ENGG2430A Probability and Statistics for Engineers

Chapter 3: General Random Variables

Instructor: Shengyu Zhang

Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

Continuous Random Variables

- We've learned discrete random variables, which can be used for dice rolling, coin flipping, etc.
- Random variables with a continuous range of possible values are quite common.
  - velocity of a vehicle traveling along the highway
- Continuous random variables are useful:
  - finer-grained than discrete random variables
  - able to exploit powerful tools from calculus.

Continuous r.v. and PDFs

- A random variable X is called continuous if there is a function $f_X \geq 0$, called the probability density function of X, or PDF, s.t.

$$P(X \in B) = \int_B f_X(x) dx$$

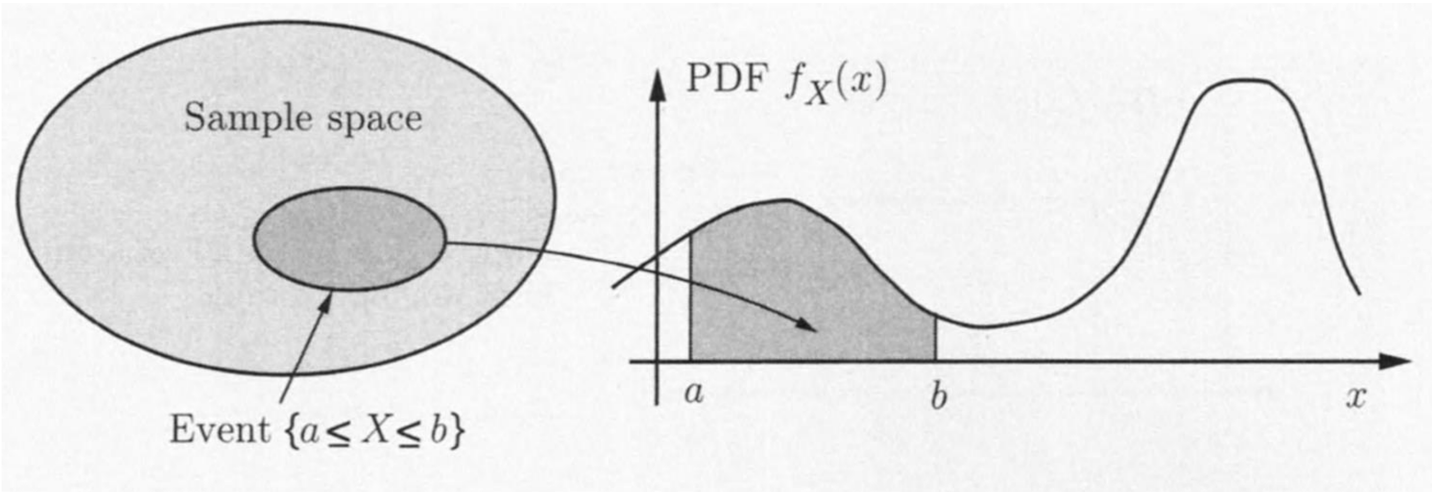
  for every subset $B \subseteq \mathbb{R}$.
  - We assume the integral is well-defined.
- Compared to discrete case: replace summation by integral.

PDF

- In particular, when $B = [a, b]$,

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

is the area under the graph of PDF.



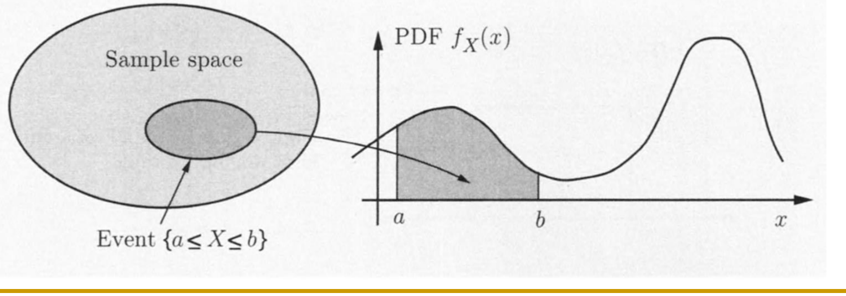
PDF

- $P(a \leq X \leq a) = \int_a^a f_X(x)dx = 0.$

$$\therefore \quad P(a \leq X \leq b) = P(a < X \leq b)$$
$$= P(a \leq X < b) = P(a < X < b)$$

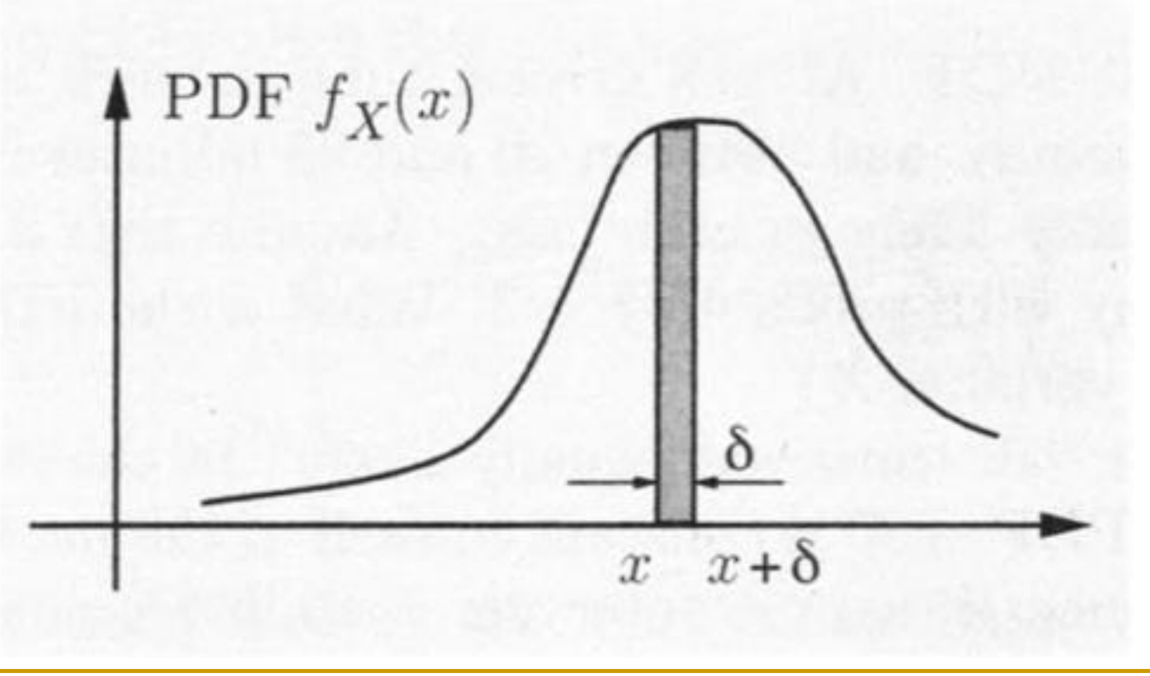
- The entire area under the graph is equal to 1.

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



Interpretation of PDF

- $f_X(x)$: "probability mass per unit length"
- $P([x, x+\delta]) = \int_x^{x+\delta} f_X(t)dt \approx f_X(x) \cdot \delta$



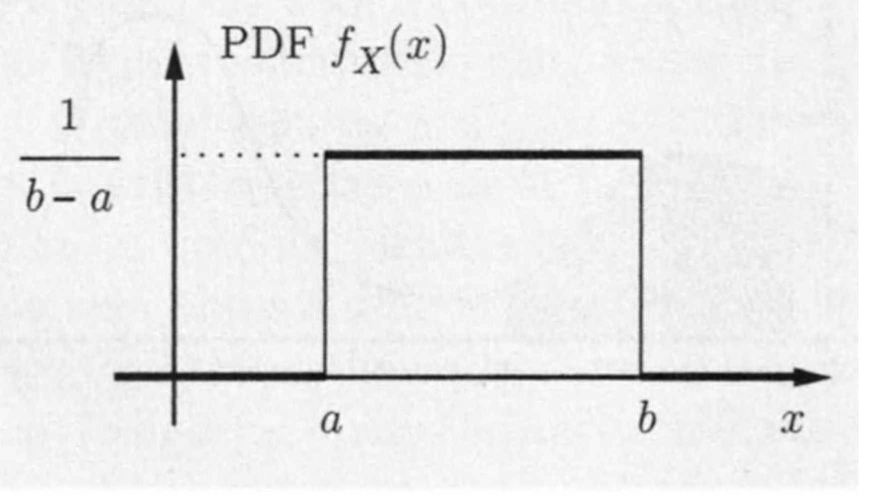
Example 1: Uniform

- Consider a random variable X takes value in interval $[a, b]$.
- Any subintervals of the same length have the same probability.
- It is called uniform random variable.

Example 1: Uniform

- Its PDF has the form

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



Example 2: Piecewise Constant

- When sunny, driving time is 15-20 minutes.
- When rainy, driving time is 20-25 minutes.
- With all times equally likely in each case.
- Sunny with prob. $2/3$, rainy with prob. $1/3$
- The PDF of driving time X is

$$f_X(x) = \begin{cases} c_{1,} & \text{if } 15 \le x \le 20 \\ c_2, & \text{if } 20 \le x \le 25 \\ 0, & \text{otherwise} \end{cases}$$

Example 2: Piecewise Constant

- $f_X(x) = \begin{cases} c_{1,} & \text{if } 15 \leq x \leq 20 \\ c_2, & \text{if } 20 \leq x \leq 25 \\ 0, & \text{otherwise} \end{cases}$
- $\frac{2}{3} = P(sunny) = \int_{15}^{20} f_X(x)dx = 5c_1$
- $\frac{1}{3} = P(rainy) = \int_{20}^{25} f_X(x)dx = 5c_2$
- Solving this gives $c_1 = \frac{2}{15}$, $c_2 = \frac{1}{15}$.

Example 3: large values

- Consider a random variable X with PDF

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

- Note that $\int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 \frac{1}{2\sqrt{x}} dx = \sqrt{x}\Big|_0^1 = 1$
  - So it's a valid PDF.
- But $\lim_{x \to 0^+} f_X(x) = \lim_{x \to 0^+} \frac{1}{2\sqrt{x}} = +\infty$.
- There, a PDF can take arbitrarily large values.

Expectation

- The expectation of a continuous random variable X is defined by

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

- As for discrete random variables, the expectation can be interpreted as
  - "center of gravity" of the PDF
  - anticipated average value of X in a large number of independent repetitions of the experiment.

Function of random variable

- For any real-valued function g, $Y = g(X)$ is also a random variable.

- The expectation of $g(X)$ is

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

Moments and variance

- The nth moment of X is defined by $\mathbf{E}[X^n]$.
- The variance of X is defined by

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

- $0 \leq \mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2$
  - Please verify the equality.
- If $Y = aX + b$, then

$\mathbf{E}[Y] = a\mathbf{E}[X] + b,$ $\qquad$ $\mathbf{Var}[Y] = a^2\mathbf{Var}[X]$.

Example: Uniform

- Consider a uniform random variable with PDF

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

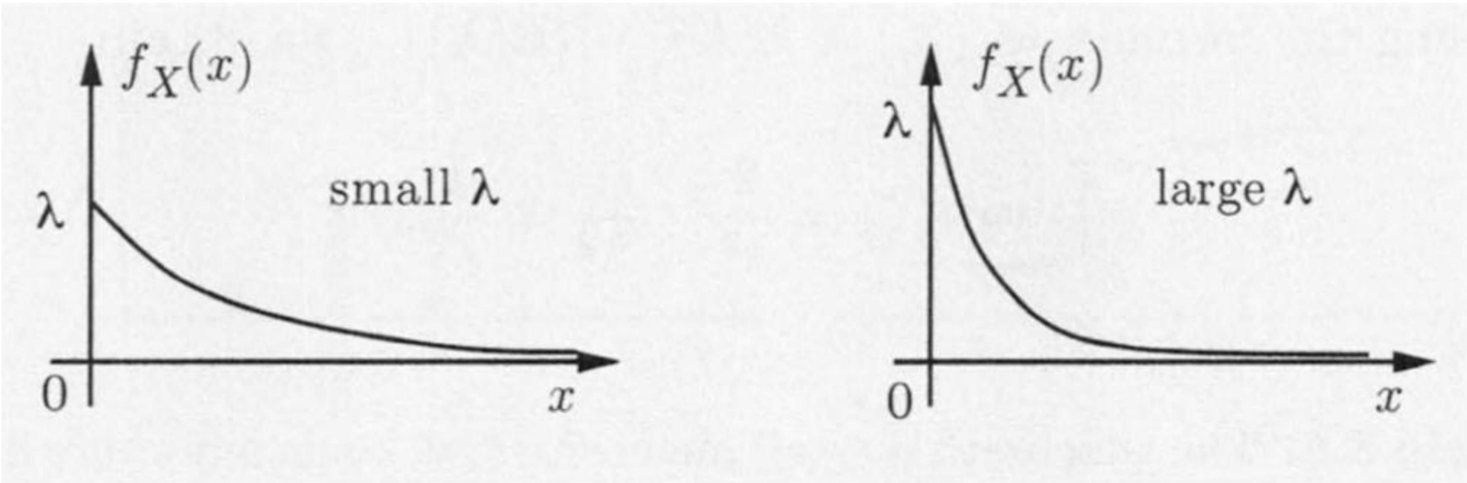
- $\mathbf{E}[X] = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{2} x^2 \Big|_a^b = \frac{a+b}{2}$.
- $\mathbf{E}[X^2] = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{3} x^3 \Big|_a^b = \frac{a^2+ab+b^2}{3}$.
- $\mathbf{Var}[X] = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{(b-a)^2}{12}$.

Example: Exponential

- An exponential random variable has PDF

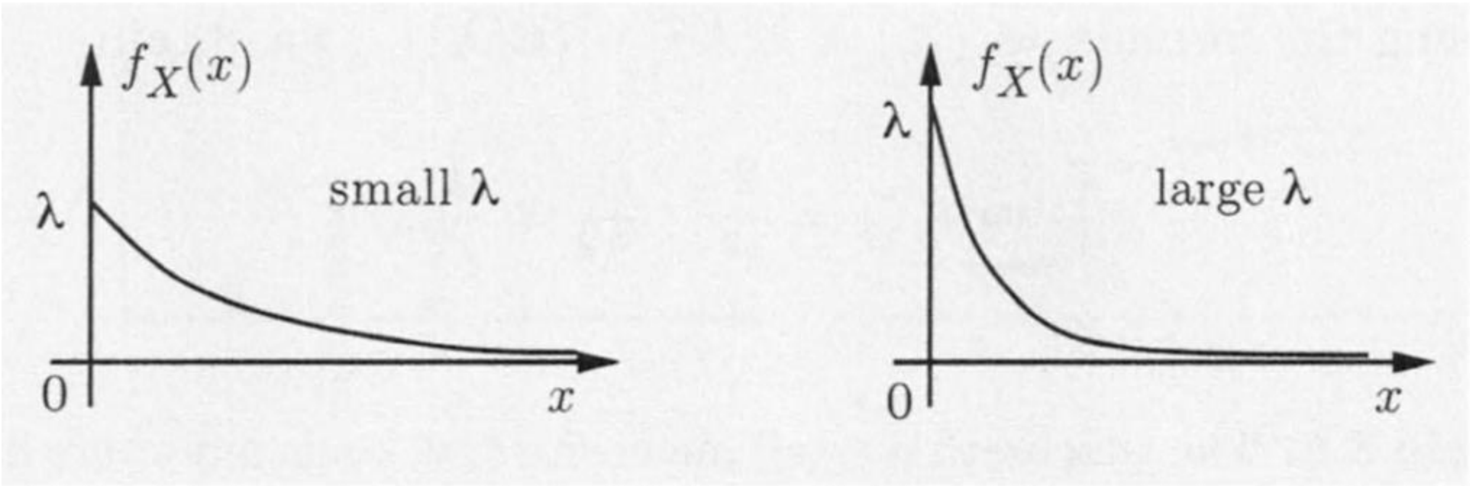
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

- Note: $f_X(0) = \lambda$.



Example: Exponential

- Note: $\int_0^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\big|_0^\infty = 1$.
  - $d(e^{-\lambda x})/dx = -\lambda e^{-\lambda x}$.
- Tail: $P(X \geq a) = \int_a^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\big|_a^\infty = e^{-\lambda a}$



Example: Exponential

- $\mathbf{E}[X] = \int_0^\infty x\lambda e^{-\lambda x} dx = -\int_0^\infty x\, de^{-\lambda x}$

$$= -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx = 0 - \frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda}$$

  - *Recall integral by parts*: $\int udv = uv - \int vdu$.

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

$$= -x^2 e^{-\lambda x}\Big|_0^\infty + \int_0^\infty 2xe^{-\lambda x} dx = \frac{2}{\lambda}\mathbf{E}[X] = \frac{2}{\lambda^2}$$

- $\mathbf{Var}[X] = \mathbf{E}[X^2] - \mathbf{E}[\mathrm{X}]^2 = 1/\lambda^2$

Example

- Time X of a meteorite first lands in Sahara.
- An exponential r.v. with mean of 10 days.
- Since $\mathbf{E}[X] = 1/\lambda$, we have $\lambda = {}^{1}/_{10}$.
- *Question: What's the probability of it first lands in 6am – 6pm of the first day?*
- $P\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) = P\left(X \geq \frac{1}{4}\right) - P\left(X \geq \frac{3}{4}\right)$

$$= e^{-\frac{1}{40}} - e^{-\frac{3}{40}} = 0.0476.$$

Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

Cumulative Distribution Function

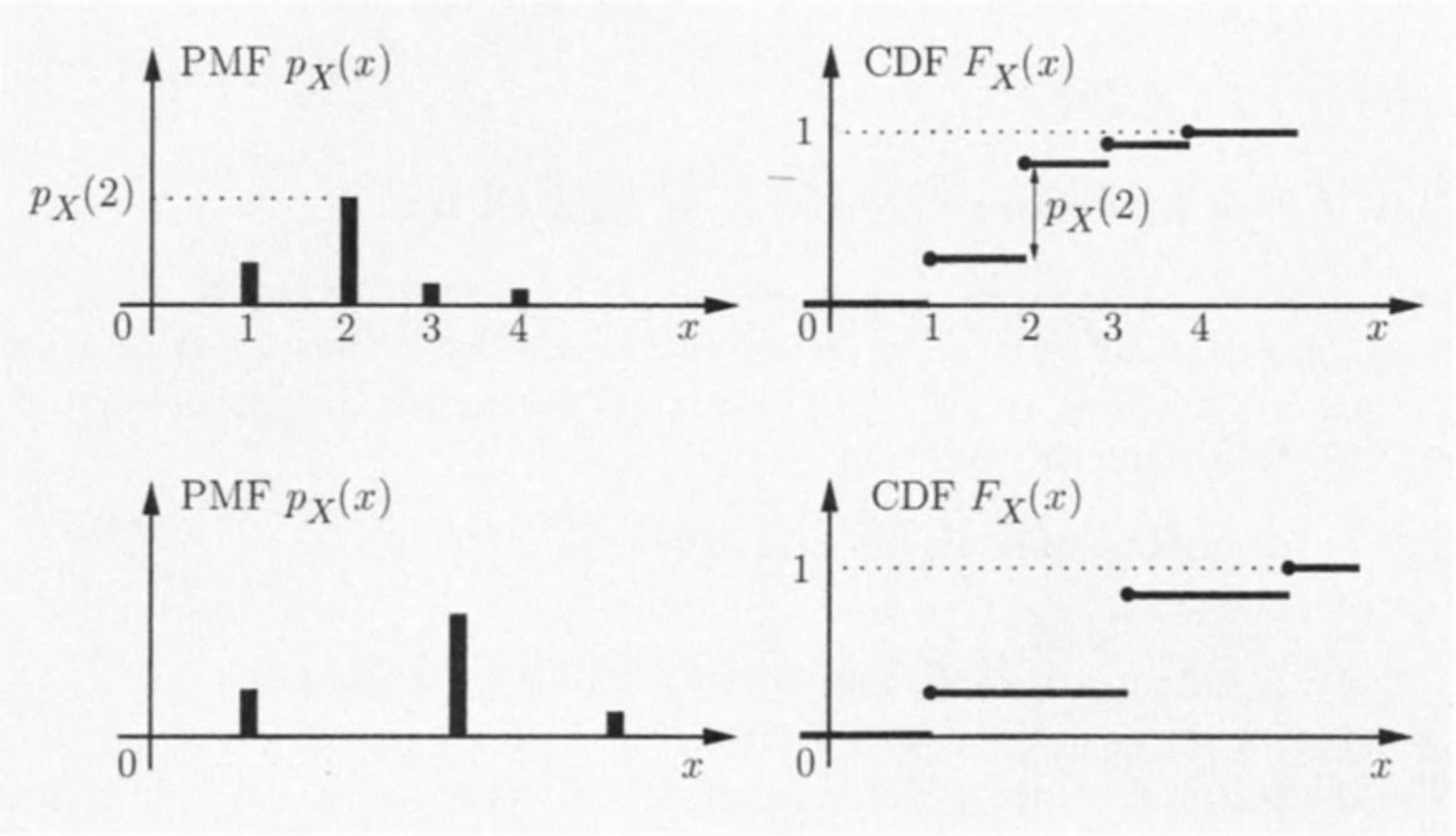
- The cumulative distribution function, or CDF, of a random variable X is

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

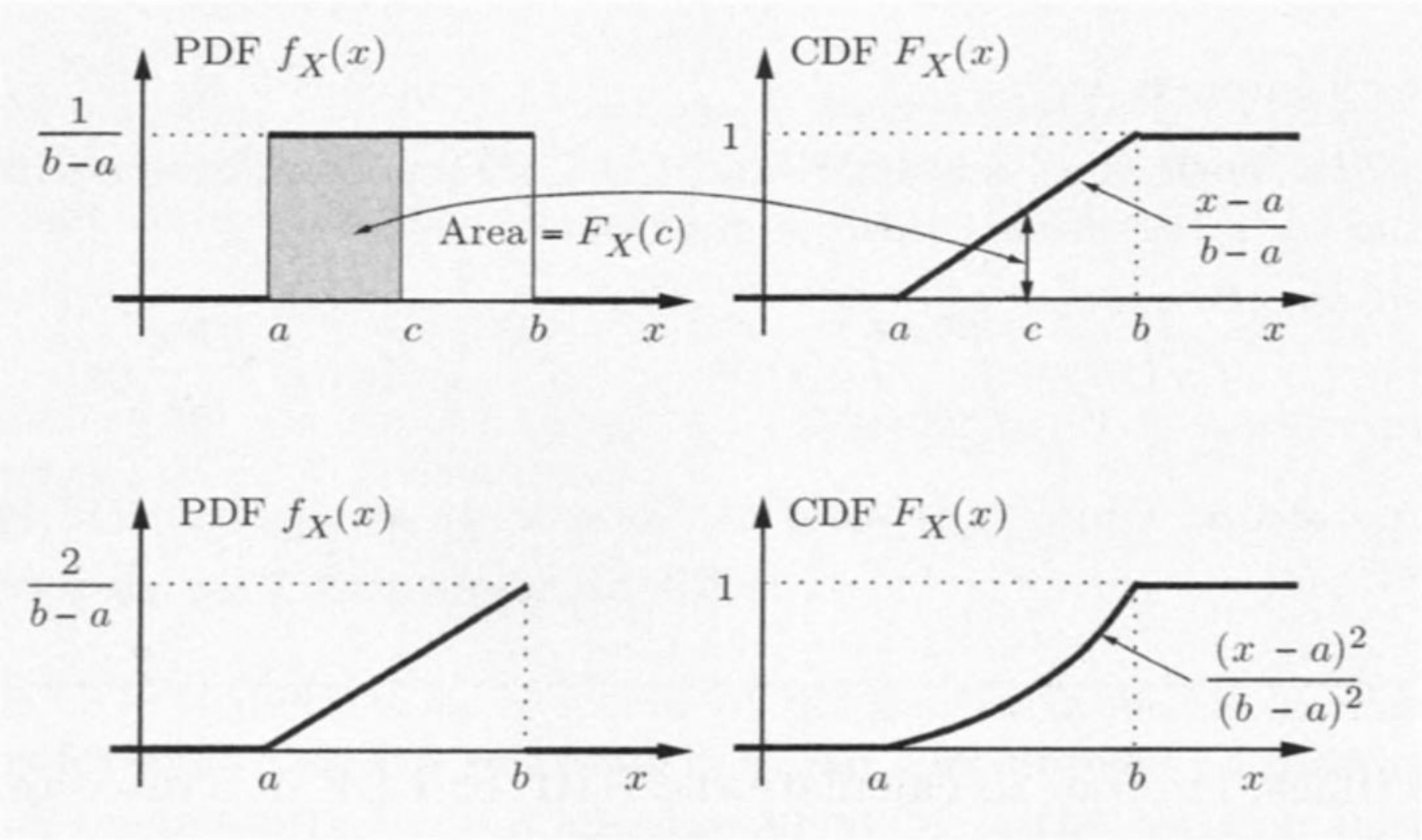
$$= \begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^{x} f_X(y)\,dy, & \text{continuous} \end{cases}$$

- The CDF $F_X(x)$ "accumulates" probability "up to" the value x.

CDF for discrete case



CDF for continuous case



Properties

- F_X is monotonically increasing:

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

- $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to +\infty} F_X(x) = 1$
- If X is discrete, F_X is piecewise constant.
- If X is continuous, F_X is continuous and

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

Example: maximum of several random variables

- Take a test three times with score in $\{1, .., 10\}$
- The final score is the maximum of the scores

$$X = \max(X_1, X_2, X_3)$$

- Each X_i takes values $\{1, .., 10\}$ eually likely, and different X_i's are independent.
- The CDF of the final score X is
- $F_X(k) = P(X \leq k)$

$$= P(X_1 \leq k) P(X_2 \leq k) P(X_3 \leq k)$$

$$= (k/10)^3$$

Example: maximum of several random variables

- Take a test three times with score in $\{1, .., 10\}$
- The final score is the maximum of the scores
$$X = \max(X_1, X_2, X_3)$$
- Each X_i takes values $\{1, .., 10\}$ eually likely, and different X_i's are independent.
- The PDF of the final score X is
- $P_X(k) = F_X(k) - F_X(k-1) = \left(\frac{k}{10}\right)^3 - \left(\frac{k-1}{10}\right)^3$

Example: Geometric and Exponential

- CDN for geometric random variable:

$$F_{geo}(n) = \sum_{k=1}^{n} p(1-p)^{k} = p\,\frac{1-(1-p)^{n}}{1-(1-p)}$$

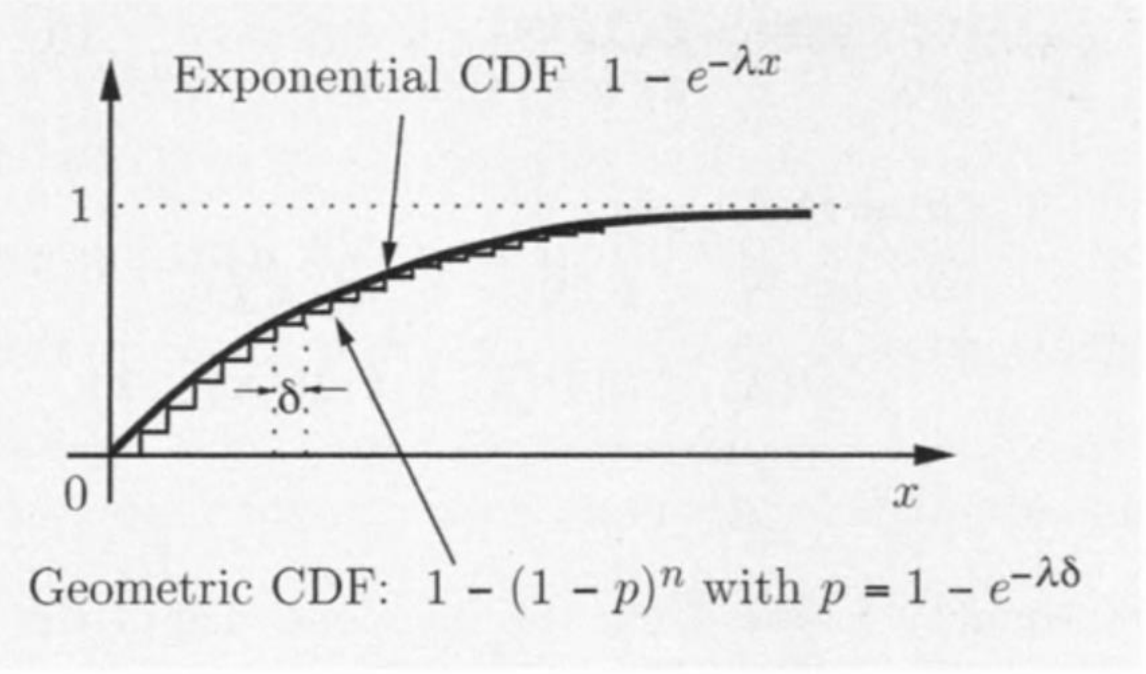
$$= 1-(1-p)^{n} \text{ for } n = 1,2,\ldots$$

- CDN for exponential random variable:
- When $x \leq 0$: $F_{exp}(x) = P(X \leq 0) = 0$
- When $x \geq 0$: $F_{exp}(x) = \int_0^x \lambda e^{-\lambda t} dt$

$$= -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$$

Example: Geometric and Exponential

- $F_{geo}(n) = 1 - (1-p)^n$, $F_{exp}(x) = 1 - e^{-\lambda x}$.
- If $e^{-\lambda\delta} = 1 - p$, then $F_{exp}(n\delta) = F_{geo}(n)$.



Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

Normal Random Variable

- A continuous random variable X is normal, or Gaussian, if it has a PDF

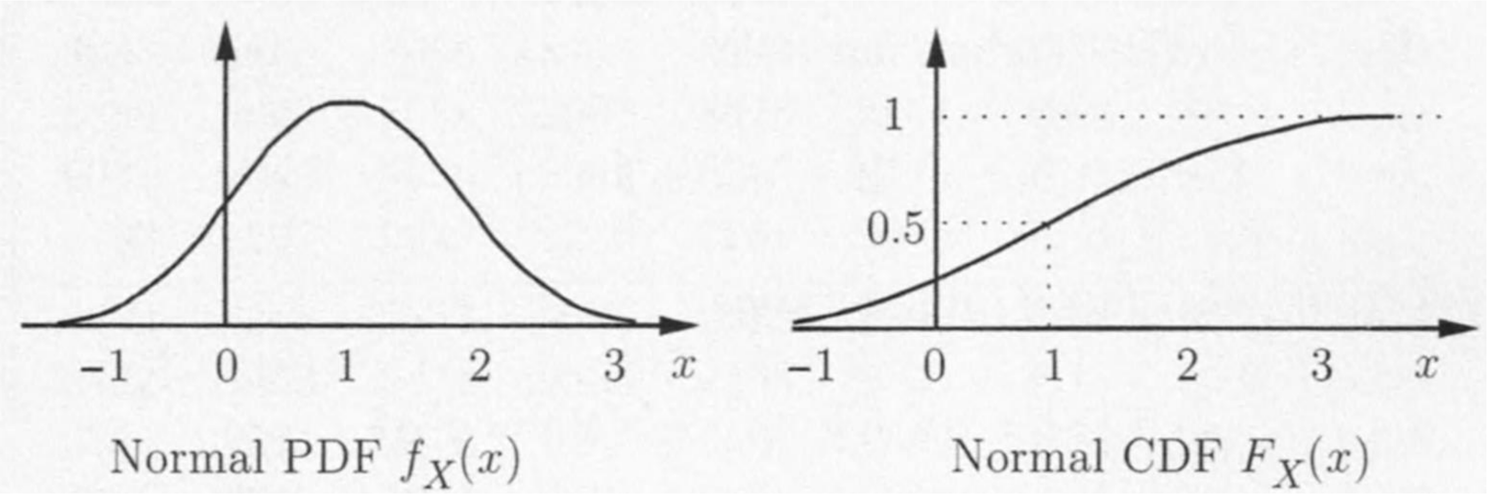
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

for some $\sigma > 0$.

- It can be verified that

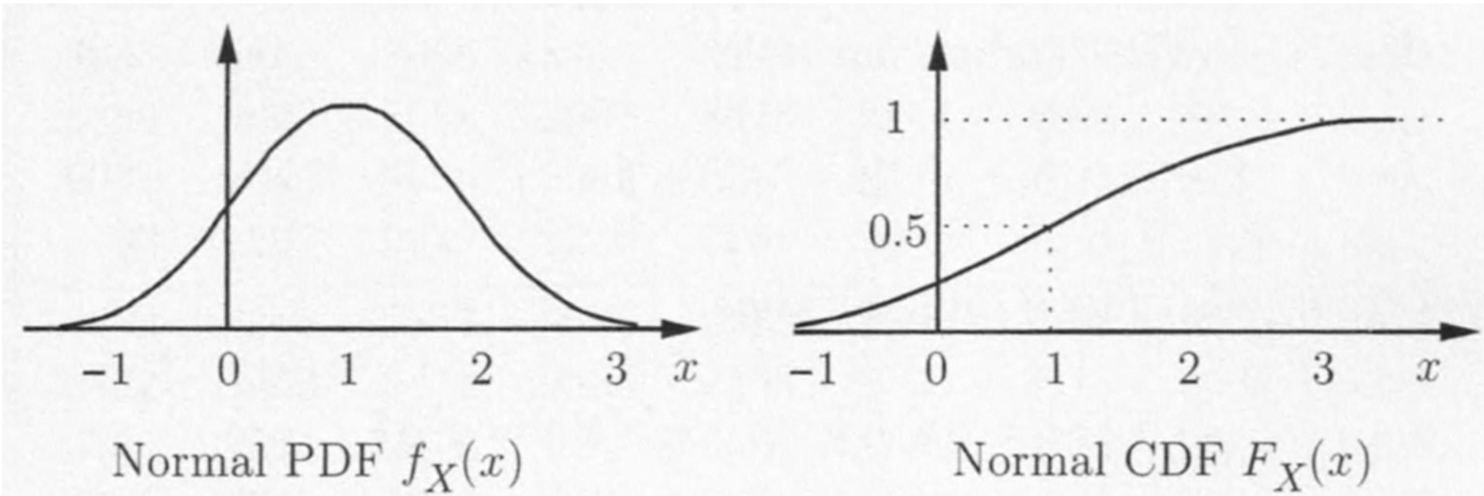
$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Example



- A normal PDF and CDF with $\mu = 1, \sigma^2 = 1$.
  - The PDF is symmetric around its mean μ, and has a characteristic bell shape.

Example



- A normal PDF and CDF with $\mu = 1, \sigma^2 = 1$.
  - As x gets further from μ. the term $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ decreases very rapidly. In this figure, the PDF is very close to zero outside the interval $[-1,3]$.

Mean and variance

- The PDF is symmetric around μ, so
$$\mathbf{E}[X] = \mu$$

- It turns out that $\mathbf{Var}[X] = \sigma^2$.

Variance

- $\mathbf{Var}[X]$

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

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy \qquad \left(\because y = \frac{x-\mu}{\sigma}\right)$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(-ye^{-\frac{y^2}{2}}\right)\Big|_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

(integral by parts)

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

$$= \sigma^2$$

Standard Normal

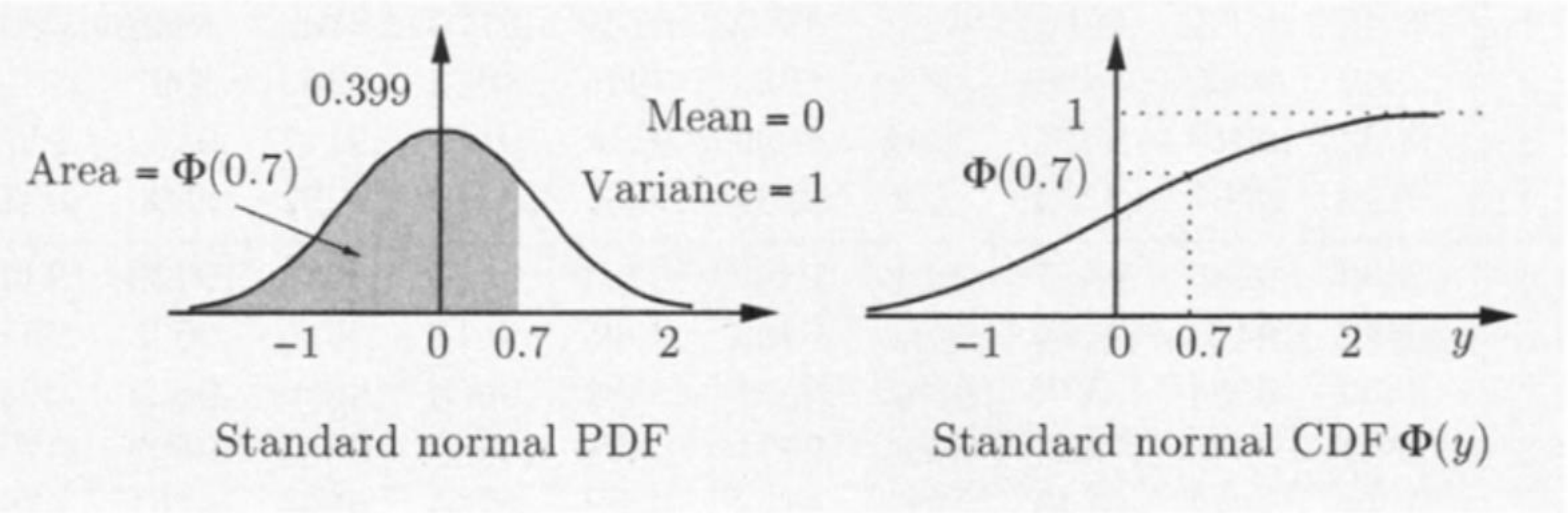
- The normal random variable with zero mean and unit variance is a standard normal. Its CDF is denoted by Φ:

$$\Phi(y) = P(Y \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt$$

- By symmetry, it holds

$$\Phi(-y) = 1 - \Phi(y)$$

Standard Normal



Standard Normal

- Table of $\Phi(x)$ for positive x.
  - $\Phi(-0.5) = 1 - \Phi(0.5) = 1 - .6915 = .3085$

	.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

Standard Normal

- Let X be a normal random variable with mean μ and variance σ^2. Then

$$Y = \frac{X - \mu}{\sigma}$$

is normal and

$$\mathbf{E}[Y] = \frac{\mathbf{E}[X] - \mu}{\sigma} = 0, \quad \mathbf{Var}[\mathrm{Y}] = \frac{\mathbf{Var}[X]}{\sigma^2} = 1$$

Thus, Y is standard normal.

  - This fact allows us to calculate the probability of any event defined in terms of X: we redefine the event in terms of y, and then use the standard normal table.

Example 1: Using the normal table

- annual snowfall at a certain place
- normal r.v. with mean $\mu = 60$ and standard deviation $\sigma = 20$.
- *Question: What is the probability that this year's snowfall will be at least 80 inches?*
- $P(X \geq 80) = P\left(\frac{X-60}{20} \geq \frac{80-60}{20}\right) = P(Y \geq 1)$

$= 1 - \Phi(1) = 1 - .8413 = .1587$

Example 1: Using the normal table

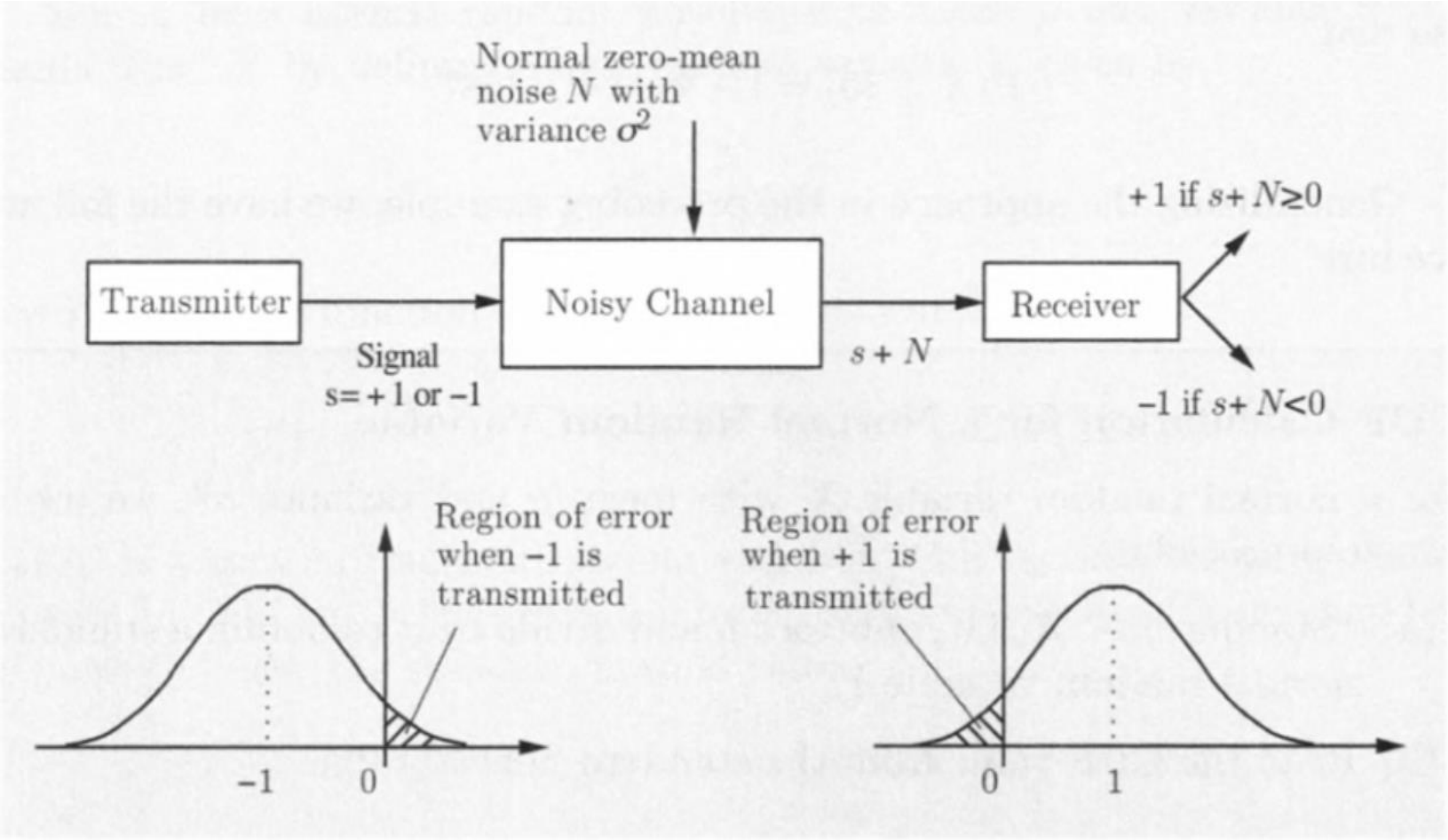
- In general, we can calculate the CDF for a normal random variable as follows.
- For a normal random variable X with mean μ and variance σ^2, we
  - first “standardize" X, i.e., subtract μ and divide by σ^2, to obtain a standard normal random variable Y
  - read the CDF value from standard normal table:

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

Example 2: Signal detection

- A binary message is transmitted as a signal s, which is either $+1$ or -1.
- The communication corrupts the transmission with additional 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).

Example 2: Signal detection 2



Example 2: Signal detection 3

- *Question*: *What is the probability of error?*
- The error occurs whenever -1 is transmitted and the noise N is at least 1, or whenever $+1$ is transmitted and the noise is smaller than -1.
- $$P(N \geq 1) = 1 - P(N < 1)$$
$$= 1 - P\left(\frac{N-\mu}{\sigma} < \frac{1-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{1-\mu}{\sigma}\right)$$
$$= 1 - \Phi\left(\frac{1}{\sigma}\right)$$

Normal Random Variable

- Normal random variables play an important role in a broad range of probabilistic models.
- The main reason is that they model well the additive effect of many independent factors.
- The sum of a large number of independent and identically distributed (not necessarily normal) random variables $\approx$ normal CDF.
  - regardless of CDF of individual random variables.
  - More on this in Chapter 5.

Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

Joint PDFs

- 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 nonnegative function that satisfies

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

for every subset B of the two-dimensional plane.

Joint PDFs

- Normalization: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\,dy = 1$
- To interpret the joint PDF, let δ be a small positive number,
- $P(a \le X \le a+\delta, c \le Y \le c+\delta)$

$$= \int_{c}^{c+\delta}\int_{a}^{a+\delta} f_{X,Y}(x,y)dx\,dy \approx f_{X,Y}(a,c)\delta^2$$

- $f_{X,Y}(a,c)$: "probability per unit area" in the vicinity of (a,c).

Marginal Probability

- $P(X \in A) = P\big(X \in A, Y \in (-\infty, \infty)\big)$

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

- Recall $P(X \in A) = \int_{x \in A} f_X(x)\,dx$
- Thus the marginal PDF of X is given by

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

- Similarly, the marginal PDF of Y is

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

Example: 2D Uniform PDF

- Romeo and Juliet have a date at a given time
- Each will arrive with a delay between 0 and 1 hour.
- Let X, Y denote their delays.
- Assume that no pairs in the unit square is more likely than others

Example: 2D Uniform PDF

- Then the joint PDF is of the form

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

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

$$c = 1$$

2D Uniform PDF

- In general, let S be a subset of the two dimensional plane. The corresponding uniform joint PDF on S is defined by

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

2D Uniform PDF

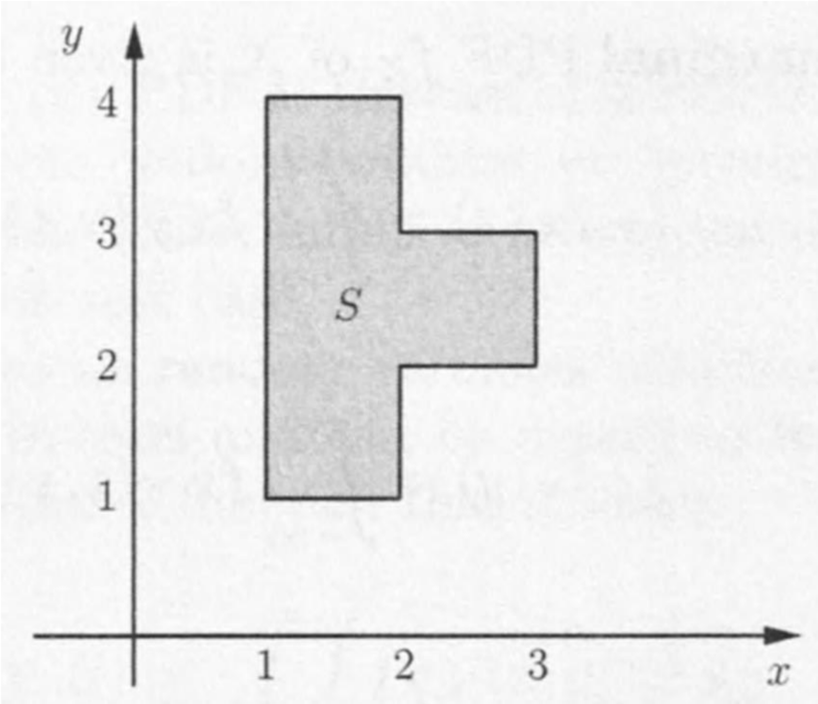
- For any subset $A \subset S$, the probability that (X, Y) lies in A is

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

$$= \frac{1}{\text{area of } S} \iint_{(x,y)\in A} dxdy = \frac{\text{area of } A}{\text{area of } S}$$

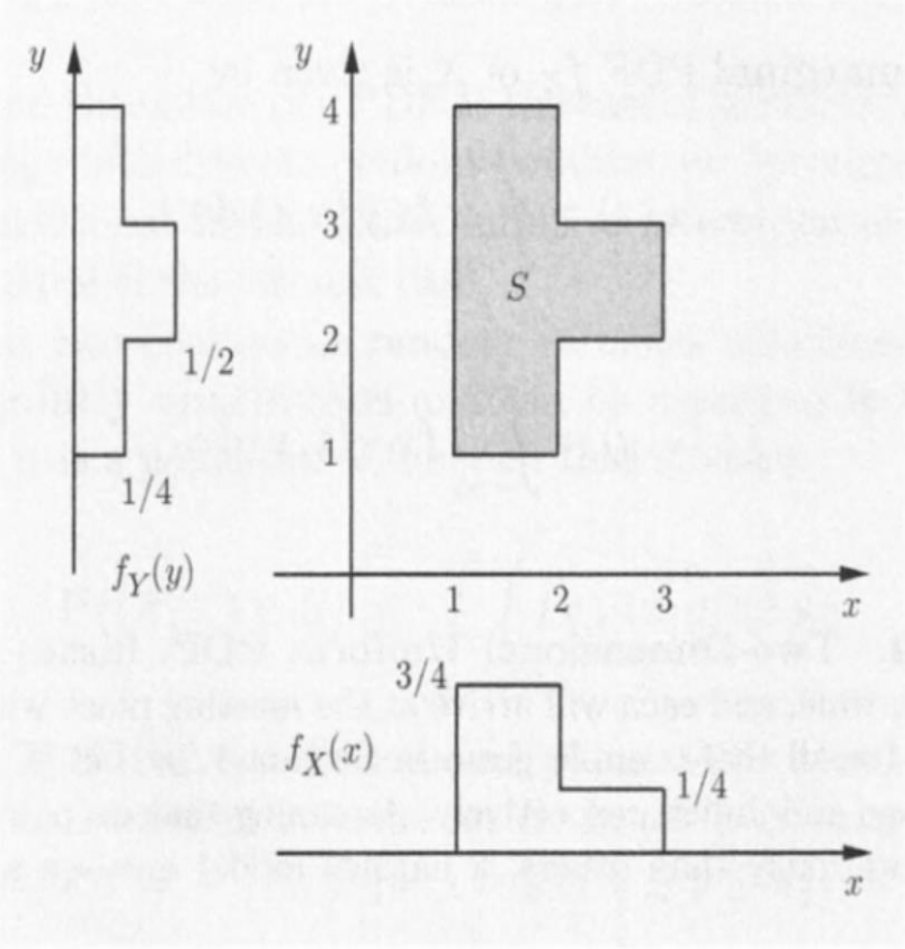
Example 2

- Suppose the joint PDF of X, Y is a constant c on S and 0 outside.
- *Question*: *What is c*?
- *Question*: *What are the marginal PDFs of X and Y?*



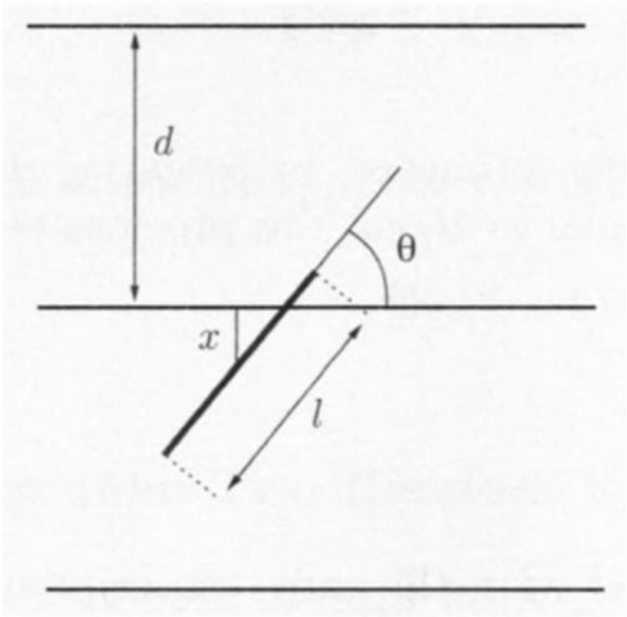
Example 2

- Suppose the joint PDF of X, Y is a constant c on S and 0 outside.
- The area of S is 4, so $c = 1/4$.
- The marginal PDFs of X, Y are shown in the figure.



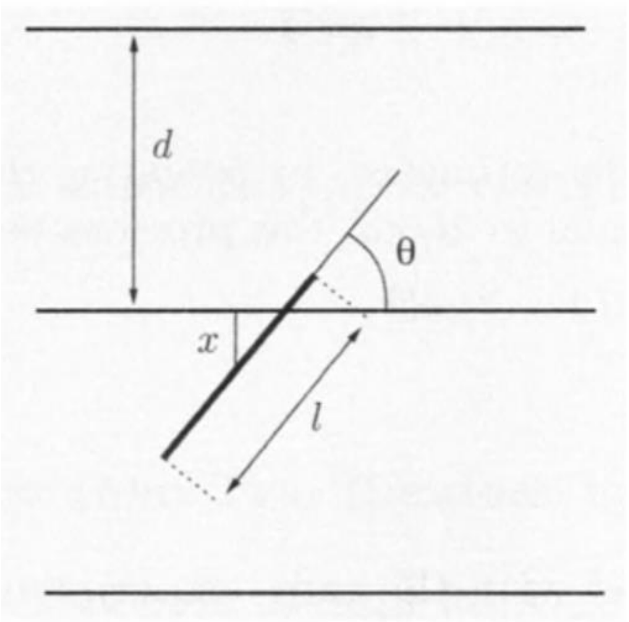
Example: Buffon's Needle

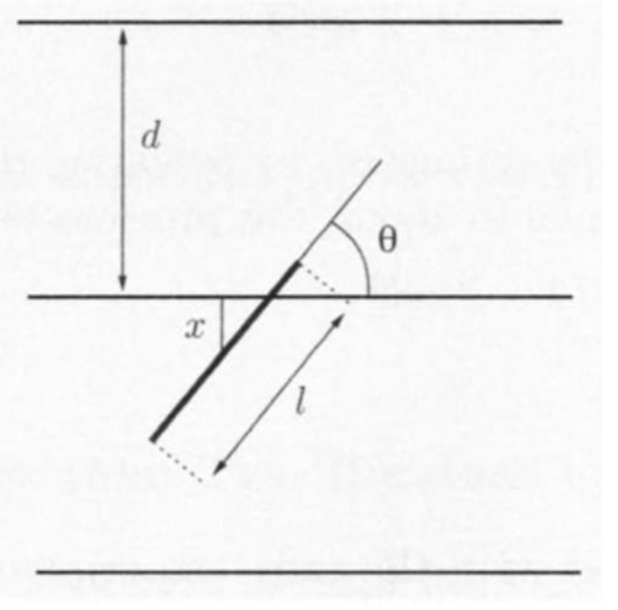
- A surface is ruled with parallel lines, which at distance d from each other.
- Suppose we throw a needle of length l randomly.
- *Question: What is the probability that the needle will intersect one of the lines?*
  - Assume $l < d$ so that the needle cannot intersect two lines simultaneously.



Example: Buffon's Needle

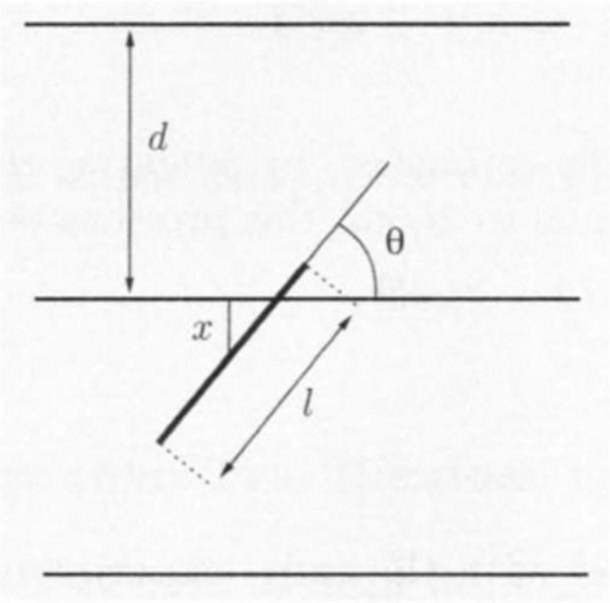
- X: the distance from the middle point of the needle and the nearest of the parallel lines
- Θ: the acute angle formed by the needle and the lines





- We model (X, Θ) with a uniform joint PDF:

$$f_{X,\Theta}(x,\theta) = \begin{cases} \dfrac{4}{\pi d} & \text{if } x \in \left[0, \dfrac{d}{2}\right] \text{ and } \theta \in \left[0, \dfrac{\pi}{2}\right] \\ 0 & \text{otherwise} \end{cases}$$



- The needle will intersect one of the lines if and only if

$$X \leq \frac{l}{2} \sin \Theta$$

Example: Buffon's Needle

- So the probability of intersection is

$$P\left(X \le \frac{l}{2}\sin\Theta\right) = \iint_{X \le \frac{l}{2}\sin\theta} f_{X,\Theta}(x,\theta)\,dx\,d\theta$$

$$= \frac{4}{\pi d}\int_0^{\pi/2}\int_0^{\left(\frac{l}{2}\right)\sin\theta} dx\,d\theta = \frac{4}{\pi d}\int_0^{\pi/2}\left(\frac{l}{2}\right)\sin\theta\, d\theta$$

$$= \frac{2l}{\pi d}(-\cos\theta)\Big|_0^{\pi/2} = \frac{2l}{\pi d}$$

Joint CDFs

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

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

Joint CDFs

- If X and Y are described by a joint PDF $f_{X,Y}$, then

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

and

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

Example

- Let X and Y be described by a uniform PDF on the unit square $[0,1]^2$.
- The joint CDF is given by

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

- Then

$$\frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y) = \frac{\partial^2 (xy)}{\partial x \partial y}(x,y) = 1 = f_{X,Y}(x,y)$$

for all (x,y) in the unit square.

Expectation

- If X and Y are jointly continuous random variables and g is some function, then
$$Z = g(X,Y)$$
 is also a random variable.
- And
$$\mathbf{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dxdy$$

Expectation

- If $g(X,Y)$ is a linear function:

$$g(X,Y) = aX + bY + c$$

  for some scalars a, b, c.

- then

$$\mathbf{E}[aX + bY + c] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c$$

- “linearity of expectation”

More than two random variables

- The joint PDF of X, Y and Z satisfies

$$P\big((X, Y, Z) \in B\big) = \iiint_{(x,y,z)\in B} f_{X,Y,Z}(x, y, z)dxdydz$$

  for any set B.
- Marginal:
  - $f_{X,Y}(x, y) = \int_{-\infty}^{\infty} f_{X,Y,Z}(x, y, z)dz$
  - $f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x, y, z)dy\,dz$

More than two random variables

- $\mathbf{E}[g(X, Y, Z)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y, z) f_{X,Y,Z}(x, y, z) dx\, dy\, dz$

- If g is linear, of the form $aX + bY + cZ$, then
  $\mathbf{E}[aX + bY + cZ] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c\mathbf{E}[Z]$

- In general,
  $\mathbf{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbf{E}[X_1] + \cdots + a_n \mathbf{E}[X_n]$

Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

Conditioning

- Similar to the case of discrete random variables, we can condition a random variable
  - on an event, or
  - on another random variable,
- and define the concepts of conditional PDF and conditional expectation.

Conditioning a r.v. on an event

- The conditional PDF of a continuous random variable X, given an event A with $P(A) > 0$, is defined as a nonnegative function $f_{X|A}$ that satisfies

$$P(X \in B|A) = \int_B f_{X|A}(x)dx\,,$$

for any subset B of the real line.

Conditioning a r.v. on an event

- In particular, taking $B = \mathbb{R}$

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

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

Conditioning on event $\{X \in A\}$

- If we condition on event $\{X \in A\}$, with $P(X \in A) > 0$, then

$$P(X \in B | X \in A)$$

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

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

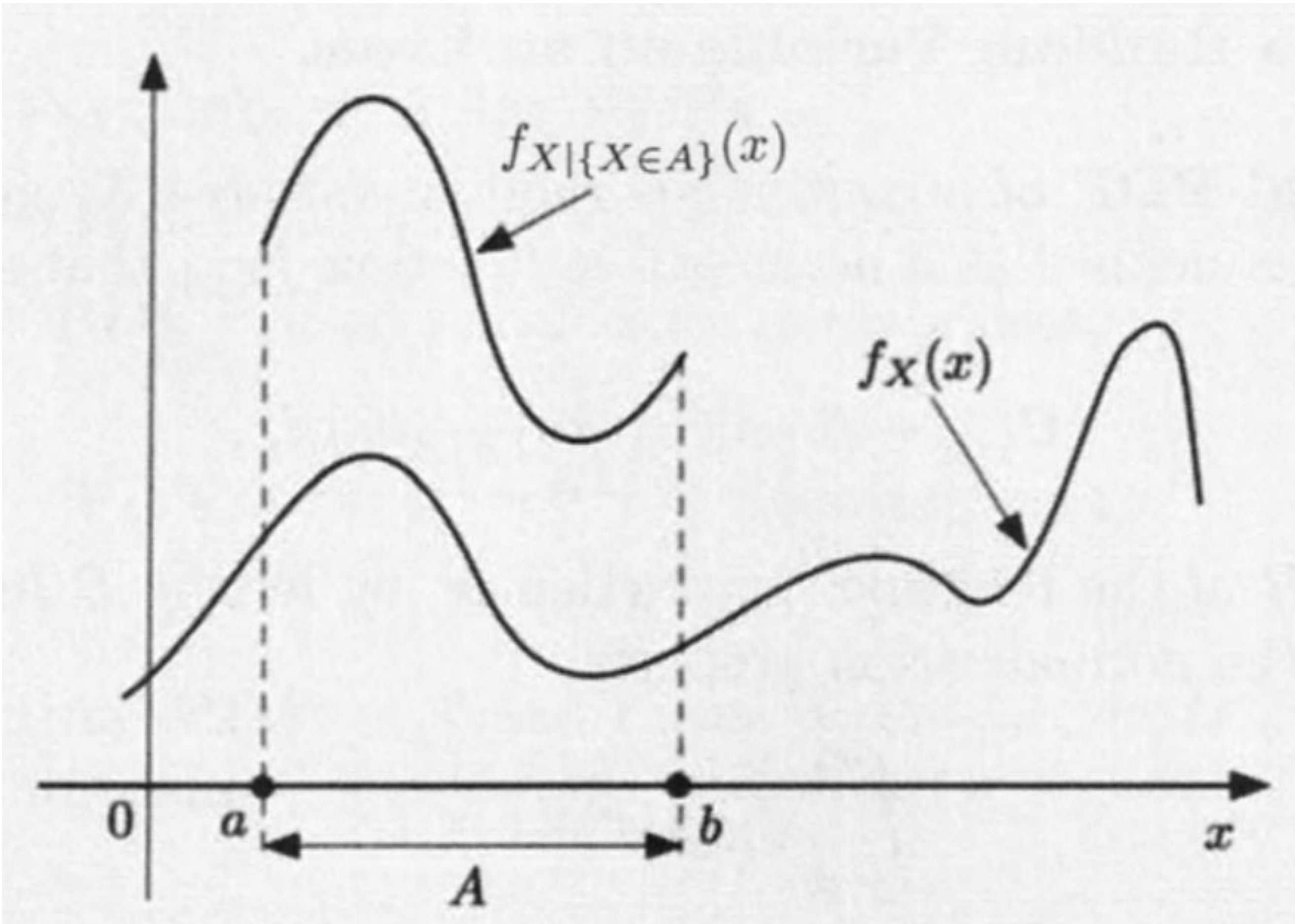
Conditioning on event $\{X \in A\}$

- Comparing with the earlier formula gives

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

  - The conditional PDF is zero outside the conditioning set.
  - Within the conditioning set, the conditional PDF has the same shape as the unconditional one, except that scaled by a factor $1/P(X \in A)$

Conditioning on event $\{X \in A\}$



Example: exp. r.v. is memoryless

- The time T until a new light bulb burns out is an exponential random variable with parameter λ.
- Alice turns the light on, leaves the room, and when she returns, t time units later, finds the light bulb is still on, which corresponds to the event

$$A = \{T > t\}$$

Example: exp. r.v. is memoryless

- Let X be the additional time until the light bulb burns out.
- *Question: What's the conditional CDF of X given the event A?*
- $P(X > x|A) = P(T > t + x|T > t)$

$$= \frac{P(T > t + x \text{ and } T > t)}{P(T > t)} = \frac{P(T > t + x)}{P(T > t)}$$

$$= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x}$$

Example: exp. r.v. is memoryless

- Last slide: $P(X > x|A) = e^{-\lambda x}$.
- Recall tail probability of exponential r.v.:
$$P(X \geq a) = e^{-\lambda a}.$$
- Observation: The conditional CDF of X is exponential with parameter λ, regardless of the time t that elapsed between the lighting of the bulb and Alice's arrival.
- Thus the exponential random variable is called *memoryless*.

Conditioning with multiple r.v.

- Suppose X and Y are jointly continuous random variables, with joint PDF $f_{X,Y}$.
- If we condition on a positive probability event of the form $C = \{(X, Y) \in A\}$, we have

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

- The conditional PDF of X, given event C, is

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

Total probability theorem

- If the events $A_1, \cdots, A_n$ form a partition of the sample space, then

$$f_X(x) = \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x)$$

- Next we give a proof.

Proof of total probability theorem

- By the total probability theorem from Chapter 1, we have

$$P(X \le x) = \sum_{i=1}^{n} P(A_i) P(X \le x | A_i)$$

- This formula can be written as

$$\int_{-\infty}^{x} f_X(t)dt = \sum_{i=1}^{n} P(A_i) \int_{-\infty}^{x} f_{X|A_i}(t)dt$$

- Then take the derivative with respect to x and get the result.

Example: Taking train

- The metro train arrives at the station every quarter hour starting at 6:00 a.m.
- You walk into the station between 7:10-7:30 a.m. uniformly.
- *Question: What's the PDF of the time you have to wait for the first train to arrive?*

Example: Taking train

- Denote the time of your arrival by X, which is then a uniform random variable on 7:10-7:30
- Let Y be the waiting time.

- Let A and B be the events

$$A = \{7{:}10 \le X \le 7{:}15\} = \{board\ 7{:}15\ train\}$$
$$B = \{7{:}15 < X \le 7{:}30\} = \{board\ 7{:}30\ train\}$$

Example: Taking train

- Condition on event A, Y is uniform on 0-5
- Condition on event B, Y is uniform on 0-15
- Total probability theorem:

$$f_Y(y) = P(A)f_{Y|A}(y) + P(B)f_{Y|B}(y)$$

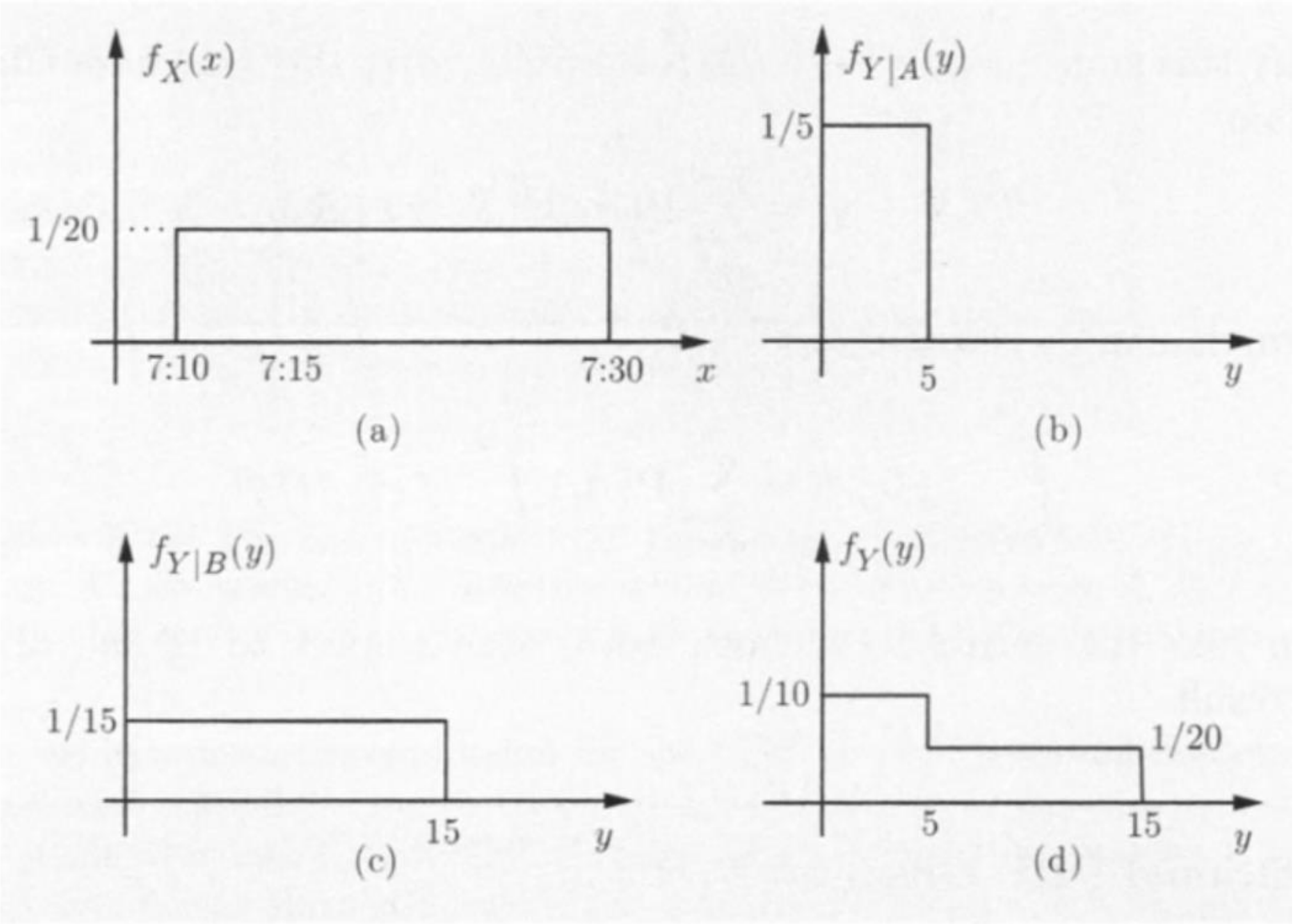
- For $0 \leq y \leq 5$,

$$f_Y(y) = \frac{1}{4} \cdot \frac{1}{5} + \frac{3}{4} \cdot \frac{1}{15} = \frac{1}{10}$$

- For $5 < y \leq 15$,

$$f_Y(y) = \frac{1}{4} \cdot 0 + \frac{3}{4} \cdot \frac{1}{15} = \frac{1}{20}$$

Example: Taking train



Conditioning one r.v. on another

- 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
$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$
- This is analogous to the formula
$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$
for the discrete case.

Conditioning one random variable on another

- Because

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

then

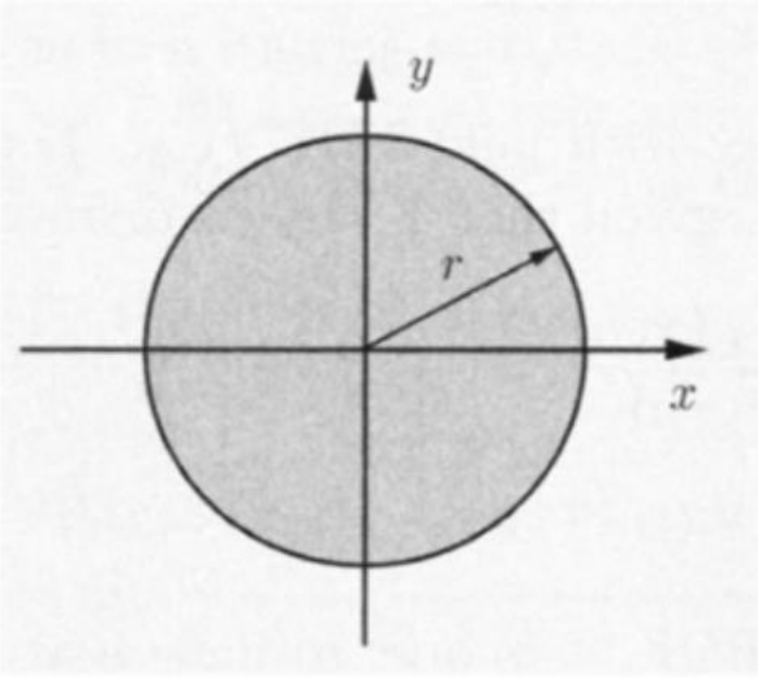
$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1,$$

for any fixed y.

- Thus $f_{X|Y}(x|y)$ is a legitimate PDF.

Example: Circular uniform PDF

- Bob throws a dart at a circular target of radius r.
- He always hits the target.
- All points of impact (x, y) are equally likely.
- Then the joint PDF of the random variables X, Y is uniform.



Example: Circular uniform PDF

- Because the area of the circle is πr^2,

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi r^2} & \text{if } x^2 + y^2 \leq r^2, \\ 0 & \text{otherwise.} \end{cases}$$

- To calculate the conditional PDF $f_{X|Y}(x|y)$, let us find the marginal PDF $f_Y(y)$.

Example: Circular uniform PDF

- For $|y| > r$, $f_{X|Y}(x|y) = 0$
- For $|y| \le r$,

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

$$= \frac{1}{\pi r^2}\int_{x^2+y^2\le r^2} dx = \frac{1}{\pi r^2}\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx$$

$$= \frac{2\sqrt{r^2-y^2}}{\pi r^2}$$

Example: Circular uniform PDF

- The conditional PDF is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{\pi r^2}}{\frac{2}{\pi r^2}\sqrt{r^2 - y^2}}$$

$$= \frac{1}{2\sqrt{r^2 - y^2}}, \qquad \text{if } x^2 + y^2 \leq r^2$$

- Thus for a fixed value of y, the conditional PDF $f_{X|Y}$ is uniform.

Conditional probability on zero event

- Let us fix some small positive numbers δ_1 and δ_2, and condition on the event
$$B = \{y \le Y \le y + \delta_2\}.$$
- Then
$$P(x \le X \le x + \delta_1 | y \le Y \le y + \delta_2)$$
$$= \frac{P(x \le X \le x + \delta_1 \text{ and } y \le Y \le y + \delta_2)}{P(y \le Y \le y + \delta_2)}$$
$$\approx \frac{f_{X,Y}(x, y)\delta_1\delta_2}{f_Y(y)\delta_2} = f_{X|Y}(x|y)\delta_1,$$
which is independent of δ_2

Conditional probability on zero event

- Let $\delta_2 \to 0$, we get
$$P(x \leq X \leq x + \delta_1 | Y = y) \approx f_{X|Y}(x|y)\delta_1$$
for δ_1 small, and more generally
$$P(X \in A | Y = y) = \int_A f_{X|Y}(x|y)dx$$
- This gives a conditional probability on the zero event $\{Y = y\}$.

Example: Vehicle speed

- The speed of a typical vehicle that drives past a police radar is an exponential random variable X with mean 50.
- The police radar's measurement Y has an error which is modeled as a normal random variable with zero mean and standard derivation equal to one tenth of the vehicle's speed.

Example: Vehicle speed

- *Question: What is the joint PDF of X and Y?*

- First, $f_X(x) = \left(\frac{1}{50}\right) e^{-\frac{x}{50}}$, for $x \geq 0$
- Also, conditioned on $X = x$, the measurement Y has a normal PDF with mean x and variance $x^2/100$. Thus

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}(x/10)} e^{-(y-x)^2/(2x^2/100)}.$$

Example: Vehicle speed

- Thus, for all $x \geq 0$ and all y,

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

$$= \frac{1}{50} e^{\frac{x}{50}} \frac{10}{\sqrt{2\pi x}} e^{-\frac{50(y-x)^2}{x^2}}$$

Conditioning for more than two random variables

- $f_{X.Y|Z}(x, y|z) = \frac{f_{X,Y,Z}(x,y,z)}{f_Z(z)}$, if $f_Z(z) > 0$.
- $f_{X|Y,Z}(x|y, z) = \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)}$, if $f_{Y,Z}(y, z) > 0$.

- There is an analog of the multiplication rule

$$f_{X,Y,Z}(x, y, z) = f_{X|Y,Z}(x|y, z) f_{Y|Z}(y|z) f_Z(z)$$

Conditional Expectation

- For a continuous random variable X, its conditional expectation $\mathbf{E}[X|A]$ given an event A with $P(A) > 0$ is

$$\mathbf{E}[X|A] = \int_{-\infty}^{\infty} x f_{X|A}(x) dx$$

- The conditional expectation of X given that $Y = y$ is

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

The expected value rule

- For a function $g(X)$, we have

$$\mathbf{E}[g(X)|A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x) dx$$

and

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

Total expectation theorem 1

- Let $A_1, \cdots, A_n$ be disjoint events that form a partition of the sample space, and assume that $P(A_i) > 0$ for all i. Then

$$f_X(x) = \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x).$$

- From here, we can get

$$\mathbf{E}[X] = \sum_{i=1}^{n} P(A_i)\mathbf{E}[X|A_i].$$

Total expectation theorem 2

- When conditioned on a random variable, we have

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

Proof

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

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

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

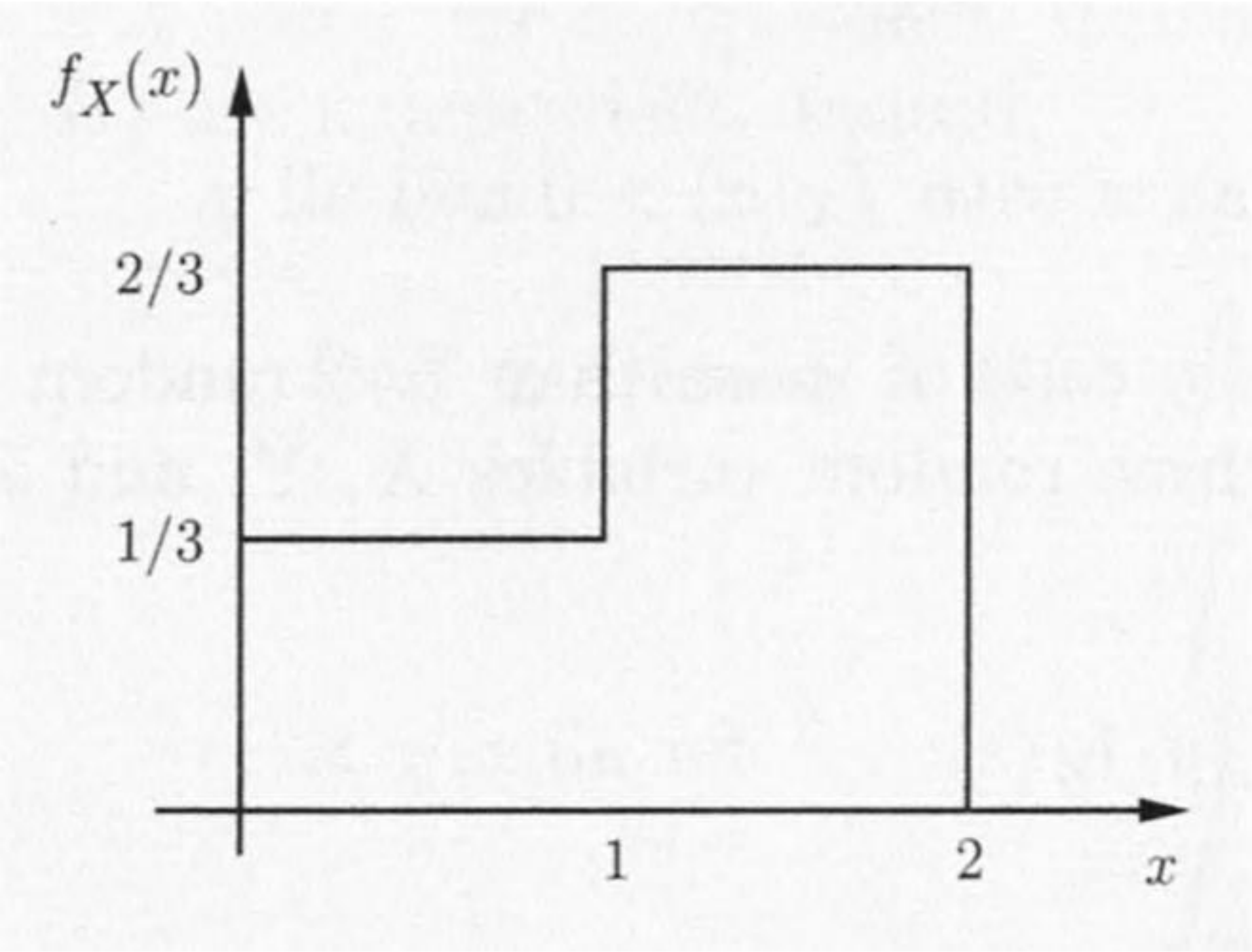
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X,Y}(x,y)\, dx dy$$

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

Example: Mean and Variance of a Piecewise Constant PDF

- Suppose the random variable X has the PDF
$$f_X(x) = \begin{cases} 1/3 & \text{if } 0 \leq x \leq 1 \\ 2/3 & \text{if } 1 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$
- Consider the events
$$A_1 = \{X \text{ lies in the first interval } [0,1]\}$$
$$A_2 = \{X \text{ lies in the second interval } (1,2]\}$$

Example: Mean and Variance of a Piecewise Constant PDF



Example: Mean and Variance of a Piecewise Constant PDF

- Then
$$P(A_1) = 1/3, P(A_2) = 2/3.$$
- And the conditional PDFs $f_{X|A_1}$ and $f_{X|A_2}$ are uniform.
- Recall previous result: Uniform random variable Y on $[a, b]$ has $\mathbf{E}[Y^2] = \frac{a^2+ab+b^2}{3}$.
- Thus $\mathbf{E}[X|A_1] = 1/2,\ \ \mathbf{E}[X|A_2] = 3/2$
$$\mathbf{E}[X^2|A_1] = 1/3, \mathbf{E}[X^2|A_2] = 7/3$$

Example: Mean and Variance of a Piecewise Constant PDF

- And

$$\mathbf{E}[X] = P(A_1)\mathbf{E}[X|A_1] + P(A_2)\mathbf{E}[X|A_2]$$

$$= \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{3}{2} = \frac{7}{6}.$$

$$\mathbf{E}[X^2] = P(A_1)\mathbf{E}[X^2|A_1] + P(A_2)\mathbf{E}[X^2|A_2]$$

$$= \frac{1}{3}\cdot\frac{1}{3} + \frac{2}{3}\cdot\frac{7}{3} = \frac{15}{9}.$$

- Thus, the variance is

$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{15}{9} - \frac{49}{36} = \frac{11}{36}.$$

-

Independence

- Two continuous random variables X and Y are independent if their joint PDF is the product of the marginal PDFs

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

for all x, y.

- It is equivalent to

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

for all y with $f_Y(y) > 0$ and all x.

Independence

- Three continuous random variables X, Y and Z are independent if
$$f_{X,Y,Z}(x, y, z) = f_X(x)f_Y(y)f_Z(z)$$
for all x, y, z.

Example: Independent normal random variables

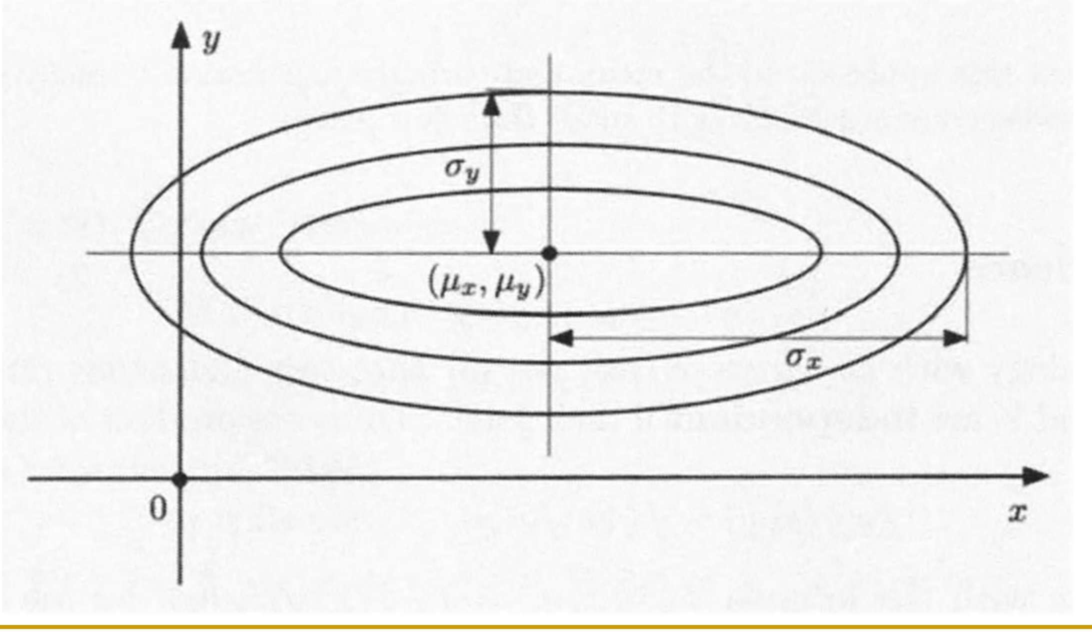
- Let X and Y be independent normal random variables with means μ_x, μ_y, and variance σ_x^2, σ_y^2, respectively. Their joint PDF is of the form

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

Example: Independent normal random variables

The ellipses are the contours of the PDF:

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



Independence

- If X and Y are independent, then any two events of the form $\{X \in A\}$ and $\{Y \in B\}$ are independent.
- $P(X \in A \text{ and } Y \in B)$

$$= \int_{x \in A} \int_{y \in B} f_{X,Y}(x, y) dy\, dx$$

$$= \int_{x \in A} \int_{y \in B} f_X(x) f_Y(y) dy\, dx$$

$$= \int_{x \in A} f_X(x) dx \int_{y \in B} f_Y(y) dy$$

$$= P(X \in A) P(Y \in B)$$

Independence

- In particular, when $A = (X \leq x)$, $B = (Y \leq y)$:

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

- Thus $F_{X,Y}(x,y) = F_X(x)F_Y(y)$.

Independence

- If X and Y are independent, then

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$$

More generally, for any two functions g, h

$$\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)]\mathbf{E}[h(Y)]$$

- Also, if X and Y are independent, then

$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$$

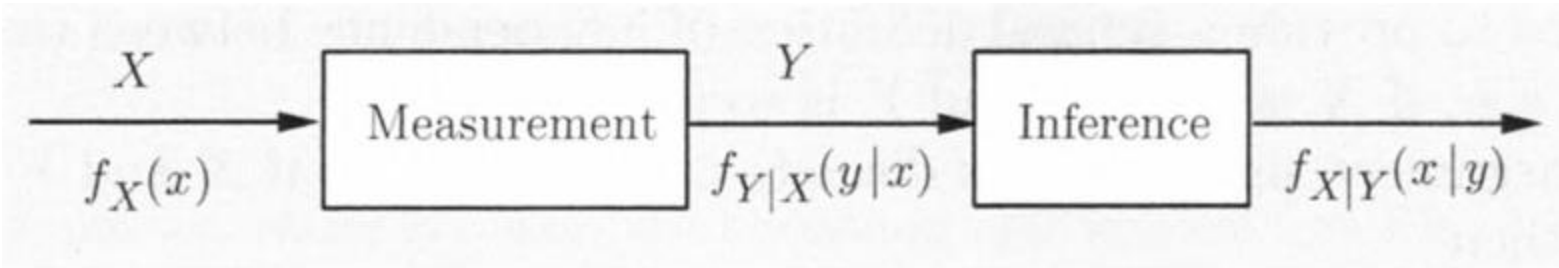
Content

- Continuous Random Variables and PDFs
- Cumulative Distribution Functions
- Normal Random Variables
- Joint PDFs of Multiple Random Variables
- Conditioning
- The Continuous Bayes' Rule

The continuous Bayes' rule

- In many situations, we represent an unobserved phenomenon by a random variable X with PDF f_X.
- 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?

The continuous Bayes' rule



- The information is provided by the conditional PDF $f_{X|Y}(x|y)$. By $f_X f_{Y|X} = f_{X,Y} = f_Y f_{X|Y}$, it follows that

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

The continuous Bayes' rule

- Based on the normalization property

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

 an equivalent expression is

$$f_{X|Y}(x|y) = \frac{f_X(x)f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(x')f_{Y|X}(y|x')dx'}$$

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

Example: Light bulbs

- A light bulb is known to have an exponentially distributed lifetime Y.
- However, the company has been experiencing quality control problems: On any given day, the parameter λ of the PDF of Y is a uniform random variable Λ on $[1, 3/2]$.
- We test a light bulb and record its lifetime.
- *Question: What can we say about λ?*
  - What is $f_{\Lambda|Y}(\lambda|y)$?

Example: Light bulbs

- The parameter λ in terms of a uniform random variable Λ with PDF

$$f_\Lambda(\lambda) = 2, \qquad \text{for } 1 \leq \lambda \leq 3/2$$

- Then by continuous Bayes' rule, for $1 \leq \lambda \leq \frac{3}{2}$

$$f_{\Lambda|Y}(\lambda|y) = \frac{f_\Lambda(\lambda) f_{Y|\Lambda}(y|\lambda)}{\int_{-\infty}^{\infty} f_\Lambda(t) f_{Y|\Lambda}(y|t)\,dt}$$

$$= \frac{2\lambda e^{-\lambda y}}{\int_1^{3/2} 2te^{-ty}\,dt}$$

.

Inference about a discrete random variable

- In some cases, the unobserved phenomenon is inherently discrete.
- *Example*. Consider a binary signal which is observed in the presence of normally distributed noise.
- *Example.* Consider a medical diagnosis that is made on the basis of continuous measurements, such as temperature and blood counts.

Inference about a discrete random variable

- Instead of working with the conditioning event $\{Y = y\}$, which has zero probability, let us first condition on the event $\{y \leq Y \leq y + \delta\}$, then take the limit as $\delta \to 0$.

$$P(A|Y = y) \approx P(A|y \leq Y \leq y + \delta)$$
$$= \frac{P(A)P(y \leq Y \leq y + \delta)}{P(y \leq Y \leq y + \delta)} \approx \frac{P(A)f_{Y|A}(y)\delta}{f_Y(y)\delta}$$
$$= \frac{P(A)f_{Y|A}(y)}{f_Y(y)}$$

.

Inference about a discrete random variable

- The denominator $f_Y(y)$ can be evaluated by total probability theorem

$$f_Y(y) = P(A) f_{Y|A}(y) + P(A^c) f_{Y|A^c}(y)$$

so that

$$P(A|Y = y) = \frac{P(A) f_{Y|A}(y)}{P(A) f_{Y|A}(y) + P(A^c) f_{Y|A^c}(y)}$$

Inference about a discrete random variable

- Consider an event A of the form $\{N = n\}$
- N is a discrete random variable with PMF p_N
- Let Y be a continuous random variable which is described by a conditional PDF $f_{Y|N}(y|n)$.

$$P(N = n|Y = y) = \frac{p_N(n) f_{Y|N}(y|n)}{f_Y(y)}$$

$$= \frac{p_N(n) f_{Y|N}(y|n)}{\sum_i p_N(i) f_{Y|N}(y|i)}$$

Example: Signal Detection

- A binary signal S is transmitted with
  $$P(S = 1) = p, \qquad P(S = -1) = 1 - p$$
- The received signal is $Y = N + S$, where N is standard normal noise.
- What is the prob. that $S = 1$, as a function of the observed value y of Y?
- Conditional on $S = s$, the random variable Y has a normal distribution with mean s and variance 1.

Example: Signal Detection

$$
\begin{aligned}
P(S = 1|Y = y) &= \frac{p_S(1) f_{Y|S}(y|1)}{f_Y(y)} \\
&= \frac{\frac{p}{\sqrt{2\pi}} e^{-(y-1)^2/2}}{\frac{p}{\sqrt{2\pi}} e^{-(y-1)^2/2} + \frac{1-p}{\sqrt{2\pi}} e^{-(y+1)^2/2}} \\
&= \frac{pe^{y}}{pe^{y} + (1-p)e^{-y}}
\end{aligned}
$$

Example: Signal Detection

- Notice that

$$\lim_{y \to -\infty} P(S = 1 | Y = y) = 0$$
$$\lim_{y \to \infty} P(S = 1 | Y = y) = 1$$

which is consistent with intuition.