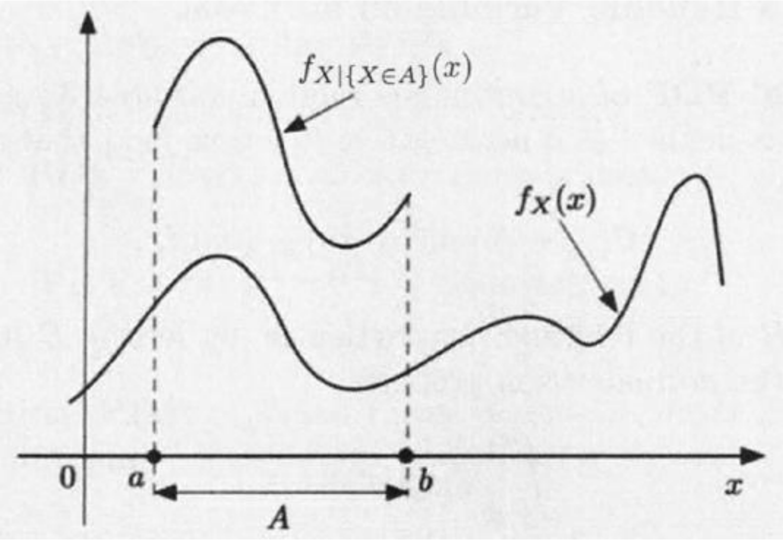
- Special Case: Conditioning on an event of the form $\{X \epsilon A\}$

In the important special case where we condition on an event of the form $\{X \in A\}$, with $\mathbf{P}(X \in A) > 0$, the definition of conditional probabilities yields

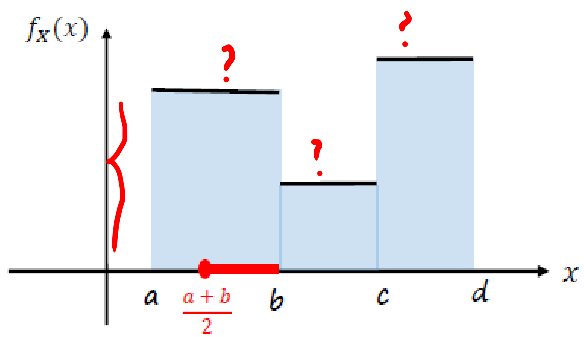
$$\mathbf{P}(X \in B \mid X \in A) = \frac{\mathbf{P}(X \in B,\ X \in A)}{\mathbf{P}(X \in A)} = \frac{\int_{A \cap B} f_X(x)\,dx}{\mathbf{P}(X \in A)}.$$

By comparing with the earlier formula

$$f_{X|\{X \in A\}}(x) = \begin{cases} \boxed{\dfrac{f_X(x)}{\mathbf{P}(X \in A)}} & \text{if } x \in A, \\ 0, & \text{otherwise.} \end{cases}$$

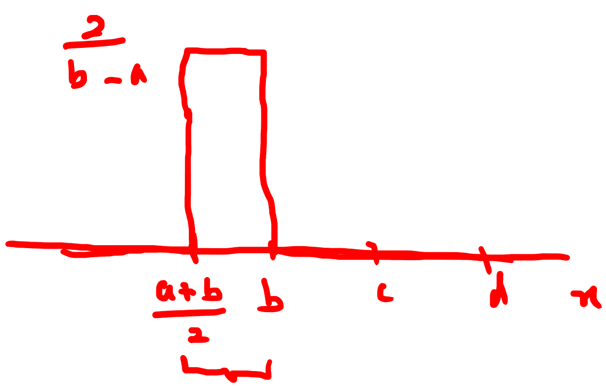


Conditioning a Random Variable on an Event: Example



$$A: \left\{ \frac{a+b}{2} \le X \le b \right\}$$

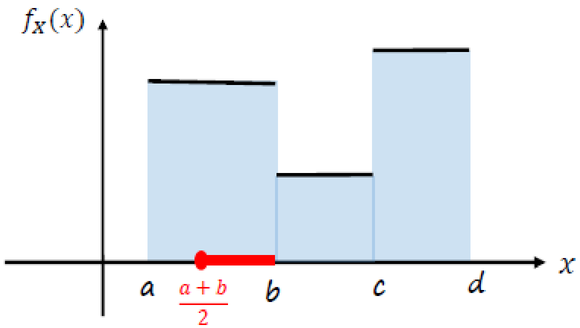
$f_{X|A}(x) = ?$



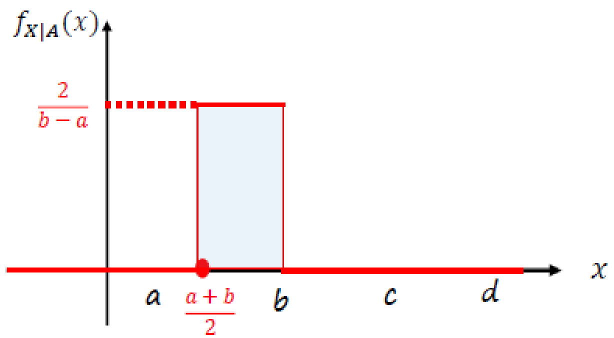
$$b - \frac{a+b}{2}$$

$$\frac{2b - a - b}{2} = \frac{b-a}{2}$$

Conditioning a Random Variable on an Event: Example



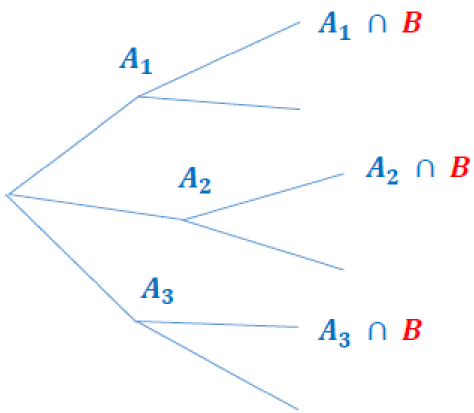
$$A: \quad \frac{a+b}{2} \leq X \leq b$$



Commute Time (T) : $f_T(t)$ $f_{T|Bike}(t)$ $f_{T|Car}(t)$ $f_{T|Van}(t)$

$E[T]$ $E[T|Bike]$ $E[T|Car]$ $E[T|Van]$

Total Expectation Theorem



$P(B) = P(A_1)\ P(B \mid A_1) + \ldots + P(A_n)\ P(B \mid A_n)$

$p_X(x) = P(A_1)\ p_{X|A_1}(x) + \ldots + P(A_n)\ p_{X|A_n}(x)$

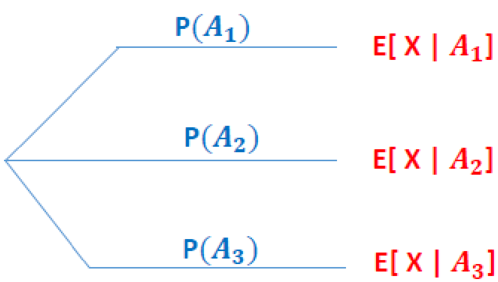
$F_X(x) = P(X \le x) = P(A_1)\ P(X \le x \mid A_1) + \ldots$

$= P(A_1)\ F_{X|A_1}(x) + \ldots$

$$\boxed{f_x(x) = P(A_1) f_{X|A_1}(x) + \ \ldots + P(A_n) f_{X|A_n}(x)}$$

$f_T(t) = P(Bike)\ f_{T|Bike}(t) + P(Car)\ f_{T|Car}(t) + P(Van) f_{T|Van}(t)$

$\int x\ f_X(x)\ dx = P(A_1) \int f_{X|A_1}(x)\ dx + \ldots$



$$\boxed{E[X] = P(A_1)E[X|A_1] + \ \ldots \ + \ P(A_n)E[X|A_n]}$$

$E[T] = P(Bike)\ E[T|Bike] + P(Car)\ E[T|Car] + P(Van)\ E[T|Van]$

X : Time from now till Bill leaves for supermarket

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

Total Expectation Theorem: Example

$f_X(x) = ?$

Bill goes to the supermarket shortly, with probability 1/3, at a time uniformly distributed between 0 and 2 hours from now;
Or with probability 2/3, later in the day
At a time uniformly distributed between 6-8 hours from now

$E[X] = ?$

$E[X|A_1] = 1$

$E[X|A_2] = 7$

$N(0,1)$

$P(H) = \frac{1}{3}$ $\Updownarrow$ $P(A_1) = \frac{1}{3}$

$U[0,2]$

$f_{X|A_1} \sim unif\ [0,2]$

$P(T) = \frac{2}{3}$

$P(A_2) = \frac{2}{3}$

$f_{X|A_2} \sim U[6,8]$

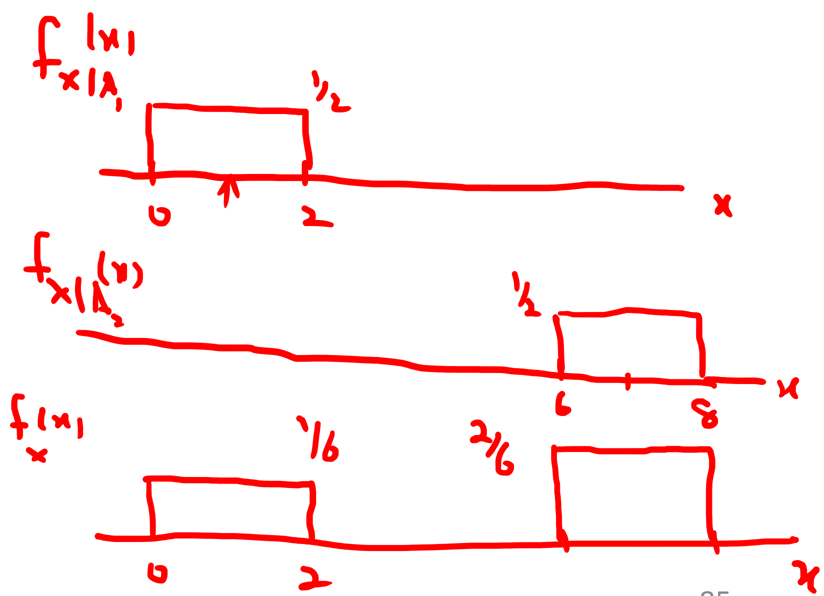
Total Prob Theorem

$f_X(x) = P(A_1) f_{X|A_1}(x) + P(A_2) f_{X|A_2}(x)$

$= \frac{1}{3} \cdot f_{X|A_1}(x) + \frac{2}{3} f_{X|A_2}(x)$

$E[X] = P(A_1) E[X|A_1] + P(A_2) E[X|A_2]$

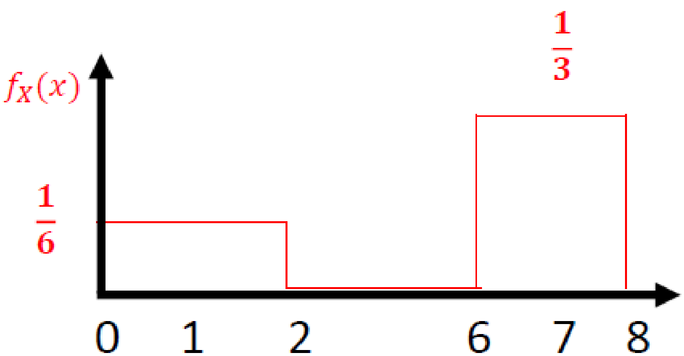
$= \frac{1}{3}(1) + \frac{2}{3}(7) =$



Total Expectation Theorem: Example

Bill goes to the supermarket shortly, with probability 1/3, at a time uniformly distributed between 0 and 2 hours from now;
Or with probability 2/3, later in the day
At a time uniformly distributed between 6-8 hours from now

$P(A_1) = \frac{1}{3}$ $\quad f_{X|A_1} \sim$ unif $[0,2]$ $\quad P(A_2) = \frac{2}{3}$ $\quad f_{X|A_2} \sim U\,[6,8]$



$$f_X(x) = P(A_1) f_{X|A_1}(x) + \ldots + P(A_n) f_{X|A_n}(x)$$

$$E[X] = P(A_1) E[X|A_1] + \ldots + P(A_n) E[X|A_1]$$

$$\frac{1}{3} * 1 + \frac{2}{3} * 7$$

Remember: Conditioning a Random Variable on another Random Variable

- Conditional PMF of X given Y: ($p_{X|Y}$)
  - Let X and Y be two random variables associated with the same experiment. The conditional PMF of X given Y is defined by:

$$p_{X\,|\,Y}(x\,|\,y) = \mathbf{P}(X = x\,|\,Y = y).$$

Using the definition of conditional probabilities, we have

$$p_{X\,|\,Y}(x\,|\,y) = \frac{\mathbf{P}(X = x, Y = y)}{\mathbf{P}(Y = y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}.$$

defined for y such that $p_Y(y) > 0$

Similar to
$P(B|A)$
with $B = \{X = x\}$
$A = \{Y = y\}$

  - Similar to the unconditional case

$$\sum_x p_{X|Y}(x|y) = 1$$

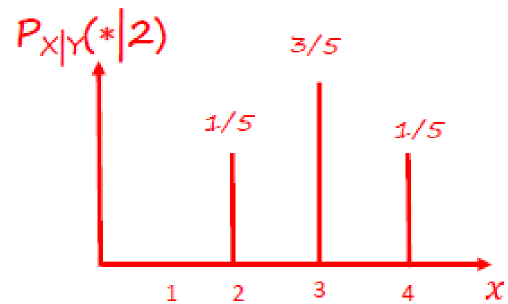
Y = Second roll				
4	1/20	2/20	2/20	
3	2/20	4/20	1/20	2/20
2		1/20	3/20	1/20
1		1/20		
	1	2	3	4

X = First roll

$P_Y(2) = 5/20$

$P_{X|Y}(1|2) = 0$

$P_{X|Y}(2|2) =$



Conditioning a Continuous RV on another RV

- <u>Conditional PDF</u>

Let X and Y be continuous random variables with joint PDF $f_{X,Y}$. For any y with $f_Y(y) > 0$, the **conditional PDF** of X given that $Y = y$, is defined by

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Conditioning a Continuous RV on another RV

- Conditional PDF

Let X and Y be continuous random variables with joint PDF $f_{X,Y}$. For any y with $f_Y(y) > 0$, the **conditional PDF** of X given that $Y = y$, is defined by

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

- Interpretation:

$$\begin{aligned}
\mathbf{P}(x \le X \le x + \delta_1 \mid y \le Y \le y + \delta_2) &= \frac{\mathbf{P}(x \le X \le x + \delta_1 \text{ and } y \le Y \le y + \delta_2)}{\mathbf{P}(y \le Y \le y + \delta_2)} \\
&\approx \frac{f_{X,Y}(x, y)\delta_1\delta_2}{f_Y(y)\delta_2} \\
&= f_{X|Y}(x \mid y)\delta_1.
\end{aligned}$$

In words, $f_{X|Y}(x \mid y)\delta_1$ provides us with the probability that X belongs to a small interval $[x, x + \delta_1]$, given that Y belongs to a small interval $[y, y + \delta_2]$. Since $f_{X|Y}(x \mid y)\delta_1$ does not depend on δ_2, we can think of the limiting case where δ_2 decreases to zero and write

$$\mathbf{P}(x \le X \le x + \delta_1 \mid Y = y) \approx f_{X|Y}(x \mid y)\delta_1, \qquad (\delta_1 \text{ small}),$$

view the conditional PDF $f_{X|Y}(x \mid y)$ (as a function of x) as a description of the probability law of X, given that the event $\{Y = y\}$ has occurred.

$$\mathbf{P}(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y)\, dx.$$

Conditioning a Continuous RV on another RV

- Conditional PDF

$$\boxed{f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.}$$

- Interpretation:

view the conditional PDF $f_{X|Y}(x \mid y)$ (as a function of x) as a description of the probability law of X, given that the event $\{Y = y\}$ has occurred.

$$\mathbf{P}(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y)\, dx.$$

$f_{W|H}(w|62)$ $\quad$ $f_{W|H}(w|72) = \dfrac{f_{W,H}(w,72)}{f_H(72)}$ $\quad$ $f_{H|W}(h|100kg) =$

Conditioning a Continuous RV on another RV

- <u>Conditional PDF</u>

Conditional PDF

$$\boxed{f_{X|Y}(x\,|\,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.}$$

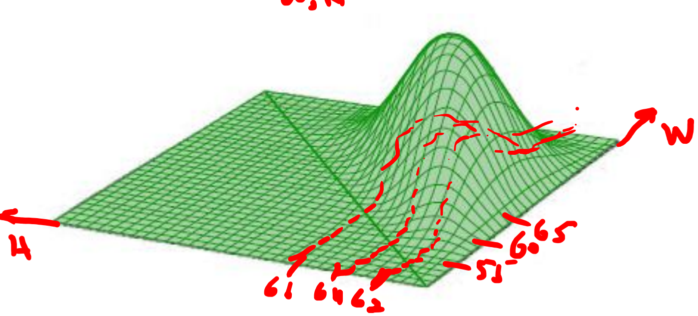
$\dfrac{\text{Joint PDF}}{\text{Marginal PDF}}$

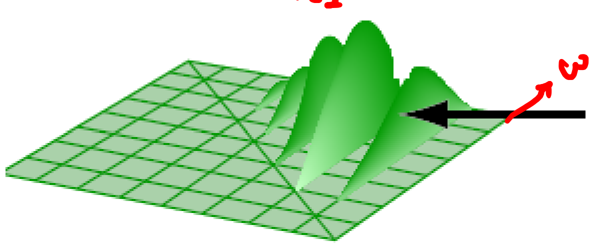
- Interpretation:

view the conditional PDF $f_{X|Y}(x\,|\,y)$ (as a function of x) as a description of the probability law of X, given that the event $\{Y = y\}$ has occurred.

$$\mathbf{P}(X \in A\,|\,Y = y) = \int_A f_{X|Y}(x\,|\,y)\,dx.$$

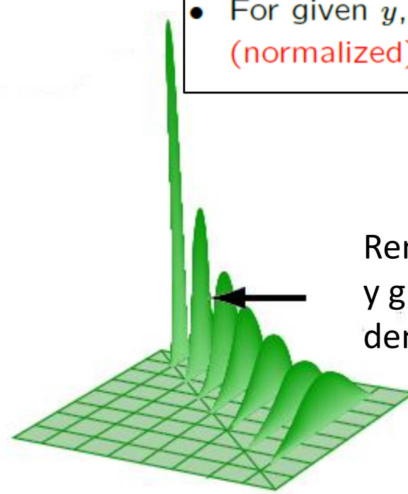
$f_{W,H}(w,h)$





Slice through density surface for fixed

- For given y, conditional PDF is a (normalized) "section" of the joint PDF



Renormalizing slices for fixed y gives the conditional densities for X given Y=y.

Joint PDF from Conditional PDF

- The conditional PDF is often convenient for calculation of the joint PDF:

$$\boxed{\begin{aligned} f_{X,Y}(x,y) &= f_{X|Y}(x|y)f_Y(y) \\ f_{x,y}(x,y) &= f_{Y|X}(y|x)f_X(x) \end{aligned}}$$

1. $f_{x|y}(x|y) = \dfrac{f_{x,y}(x,y)}{f_y(y)}$

2. $f_{y|x}(y|x) = \dfrac{f_{x,y}(x,y)}{f_x(x)}$

Conditioning a Continuous RV on another RV

- <u>Conditional PDF</u>

Let X and Y be continuous random variables with joint PDF $f_{X,Y}$. For any y with $f_Y(y) > 0$, the **conditional PDF** of X given that $Y = y$, is defined by

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

$E[W \mid H = 72]$

- <u>Conditional Expectation</u>

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx. \qquad \mathbf{E}[X \mid A] = \int_{-\infty}^{\infty} x f_{X|A}(x)\, dx. \qquad \boxed{\mathbf{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\, dx.}$$

Expected Value Rule

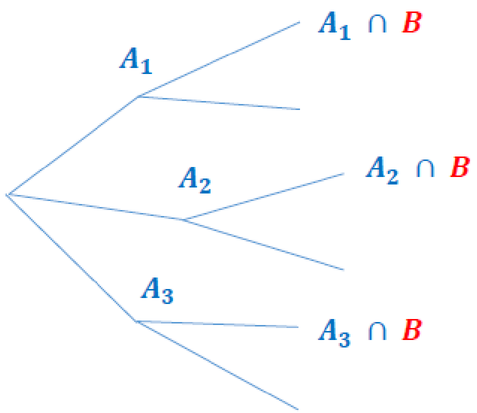
$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx \qquad \mathbf{E}[g(X) \mid A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)\, dx, \qquad \boxed{\mathbf{E}[g(X) \mid Y = y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x \mid y)\, dx.}$$

$A_1 = \{Y = y_1\}$

$A_2 = \{Y = y_2\}$

$A_3 = \{Y = y_3\}$

Remember: Total Expectation Theorem



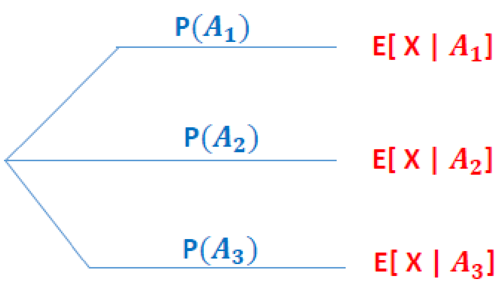
$P(B) = P(A_1)\, P(B \mid A_1) + \dots + P(A_n)\, P(B \mid A_n)$

$p_X(x) = P(A_1)\, p_{X|A_1}(x) + \dots + P(A_n)\, p_{X|A_n}(x)$

$F_X(x) = P(X \le x) = P(A_1)\, P(X \le x \mid A_1) + \dots$

$= P(A_1)\, F_{X|A_1}(x) + \dots$

$$\boxed{f_X(x) = P(A_1) f_{X|A_1}(x) + \dots + P(A_n) f_{X|A_n}(x)}$$



$\int x\, f_X(x)\, dx = P(A_1) \int f_{X|A_1}(x)\, dx + \dots$

$$\boxed{E[X] = P(A_1)E[X|A_1] + \dots + P(A_n)E[X|A_n]}$$

$E[X] = P(Y=y_1)E[X|Y=y_1] + P(Y=y_2)E[X|Y=y_2] + P(Y=y_3)E[X|Y=y_3]$

$$\boxed{E[X] = \int_{-\infty}^{\infty} f_Y(y)\, E[X|Y=y]\, dy}$$

$= \sum_{i=1}^{3} P(Y=y_i)\, E[X|Y=y_i]$

Now: Total Expectation Theorem

$$\boxed{E[\,X\,] = \mathrm{P}(\mathrm{A}_1)\,E[\,X \mid A_1\,] + \ .\ .\ . + \mathrm{P}(\mathrm{A}_n)\,E[\,X \mid An\,]}$$

Total Expectation Theorem

$$\boxed{\mathbf{E}[X] = \int_{-\infty}^{\infty} \mathbf{E}[X \mid Y = y] f_Y(y)\, dy.}$$

$$E[H] = \int_{-\infty}^{\infty} E[H \mid W = w] \cdot f_W(w) \cdot dw$$

$$E[W] = \int_{-\infty}^{\infty} E[W \mid H = h]\, f_H(h) \cdot dh$$

Now: Total Expectation Theorem

$$E[\,X\,] = P(A_1)\,E[\,X \mid A_1\,] + \,...\, + P(A_n)\,E[\,X \mid An\,]$$

Total Expectation Theorem

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} \mathbf{E}[X \mid Y = y] f_Y(y)\,dy.$$

Proof

$$\mathbf{E}[\,X\,] = \int_{-\infty}^{\infty} f_Y(y)\,\mathbf{E}[\,X \mid Y = y\,]\,\mathrm{d}y \quad = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\,dy$$

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

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

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx \quad = \quad E[\,X\,]$$

Remember: Independence

- Definition:

If

$$\mathbf{P}(A \mid B) = \mathbf{P}(A)$$

We say that A is independent of B

- Intuition
  - Remember that conditional probability encodes our "updated belief" about the likelihood of an event **IF** we have "additional partial information" about the results of the random experiment.
  - If A is independent of B, occurrence of B provides no new information about occurrence of A.
  - If A is independent of B, our beliefs are not "updated" about the likelihood of A if we have the additional information that B has occurred.

Remember: Independence

- An Equivalent Definition

If

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$$

We say that A is independent of B

- Verification

$P(A \cap B) = P(A) \cdot P(B)$

Independence of Two Continuous RVs

Let X and Y be jointly continuous random variables.

- X and Y are **independent** if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y), \qquad \text{for all } x, y.$$

$f_{W,H}(70, 66) = f_W(70) \cdot f_H(66)$

$f_{W,G}(70, 3.5) = f_W(70) f_G(3.5)$

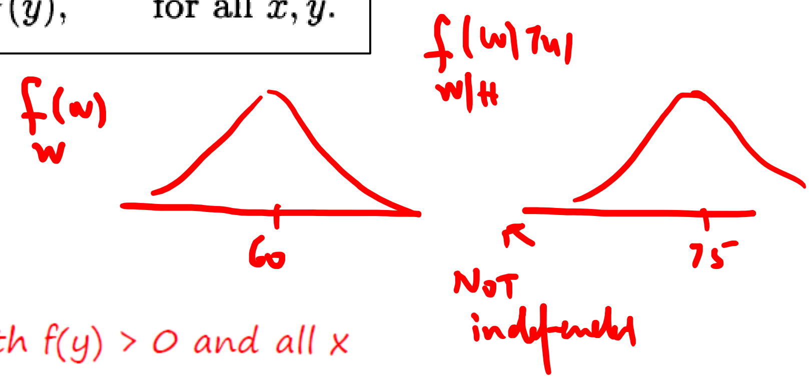
$P(A|B) = P(A)$

Independence of Two Continuous RVs

Let X and Y be jointly continuous random variables.

- X and Y are **independent** if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y), \qquad \text{for all } x, y.$$



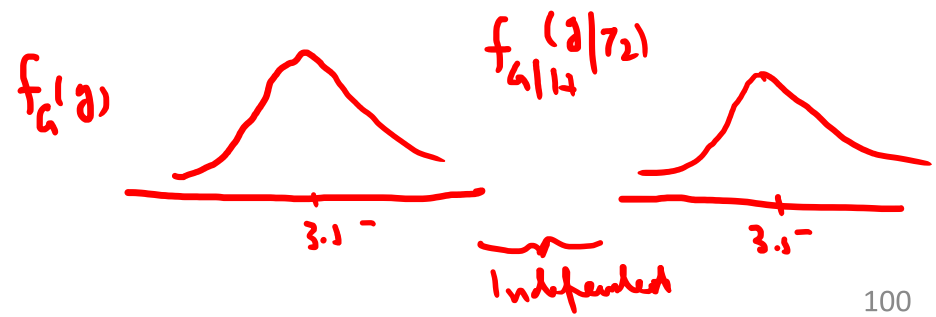
- <u>Equivalent Definitions:</u>

➡ $f_X(x) = f_{X|Y}(x|y)$ *for all y with f(y) > 0 and all x*

➡ $f_Y(y) = f_{Y|X}(y|x)$ *for all x with f(x) > 0 and all y*

$f_{W|H}(w|h) \neq f_W(w)$

$f_{G|H}(g|h) = f_G(g)$



Independence of Two Continuous RVs: "Implications"

If X and Y are independent, then any two events of the form $\{X \in A\}$ and $\{Y \in B\}$ are independent. Indeed,

H, G are independent

$$\begin{aligned}
\mathbf{P}(X \in A \text{ and } Y \in B) &= \int_{x \in A} \int_{y \in B} f_{X,Y}(x,y)\,dy\,dx \\
&= \int_{x \in A} \int_{y \in B} f_X(x) f_Y(y)\,dy\,dx \\
&= \int_{x \in A} f_X(x)\,dx \int_{y \in B} f_Y(y)\,dy \\
&= \mathbf{P}(X \in A)\,\mathbf{P}(Y \in B).
\end{aligned}$$

$P(X \in A \text{ and } Y \in B)$

$$P(68 \le H \le 72,\ 3.2 \le G \le 3.5) = P(68 \le H \le 72)\,P(3.2 \le G \le 3.5)$$

In particular, independence implies that

$$\boxed{F_{X.Y}(x.y)} = \mathbf{P}(X \le x,\ Y \le y) = \mathbf{P}(X \le x)\,\mathbf{P}(Y \le y) = \boxed{F_X(x) F_Y(y)}.$$

Independence and Expectation

If X, Y are **independent**: $\mathbf{E}[XY] = \mathbf{E}[X]\ \mathbf{E}[Y]$

g(X) and h(Y) are also independent: $\mathbf{E}[g(X)\ h(Y)] = \mathbf{E}[g(X)]\ \mathbf{E}[h(Y)]$

Independence and Variance

- Always true: $var(aX + b) = a^2 var(X)$ $var(X + a) = var(X)$

- In general: $var(X + Y) \neq var(X) + var(Y)$

If X, Y are **independent**: $var(X + Y) = var(X) + var(Y)$

Independence of Two Continuous RVs

Example:

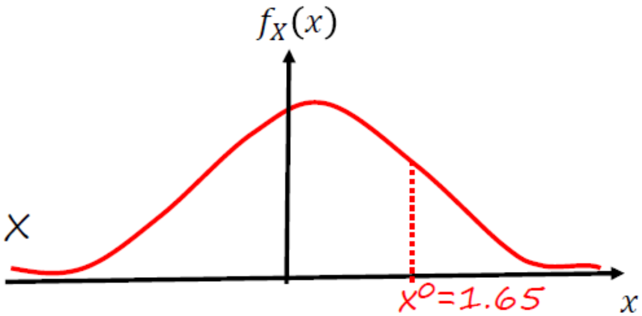
- Independence in Normal RV's X and Y:

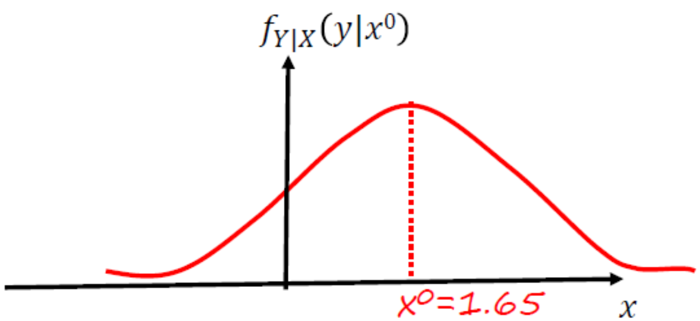
  Lets form a RV in a two step process

  mean var

  1. $X \sim N(0, 1)$;

pick an X





2. $Y \sim \underline{N(X, 1)}$

are X and Y independent?

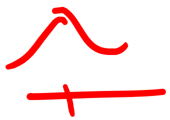
No

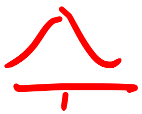
G & H are independent:

Independent Normal RVs

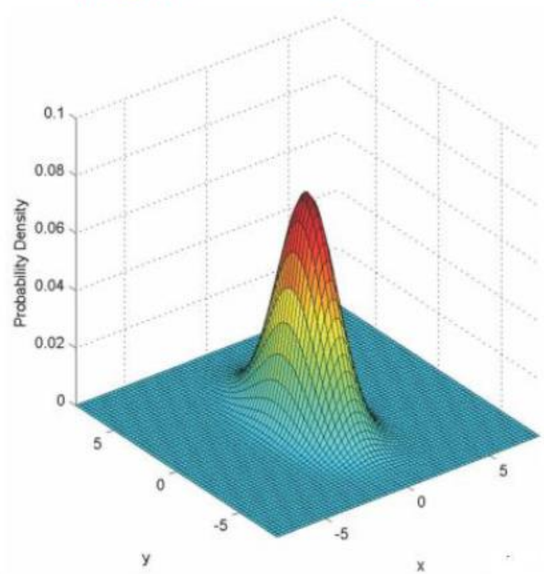
Example 3.18. Independent Normal Random Variables. Let X and Y be independent normal random variables with means μ_x, μ_y, and variances σ_x^2, σ_y^2, respectively. Their joint PDF is of the form

$$\boxed{f_{X,Y}(x,y) = f_X(x)f_Y(y)} = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right\}.$$

$f_G(g)$ 

$f_H(t)$ 

$\mu_x = \mu_y = 0;\ \sigma^2_x = \sigma^2_y = 4;$



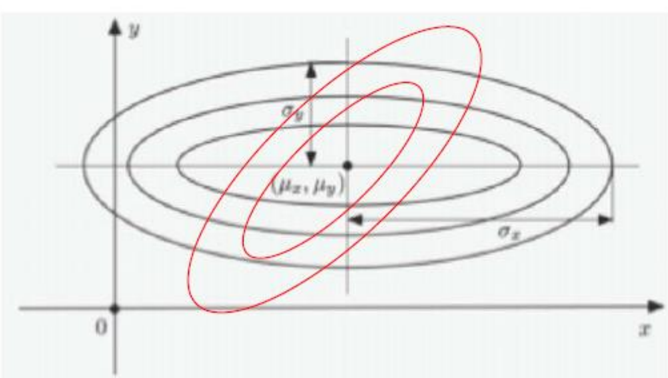
Independent Normal RVs

Example 3.18. Independent Normal Random Variables. Let X and Y be independent normal random variables with means μ_x, μ_y, and variances σ_x^2, σ_y^2, respectively. Their joint PDF is of the form

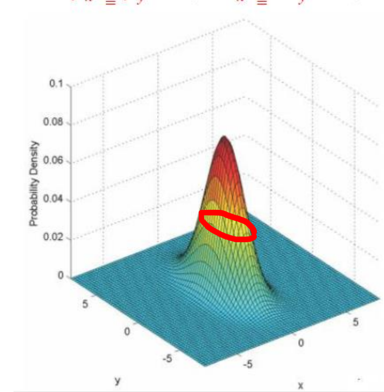
$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right\}.$$

This joint PDF has the shape of a bell centered at (μ_x, μ_y), and whose width in the x and y directions is proportional to σ_x and σ_y, respectively. We can get some additional insight into the form of this PDF by considering its contours, i.e., sets of points at which the PDF takes a constant value. These contours are described by an equation of the form

$$\frac{(x-\mu_x)^2}{\sigma_X^2} + \frac{(y-\mu_y)^2}{\sigma_Y^2} = \text{constant}.$$



$\mu_x = \mu_y = 0$; $\sigma^2_x = \sigma^2_y = 4$;

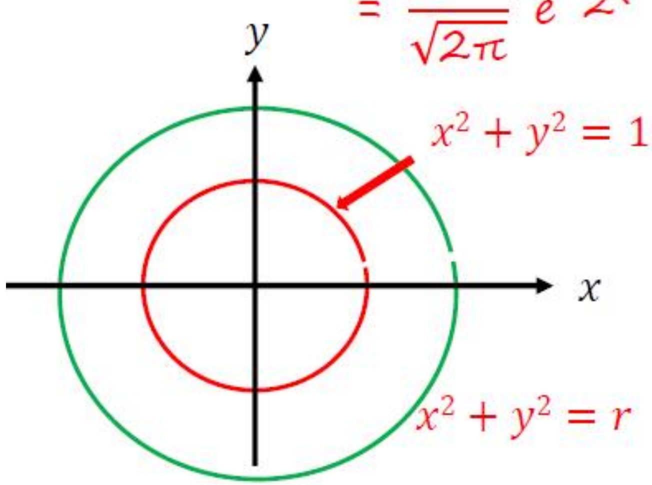


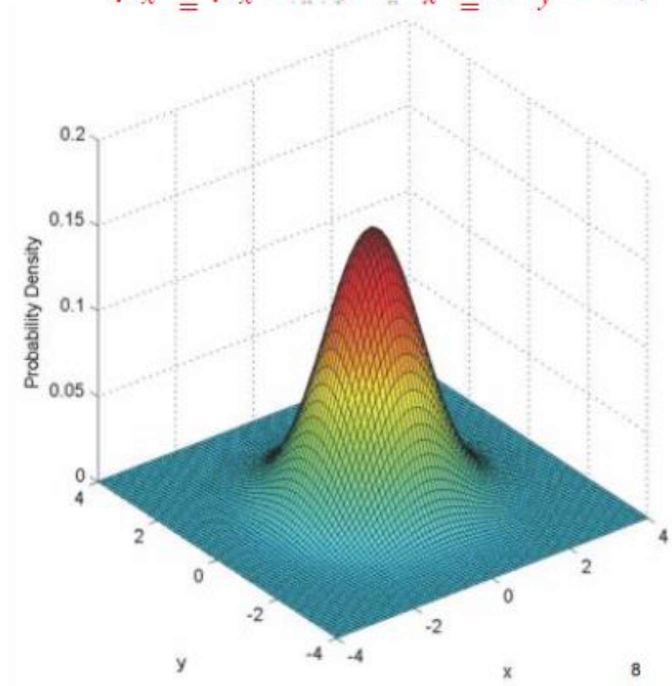
Independent Standard Normal RVs

Now let's assume X and Y independent on $\sim N(0, 1)$

$$
\begin{aligned}
f_{X,Y}(x,y) &= f_X(x)\ f_Y(y) \\
&= \frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ \frac{1}{\sqrt{2\pi}}\ e^{-\frac{y^2}{2}} \\
&= \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}(x^2+y^2)}
\end{aligned}
$$

$\mu_x = \mu_x = 0;\ \ \sigma^2{}_x = \sigma^2{}_y = 1;$

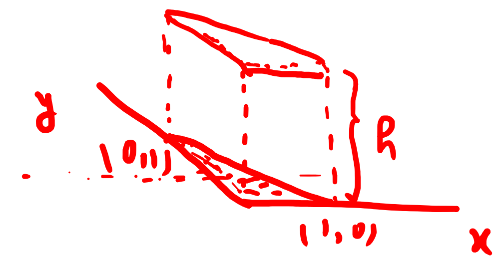




$f_Y(0.5)$

Volume under the Joint PDF = 1

$\frac{1}{2} \times 2 = 1$



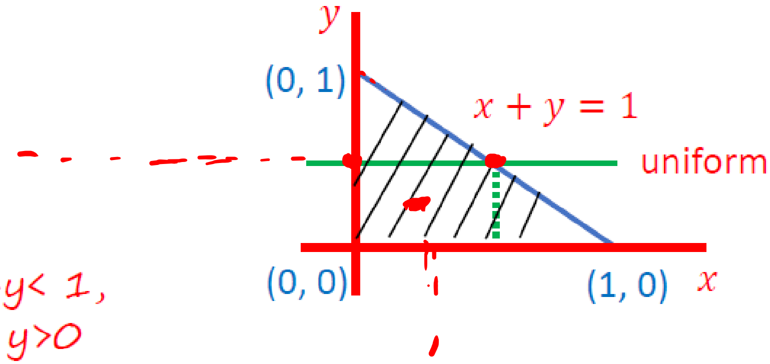
Comprehensive Example

Joint PDF :-

Uniform on a set S : [a Triangle]

Uniform probabilities on a triangle

find:

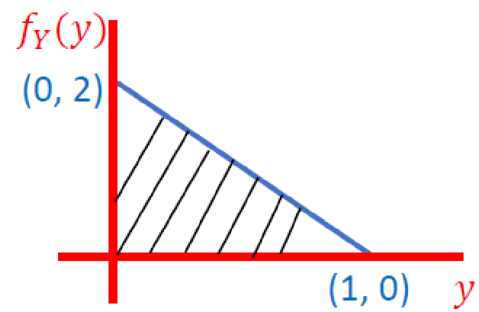


a) $f_{X,Y}(x, y)$

$f_{X,Y}(0.5, 0.5)$

$$f_{X,Y}(x, y) = \begin{cases} 2 & x+y< 1, \\ & x, y>0 \\ 0, & \text{otherwise} \end{cases}$$

b) $f_Y(y)$



$$f_Y(y) = \int f_{X,Y}(x, y)\, dx$$

$$= \begin{cases} 2(1-y), & 0\le y\le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

$$= \int_{-\infty}^{0} 0 . dx + \int_{0}^{1-y} 2 . dx + \int_{1-y}^{\infty} 0 . dx =$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) . dy$$

Comprehensive Example

c) $f_{X|Y}(x|y)$

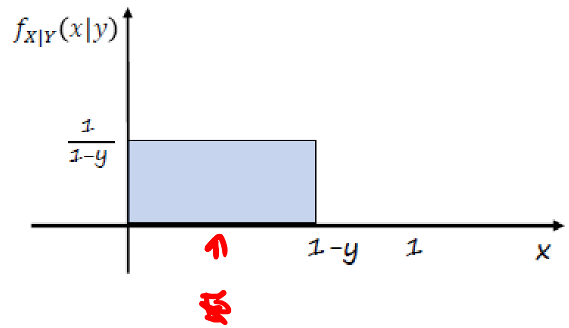
$$\boxed{f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$$

$$= \frac{2}{2(1-y)}$$

$$= \begin{cases} \frac{1}{(1-y)}, & 0 \le y \le 1;\ 0 \le x \le 1-y \\ 0, & \text{otherwise} \end{cases}$$

$0 \le y \le 1;$
$0 \le x, y < 1$

$E[X|Y=y]$



Comprehensive Example

$\boxed{\text{d) } E[X|Y=y]}$

$$E[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$$

$$E[X \mid Y=y] = \int_{0}^{1-y} x \frac{1}{(1-y)}\,dx$$

$$= \frac{1}{(1-y)} \frac{(1-y)^2}{2}$$

$$= \frac{(1-y)}{2} \qquad 0 \le y \le 1$$

Comprehensive Example

e) $E[X]$

$$E[X] = \int_{-\infty}^{\infty} f_Y(y)\, E\left[X \mid Y=y\right] dy$$

$$= \int_{-\infty}^{\infty} f_Y(y) \frac{(1-y)}{2} dy = \frac{1}{2}\int_{-\infty}^{\infty} f_Y(y) dy - \frac{1}{2}\int_{-\infty}^{\infty} y\, f_Y(y) dy$$

$$= \frac{1}{2} - \frac{1}{2}E[Y] = \frac{1}{2} - \frac{1}{2}E[X]$$

$$\left(1 + \frac{1}{2}\right) E[X] = \frac{1}{2}$$

$$E[X] = \frac{1}{3} \qquad E[Y] = \frac{1}{3}$$

Review

$$p_X(x) \qquad f_X(x)$$

$$p_{X,Y}(x,y) \qquad f_{X,Y}(x,y)$$

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} \qquad f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

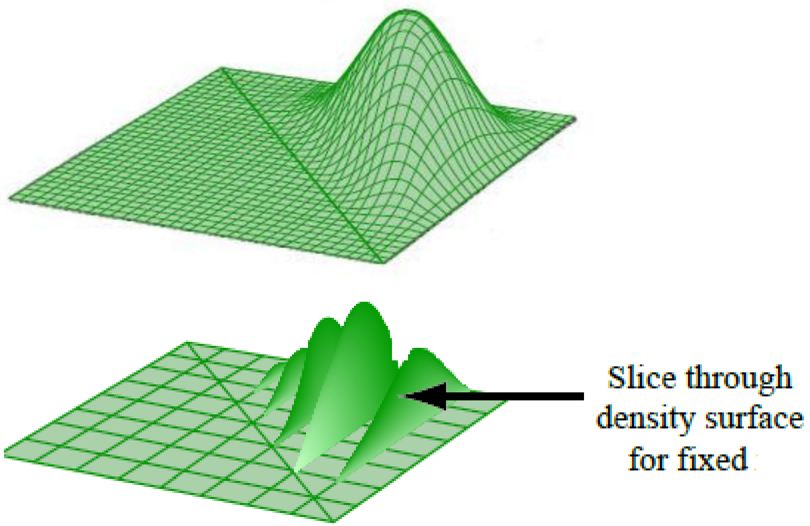
$$p_X(x) = \sum_y p_{X,Y}(x,y) \qquad f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$$F_X(x) = \mathrm{P}(X \leq x)$$

$$\mathrm{E}[X], \quad \mathrm{var}(X)$$

Review

Review



Remember: Bayes' Rule

Bayes' Rule

Let $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of the sample space, and assume that $\mathbf{P}(A_i) > 0$, for all i. Then, for any event B such that $\mathbf{P}(B) > 0$, we have

$$\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(B)}$$

$$= \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\underbrace{\mathbf{P}(A_1)\mathbf{P}(B \mid A_1)}_{P(A_1 \cap B)} + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)}.$$

$$\boxed{P(A_1/B) = \frac{P(A_1)P(B/A_1)}{P(B)}}$$

$$\boxed{P(A_2/B) = \frac{P(A_2)P(B/A_2)}{P(B)}}$$

$P(B) =$

- Verification

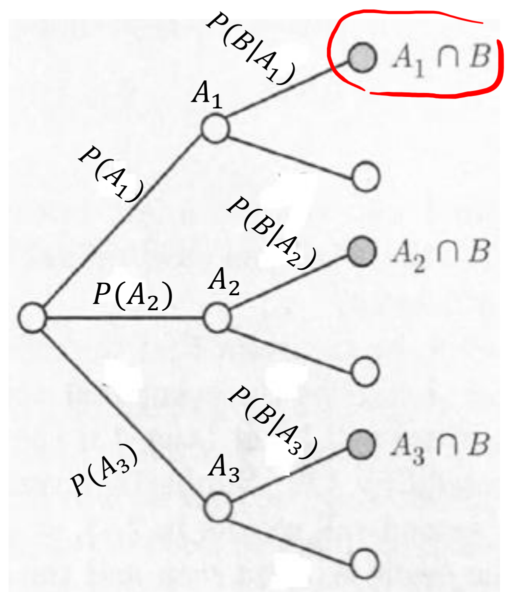
$$P(A_1/B) = \frac{P(A_1 \cap B)}{P(B)} \rightarrow 1$$

$$P(B/A_1) = \frac{P(A_1 \cap B)}{P(A_1)} \rightarrow 2$$

From

$$\Rightarrow P(A_1 \cap B) = P(A_1)\,P(B/A_1)$$

$$\boxed{P(A_1/B) = \frac{P(A_1)\,P(B/A_1)}{P(B)}}$$



Remember: Bayes' Rule

Cause → Effect

P(A₁/B)

P(B/A₁)

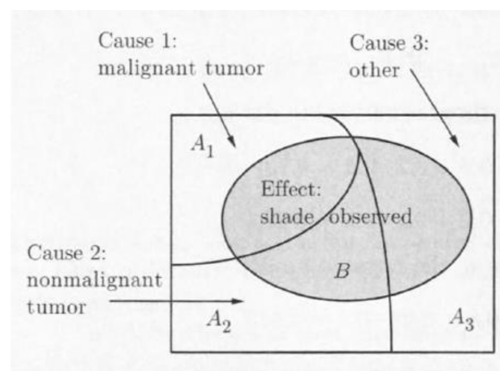
Bayes' Rule

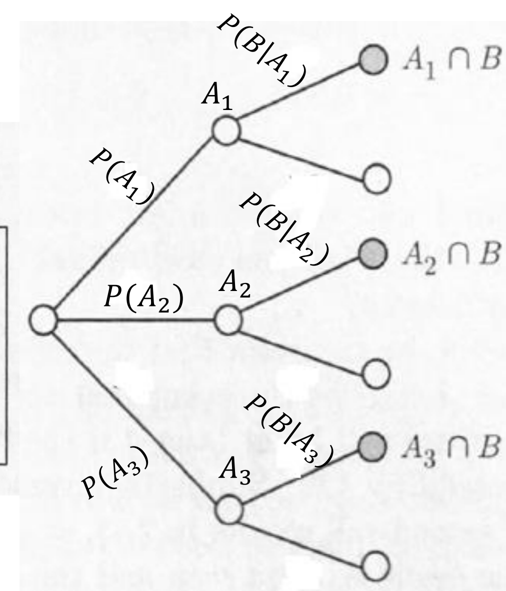
Let $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of the sample space, and assume that $\mathbf{P}(A_i) > 0$, for all i. Then, for any event B such that $\mathbf{P}(B) > 0$, we have

$$\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(A_1)\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)}.$$

- Inference Context
  - Consider a set of mutually-exclusive and collectively-exhaustive potential "causes"{A_1, A_2, A_3} that result in a certain "effect" B.
  - We observe the effect B and wish to infer the cause (out of A_1, A_2, A_3) that made B happen.
  - This is possible by evaluating following probabilities: "$\mathbf{P}(A_i \mid B)$"

    i.e conditional probabilities of A_1, A_2, A_3, given that B has occurred.





$P(A_1/B) =$

$P(A_2/B) =$

$P(A_3/B) =$

Remember: Bayes' Rule

3 CAUSES (A_1, A_2, A_3)

↓

EFFECT (B)

$P(B/A_1)$ –

$P(A_1/B)$ –

Bayes' Rule

Let $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of the sample space, and assume that $\mathbf{P}(A_i) > 0$, for all i. Then, for any event B such that $\mathbf{P}(B) > 0$, we have

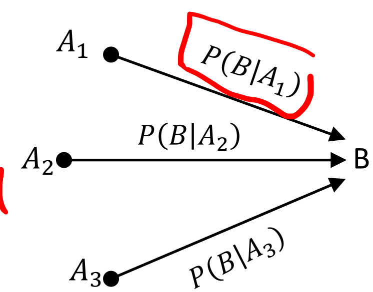
$$\begin{aligned}\mathbf{P}(A_i \mid B) &= \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(B)} \\ &= \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(A_1)\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)}.\end{aligned}$$

- Inference Context
  - Consider a set of mutually-exclusive and collectively-exhaustive potential "causes"$\{A_1, A_2, A_3\}$ that result in a certain "effect" B.
  - We observe the effect B and wish to infer the cause (out of A_1, A_2, A_3) that made B happen.
  - This is possible by evaluating following probabilities: $\mathbf{P}(A_i \mid B)$

    i.e conditional probabilities of A_1, A_2, A_3, given that B has occurred.

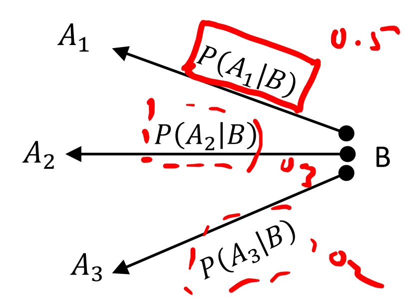
Given a cause, probability that effect will be observed

"Cause-Effect Modeling"



Given an effect has been observed, what is the prob. that it may caused by A_1.

"Inference"



$f_Y(y)$ $\qquad f_{Y|X}(y|10)$

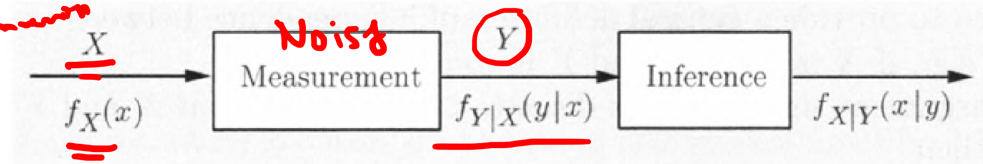
A More General Version of the Bayes' Rule

- Context

In many situations, we represent an unobserved phenomenon by a random variable X with PDF f_X and we make a noisy measurement Y, which is modeled in terms of a conditional PDF $f_{Y|X}$. Once the value of Y is measured, what information does it provide on the unknown value of X?

Unobserved Phenomenon

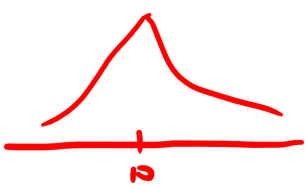
Noisy



10 A $\qquad$ 10A $\qquad$ Y = 11.5 A

10.1 A

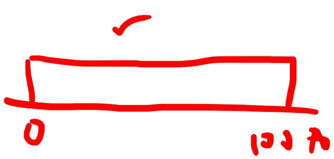
9.8 A

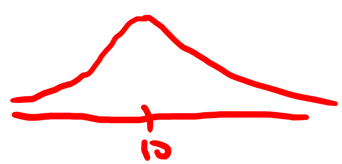


$f_{Y|X}(y|10A)$

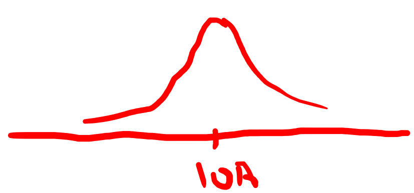
$f_{X|Y}(x|11.5)$

PKR 20,000





PKR 200,000

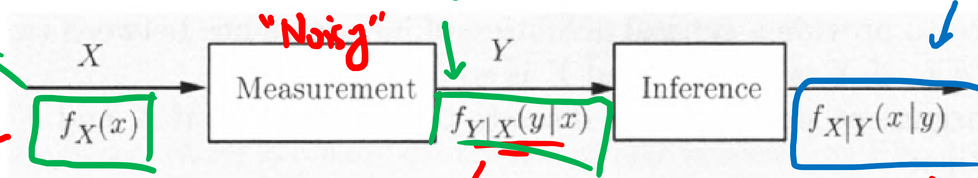


A More General Version of the Bayes' Rule

- <u>Context</u>

In many situations, we represent an unobserved phenomenon by a random variable X with PDF f_X and we make a noisy measurement Y, which is modeled in terms of a conditional PDF $f_{Y|X}$. Once the value of Y is measured, what information does it provide on the unknown value of X?

Known "Noisy" Known Unknown



"Prior Belief"

"Quality of Measurement"

"POSTERIOR BELIEF"

"What is the distribution of Y Given the state of the world is $X=x_0$"

"Given the measurement $Y=y_0$, what is the distribution of X."

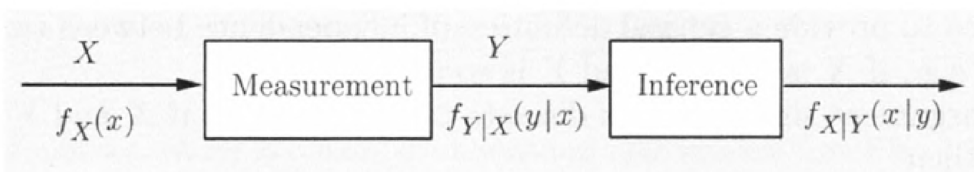
X is continuous, Y continuous	$f_X(x)$	$f_{Y\|X}(y\|x)$	$f_{X\|Y}(x\|y)$
X is discrete, Y is discrete	$p_X(x)$	$p_{Y\|X}(y\|x)$	$p_{X\|Y}(x\|y)$
X is discrete, Y is continuous	$p_X(x)$	$f_{Y\|X}(y\|x)$	$p_{X\|Y}(x\|y)$
X continuous, Y discrete	$f_X(x)$	$p_{Y\|X}(y\|x)$	$f_{X\|Y}(x\|y)$

$P(B|A) = \frac{P(B)\,P(A|B)}{P(A)}$

A More General Version of the Bayes' Rule

- <u>Context</u>

In many situations, we represent an unobserved phenomenon by a random variable X with PDF f_X and we make a noisy measurement Y, which is modeled in terms of a conditional PDF $f_{Y|X}$. Once the value of Y is measured, what information does it provide on the unknown value of X?



Note that whatever information is provided by the event $\{Y = y\}$ is captured by the conditional PDF $f_{X\|Y}(x\,\|\,y)$. It thus suffices to evaluate this PDF.

$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

$$\boxed{f_{X|Y}(x \mid y) = \frac{f_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}.}$$

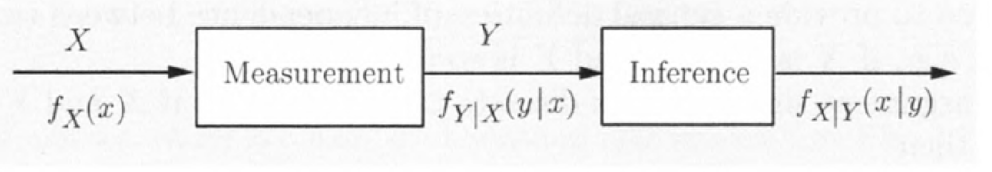
$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$

$f_{X,Y}(x,y) = f_X(x)\, f_{Y|X}(y|x)$

Variations of The Bayes' Rule

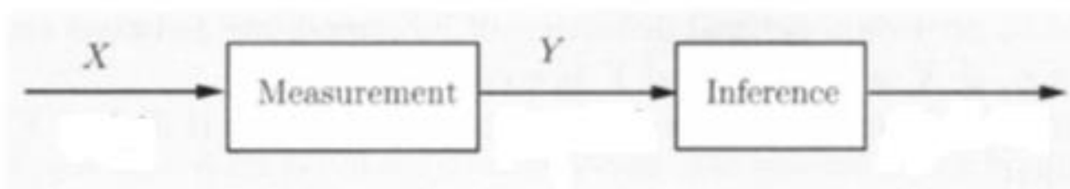
- <u>Context</u>

In many situations, we represent an unobserved phenomenon by a random variable X with PDF f_X and we make a noisy measurement Y, which is modeled in terms of a conditional PDF $f_{Y|X}$. Once the value of Y is measured, what information does it provide on the unknown value of X?



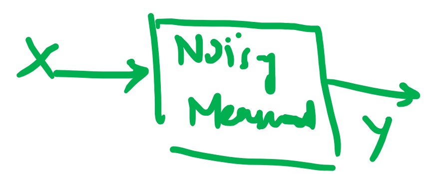
- Variations
  - Discrete X, Discrete Y
  - Continuous X, Continuous Y
  - Discrete X, Continuous Y
  - Continuous X, Discrete Y

Variations of The Bayes' Rule: Discrete X, Discrete Y



$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{p_X(x)p_{Y|X}(y \mid x)}{p_Y(y)}$$

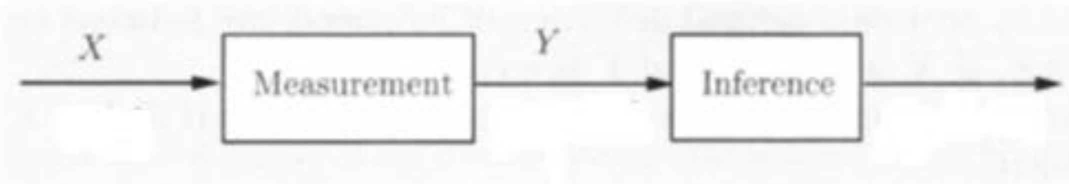
$$p_Y(y) = \sum_x p_X(x)p_{Y|X}(y \mid x)$$



Example:

- $X = 1, 0$: airplane present/not present
- $Y = 1, 0$: something did/did not register on radar

Variations of The Bayes' Rule: Continuous X, Continuous Y

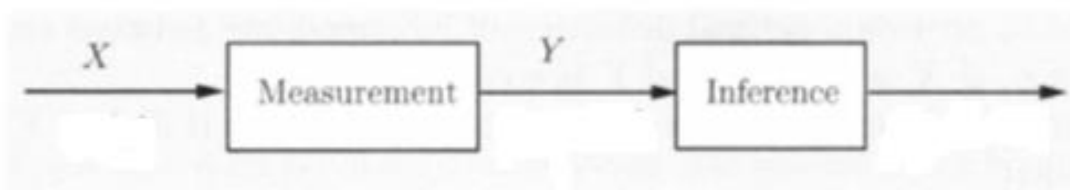


$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}$$

$$f_Y(y) = \int_x f_X(x) f_{Y|X}(y \mid x)\, dx$$

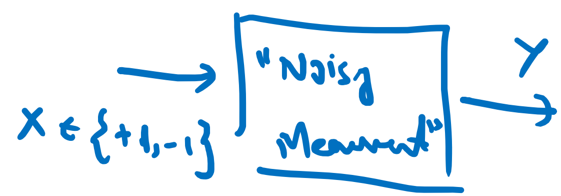
- Example:
  - X: Current through a resistor
  - Y: Analog measurement of the current through a resistor

Variations of The Bayes' Rule: Discrete X, Continuous Y



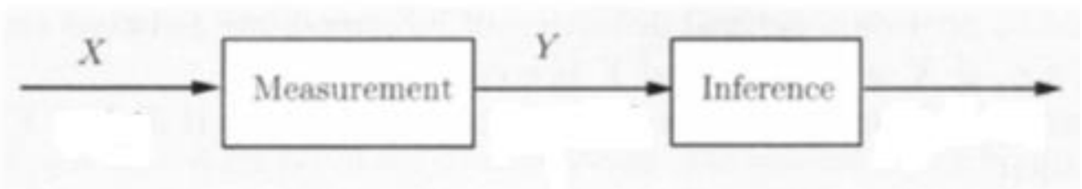
$$p_{X|Y}(x \mid y) = \frac{p_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}$$

$$f_Y(y) = \sum_x p_X(x) f_{Y|X}(y \mid x)$$



- Example:
  - X: A binary signal (+1 or -1) S is transmitted.
  - Y: The received signal is Y = S+N where N is normal noise with zero mean and unit variance.
  - Inference: What is the probability that S=1 as a function of the observed value of y of Y.

Variations of The Bayes' Rule: Continuous X, Discrete Y



$$f_{X|Y}(x \mid y) = \frac{f_X(x)p_{Y|X}(y \mid x)}{p_Y(y)}$$

$$p_Y(y) = \int_x f_X(x)p_{Y|X}(y \mid x)\,dx$$

- Example:
  - X: Intensity of the light beam
  - Y: The measured photon count

Bayes' Rule: Example

The incidence rate of a certain disease is 15/100000. There is a test for the disease which is 95% accurate(i.e. If a person has the disease, the test comes back positive with probability 0.95. If a person does not have the disease, it comes back negative with probability 0.95). Given that a person tested positive, what is the probability they have the disease.

$A_1 = \{\text{Has Disease}\}$ $\quad B = \{\text{Test +ve}\}$

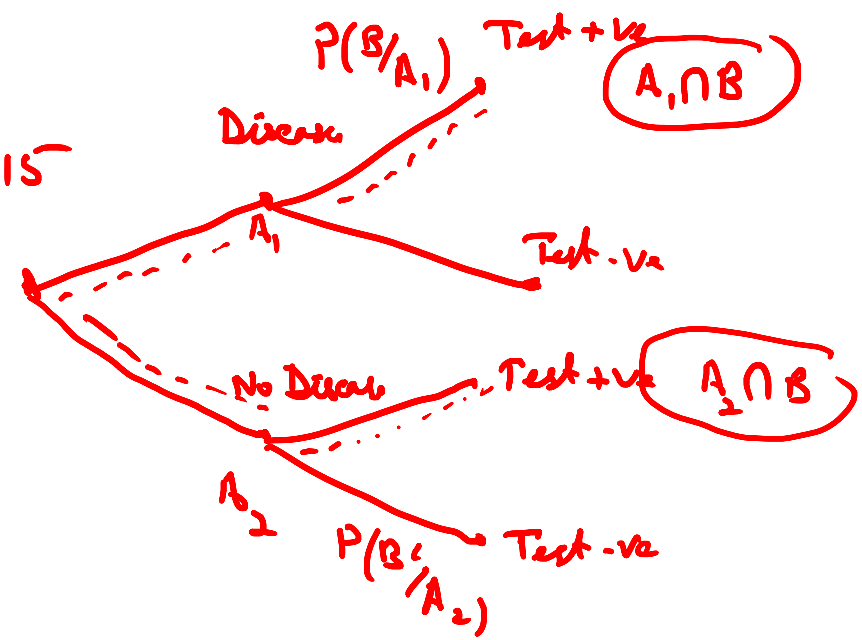
$A_2 = \{\text{Has No Disease}\}$ $\quad P(A_1) = 0.00015$

$P(B/A_1) = 0.95$

$P(A_1/B) = ?$ ← Inference Content

$$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{\boxed{P(B)}}$$

$=$



Bayes' Rule: Example (The new approach)

The incidence rate of a certain disease is 15/100000. There is a test for the disease which is 95% accurate(i.e. If a person has the disease, the test comes back positive with probability 0.95. If a person does not have the disease, it comes back negative with probability 0.95). Given that a person tested positive, (+ve) what is the probability they have the disease.

X: Somebody has the disease or not $X \in \{0, 1\}$

Y: Somebody test +ve or -ve $Y \in \{0, 1\}$

$X \rightarrow$ Noisy Measurement $\rightarrow y$

$p_X(x)$ $\quad$ $p_{Y|X}(y|x)$ $\quad$ $p_{X|Y}(x|y)$

$$p_X(x) = \begin{cases} 0.99985 & x = 0 \\ 0.00015 & x = 1 \end{cases}$$ Prior Belief

$$p_{Y|X}(y|0) = \begin{cases} 0.95 & y = 0 \\ 0.05 & y = 1 \end{cases}$$

Quality of Measurement

$$p_{Y|X}(y|1) = \begin{cases} 0.05 & y = 0 \\ 0.95 & y = 1 \end{cases}$$

Bayes' Rule:

$$p_{X|Y}(x|y) = \frac{p_X(x)\, p_{Y|X}(y|x)}{p_Y(y)}$$

$$p_{X|Y}(x|y) = \frac{p_X(x)\, p_{Y|X}(y|x)}{p_X(0) \cdot p_{Y|X}(y|0) + p_X(1)\, p_{Y|X}(y|1)}$$

$$p_{X|Y}(1|1) = \frac{p_X(1)\, p_{Y|X}(1|1)}{p_X(0)\, p_{Y|X}(1|1) + p_X(1)\, p_{Y|X}(1|1)}$$

posterior Belief

The Bayes' Rule Example: Discrete X, Continuous Y

Example 3.20. Signal Detection. A binary signal S is transmitted, and we are given that $\mathbf{P}(S = 1) = p$ and $\mathbf{P}(S = -1) = 1 - p$. The received signal is $Y = N + S$, where N is normal noise, with zero mean and unit variance, independent of S. What is the probability that $S = 1$, as a function of the observed value y of Y?

The Bayes' Rule Example: Discrete X, Continuous Y

Example 3.20. Signal Detection. A binary signal S is transmitted, and we are given that $\mathbf{P}(S = 1) = p$ and $\mathbf{P}(S = -1) = 1 - p$. The received signal is $Y = N + S$, where N is normal noise, with zero mean and unit variance, independent of S. What is the probability that $S = 1$, as a function of the observed value y of Y?

$X = S = \{1, -1\}$ $\boxed{\text{Comm. Channel}} \rightarrow Y = S + N$ $N \sim N(0,1)$

$X \rightarrow \boxed{\text{Measurement Process}} \rightarrow Y$

$p_X(x)$ $f_{Y|X}(y|x)$ $p_{X|Y}(x|y)$

$$p_X(x) = \begin{cases} p & x = +1 \\ 1 - p & x = -1 \end{cases}$$

$p = 0.5$

Q: "What is the prob. that $S=1$", given the received signal is 0.55?

The Bayes' Rule Example: Discrete X, Continuous Y

Example 3.20. Signal Detection. A binary signal S is transmitted, and we are given that $\mathbf{P}(S=1) = p$ and $\mathbf{P}(S=-1) = 1-p$. The received signal is $Y = N + S$, where N is normal noise, with zero mean and unit variance, independent of S. What is the probability that $S = 1$, as a function of the observed value y of Y?

$X = S = \{1,-1\}$ $\xrightarrow{}$ Comm. Channel $\xrightarrow{}$ $Y = S+N$ $\quad N \sim N(0,1)$

$X \rightarrow$ Measurement Process $\rightarrow Y$

$p_X(x) \quad f_{Y|X}(y|x) \quad p_{X|Y}(x|y)$

$$p_X(x) = \begin{cases} p & x=+1 \\ 1-p & x=-1 \end{cases}$$ Prior Belief

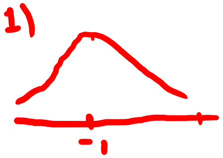
$$f_{Y|X}(y|x) = \begin{cases} f_{Y|X}(y|1) \\ f_{Y|X}(y|-1) \end{cases}$$

Q. of Measurement

$f_{Y|X}(y|1) = N(1,1)$

$\rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}$

$f_{Y|X}(y|-1) = N(-1,1)$



$\rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}}$

$$\boxed{p_{X|Y}(x|y) = \frac{p_X(x)\, f_{Y|X}(y|x)}{p_X(1) f_{Y|X}(y|1) + p_X(-1) f_{Y|X}(y|-1)}}$$

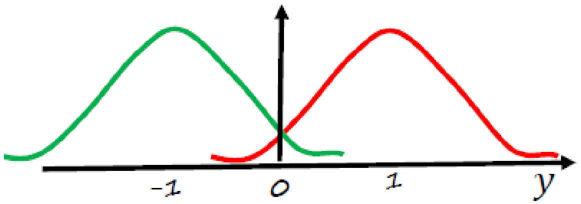
$$p_{X|Y}(1|0.55) = \frac{p_X(1)\, f_{Y|X}(0.55|1)}{p_X(1) f_{Y|X}(0.55|1) + p_X(-1) f_{Y|X}(0.55|-1)}$$

The Bayes' Rule Example: Discrete X, Continuous Y

Example 3.20. Signal Detection. A binary signal S is transmitted, and we are given that $\mathbf{P}(S=1)=p$ and $\mathbf{P}(S=-1)=1-p$. The received signal is $Y=N+S$, where N is normal noise, with zero mean and unit variance, independent of S. What is the probability that $S=1$, as a function of the observed value y of Y?

$$p_{X|Y}(x \mid y) = \frac{p_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}$$

$$f_Y(y) = \sum_x p_X(x) f_{Y|X}(y \mid x)$$



Conditioned on $S=s$, the random variable Y has a normal distribution with mean s and unit variance.

$$\mathbf{P}(S=1 \mid Y=y) = \frac{p_S(1) f_{Y|S}(y \mid 1)}{f_Y(y)} = \frac{\frac{p}{\sqrt{2\pi}} e^{-(y-1)^2/2}}{\frac{p}{\sqrt{2\pi}} e^{-(y-1)^2/2} + \frac{1-p}{\sqrt{2\pi}} e^{-(y+1)^2/2}},$$

which simplifies to

$$\mathbf{P}(S=1 \mid Y=y) = \frac{pe^y}{pe^y + (1-p)e^{-y}}.$$

Note that the probability $\mathbf{P}(S=1 \mid Y=y)$ goes to zero as y decreases to $-\infty$, goes to 1 as y increases to ∞, and is monotonically increasing in between, which is consistent with intuition.

Statistics: Descriptive vs Inferential

- Statistics
  - Area of applied math that deals with collection, organization, analysis, interpretation, and presentation of data

- Descriptive Statistics
  - Focuses on describing the visible characteristics of a dataset
  - Examples of characteristics: Distribution, central tendency, variability

- **Inferential Statistics (Statistical Inference**)
  - Focuses on extracting information about the unknown parameters of a model from available data
  - Focuses on making predictions or generalizations from a dataset

n, p

Probability and Statistical Inference: Summary

X: "Num of ppl that are outside ATM"

X : Binomial $(1000, 0.001)$

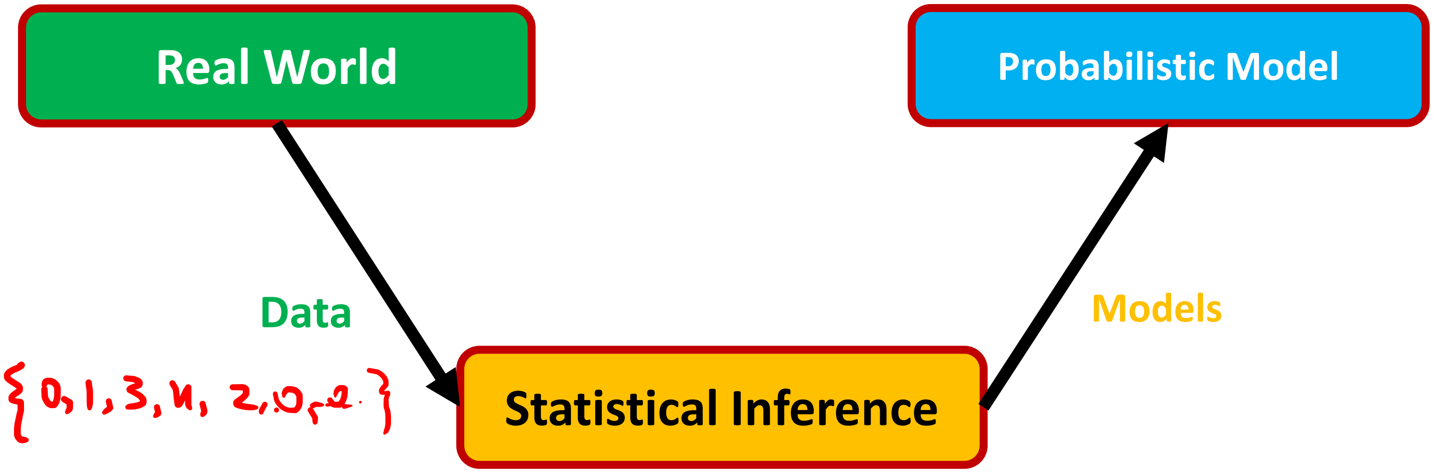
Real World

Probabilistic Model

Probability and Statistical Inference: Summary

"Num of ppl outside ATM"

X : Binomial (n, p)



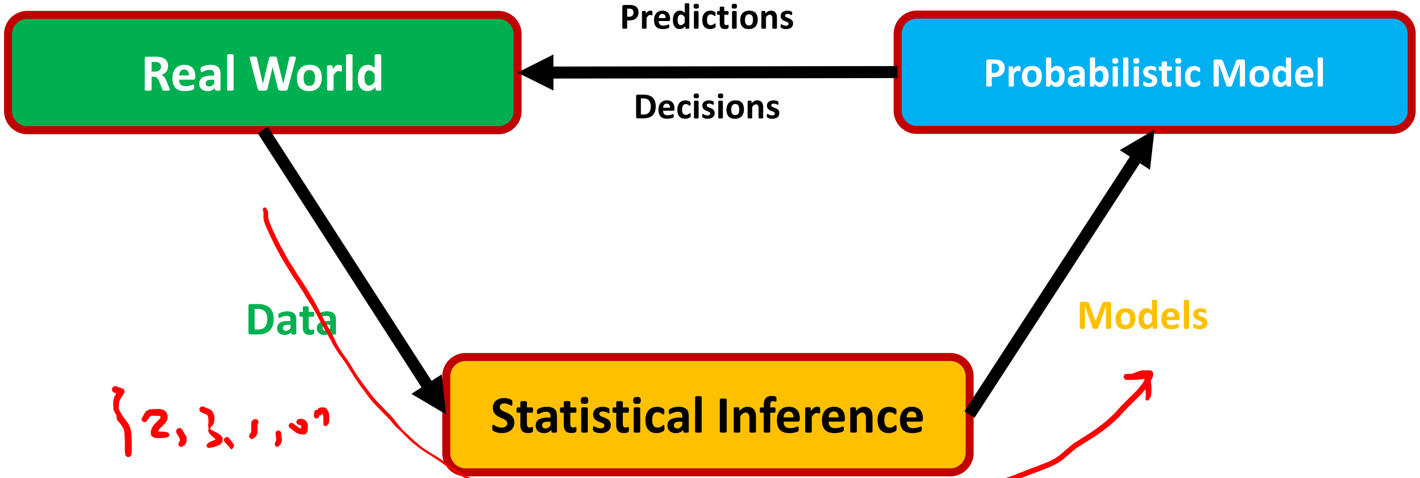
$\{0, 1, 3, 4, 2, 0, 2 \}$

Probability and Statistical Inference: Summary

300, 0.025

Num of ppl outside ATM

" INFERENCE FROM MODEL " Binomial (n, p)



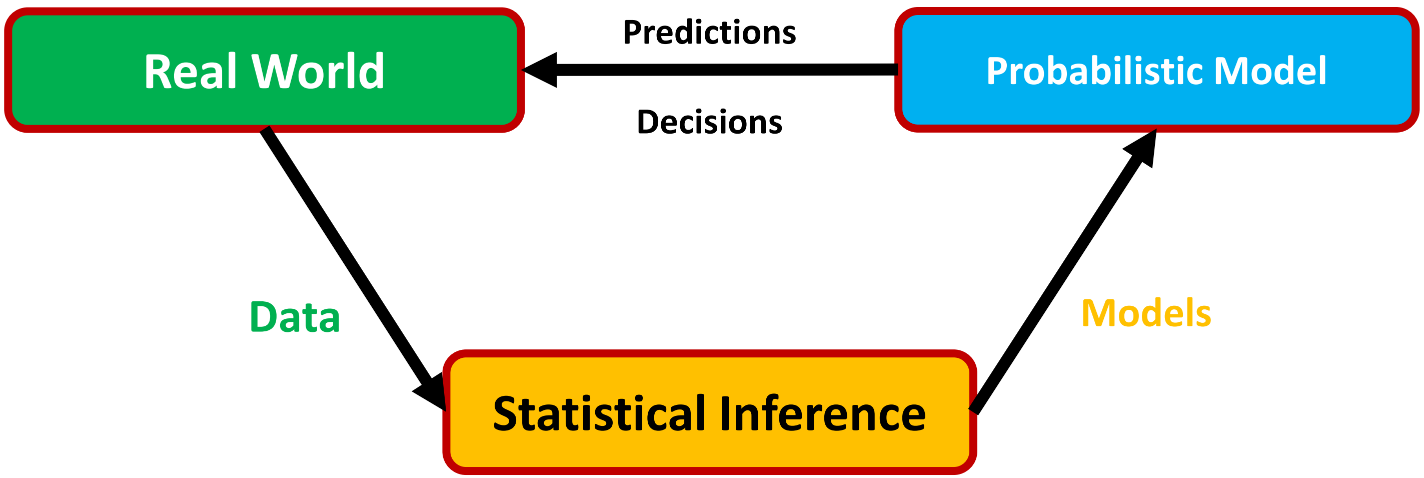
{2, 3, 1, ...}

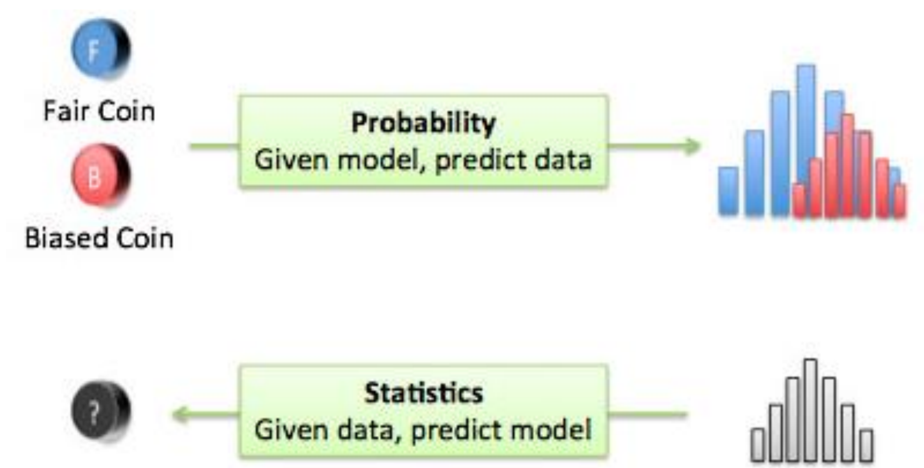
" TRAINING THE MODEL "

PROBABILITY → " STATISTICS INFERENCING " → ML

PROBABILITY → Add Stochastic Process → RL

Probability and Statistical Inference: Summary

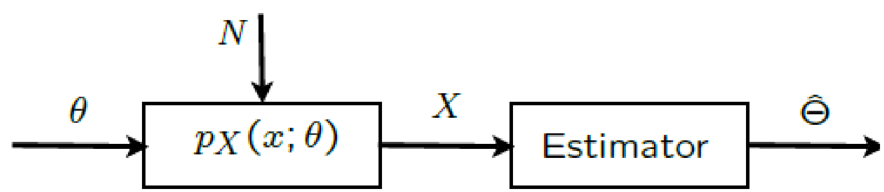




Statistical Inference: Two Approaches

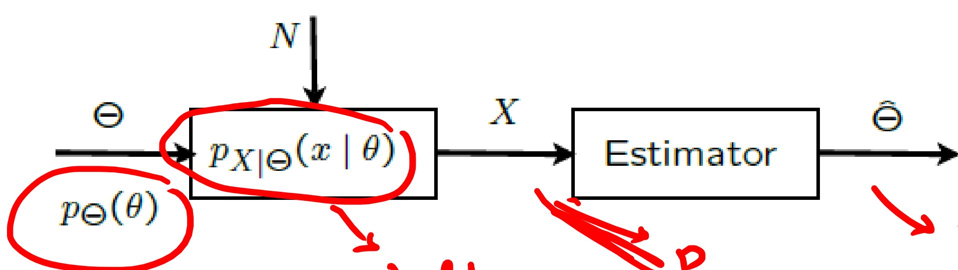
Data → $X \sim$ Binomial RV $(100, \underline{\theta})$ p

- Classical



θ: unknown parameter (not a r.v.)

- Bayesian:



PRIOR BELIEF

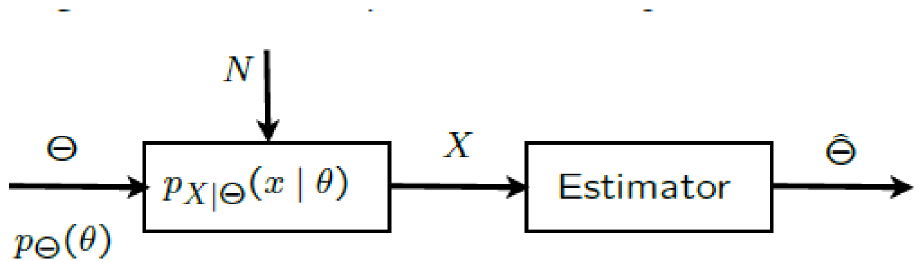
Quality of Measurement

P

Posterior Belief

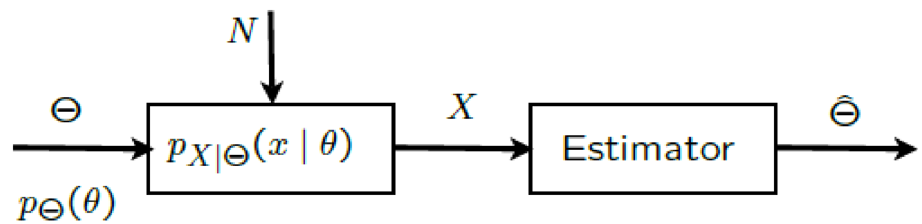
$p_{\hat{\Theta}|X}(\theta|x)$

Bayesian Statistical Inference: What is the Output?



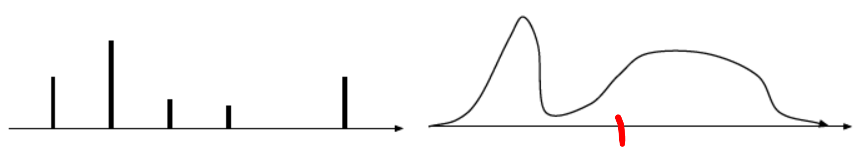
- What is the output?

Bayesian Statistical Inference: What is the Output?



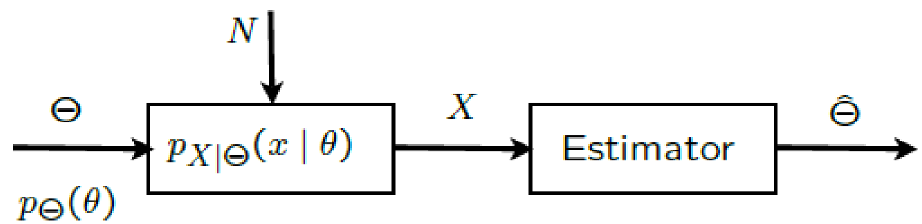
- What is the output?

$p_{\theta|x}(\theta|x)$



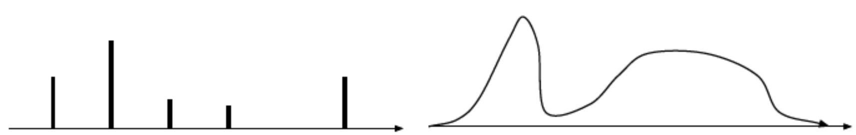
- What if we are interested in a single answer::

Bayesian Statistical Inference: Estimation vs Hypothesis Testing



- What is the output?

$f_{\Theta|X}(\theta|x)$ $f_{\Theta|X}(\theta|x)$



"HYPOTHESIS TESTING" "ESTIMATION"