Reasoning Under Uncertainty
Probability theory

References:

- Mathematical methods in artificial intelligence, Bender, Chapter 7.
- Expert systems: Principles and programming, Giarratano and Riley, pag. 183-243.
- Artificial intelligence: A modern approach, Russell and Norvig, pg. 413-470.

Review of Probability and Statistics

- Probability
  - Definition of probability
  - Axioms and properties
  - Conditional probability
  - Bayes theorem
- Random Variables
  - Definition
  - Probability density function
  - Statistical characterization of random variables
- Random Vectors
  - Mean vector
  - Covariance matrix
- The Gaussian random variable

Experiment

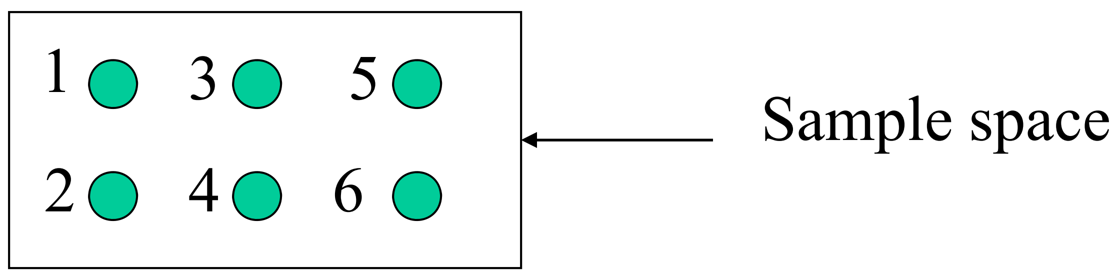
- An experiment is a process for which the outcome is not known with certainty.
- Examples
  - Rolling a fair six-sided die.
  - Randomly choosing 10 transistors from a lot of 1000.
  - Selecting a newborn child at a known hospital.

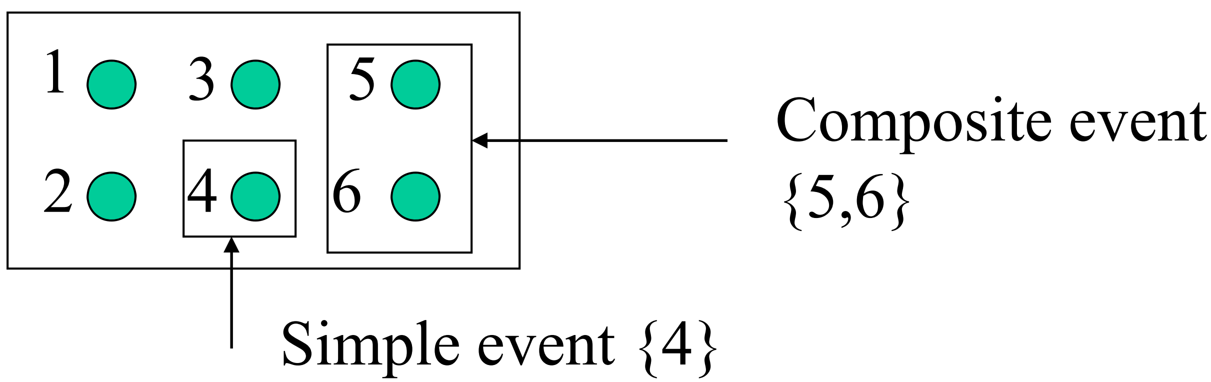
Event

- An event is an outcome or combination of outcomes from a statistical experiment.
- Examples
  - Obtaining a 6 when a die is rolled.
  - Obtaining an even number when a die is rolled.
  - Selecting a newborn child in a hospital weighing more than 8 pounds.

Sample Space

- The sample space is the event consists of all possible outcomes of a statistical experiment.

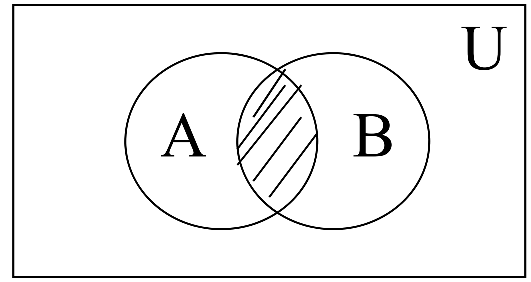




Relations between events in a sample space

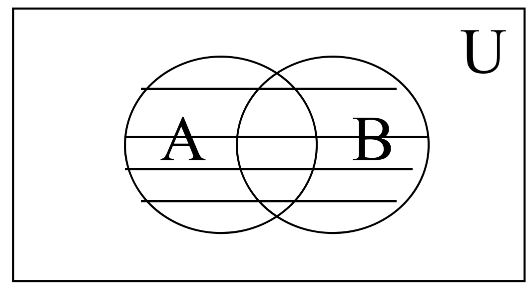
- Intersección of events

  $A \cap B$



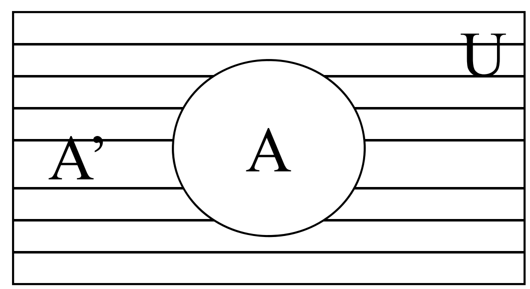
- Union of events

  $A \cup B$



- Complement of an event

  A’



Theory of Probability

- The 3 axioms of probability
- Axiom 1: $0 \le P(E) \le 1$
- Axiom 2: $P(\text{Universe})=1$ $P(\emptyset)=0$
  $P(\text{True})=1$ $P(\text{False})=0$
- Axiom 3: If $\{A \cap B = \emptyset\}$ then
  $P(A \cup B)=P(A)+P(B)$

Corollary (Axiom 2 and 3): $P(A)+P(\sim A)=1$

Probability of Composite events

$U=\{1,2,3,4,5,6\}$

$P\{1\}=P\{2\}=P\{3\}=P\{5\}=P\{6\}=1/6$

$A=\{2,4,6\}$ $B=\{3,6\}$

$A \cap B=\{6\}$

$P(A \cap B)=P(\{6\})=1/6$

Independent Events

$P(X \cap Y)=P(X)P(Y)$

Conditional Probability

$$P(A|B)=P(A\cap B)/P(B) \qquad P(B)\neq 0$$

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

General Rule

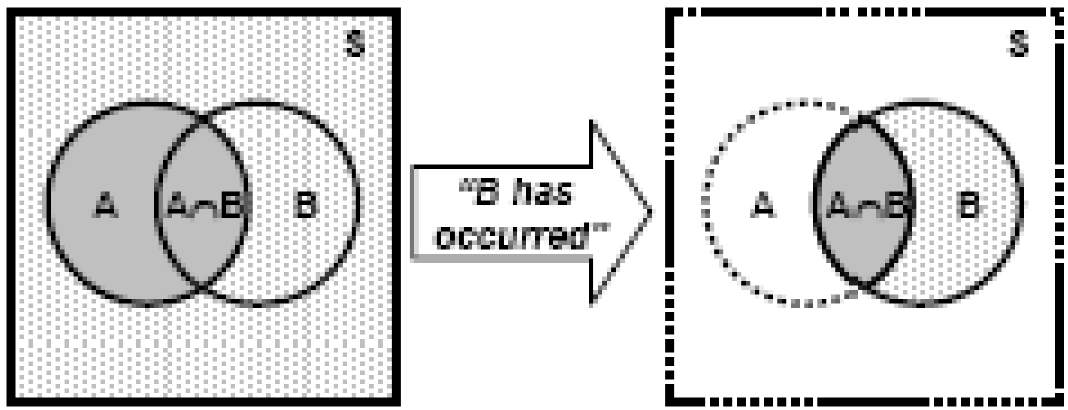
$$P(A_1\cap A_2\cap\ldots\cap A_N)=P(A_1|A_2\cap\ldots\cap A_N)*P(A_2|A_3\cap\ldots\cap A_N)*\ldots P(A_{N-1}|A_N)*P(A_N)$$

Independent Events

$$P(A|B)=P(A)$$
$$\Leftrightarrow P(B|A)=P(B)$$
$$\Leftrightarrow P(A\cap B)=P(A)*P(B)$$

Diagram: conditional probability

- **This conditional probability P[A|B] is read:**
  - the "conditional probability of A conditioned on B", or simply
  - the "probability of A given B"

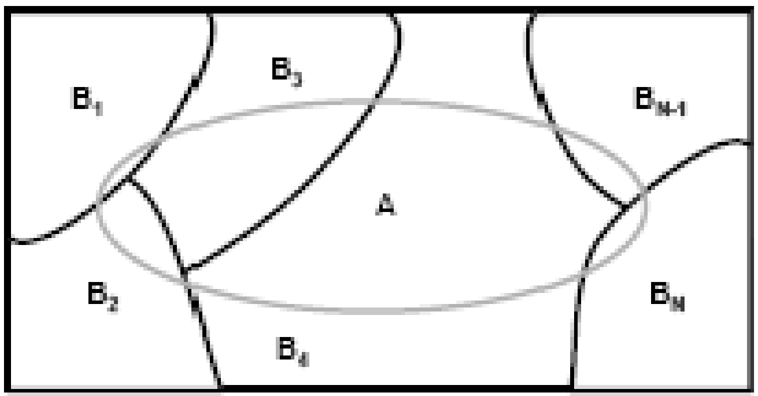


- **Interpretation**
  - The new evidence "*B has occurred*" has the following effects
    - The original sample space S (the whole square) becomes B (the rightmost circle)
    - The event A becomes A∩B
  - P[B] simply re-normalizes the probability of events that occur jointly with B

Theorem of total probability

- Let B_1, B_2, ..., B_N be mutually exclusive events whose union equals the sample space **S**. We refer to these sets as a partition of **S**.
- An event **A** can be represented as:

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \ldots \cup B_N) = (A \cap B_1) \cup (A \cap B_2) \cup \ldots (A \cap B_N)$$



- Since B_1, B_2, ..., B_N are mutually exclusive, then

$$P[A] = P[A \cap B_1] + P[A \cap B_2] + \ldots + P[A \cap B_N]$$

- and, therefore

$$P[A] = P[A \mid B_1]P[B_1] + \ldots P[A \mid B_N]P[B_N] = \sum_{k=1}^{N} P[A \mid B_k]P[B_k]$$

Bayes Theorem

- **Given B_1, B_2, …, B_N, a partition of the sample space S. Suppose that event A occurs; what is the probability of event B_j?**
  - Using the definition of conditional probability and the Theorem of total probability we obtain

$$P[B_j \mid A] = \frac{P[A \cap B_j]}{P[A]} = \frac{P[A \mid B_j] \cdot P[B_j]}{\sum_{k=1}^{N} P[A \mid B_k] \cdot P[B_k]}$$

- **This is known as Bayes Theorem or Bayes Rule, and is (one of) the most useful relations in probability and statistics**
  - Bayes Theorem is definitely the fundamental relation in Statistical Pattern Recognition

Rev. Thomas Bayes (1702-1761)

Bayes Theorem and Statistical Pattern Recognition

- Bayes decision making – choosing the most likely class, given the value of the feature or features.
- Feature value – x, Class –C
- P(x) probability distribution for x in the entire population (normalization constant that does not affect the decision)
- P(C) – prior probability that a random sample is from class C (prior)
- P(x|C) – conditional probability of obtaining feature value x given sample is from class C (likelihood)
- P(C|x) – have to estimate that given a sample has feature value x, it is from class C (posterior)
- We know – probability of the joint event that a sample comes from class C and has feature value x is
- $P(C \cap x)=P(C)P(x|C)=P(x)P(C|x)$
- Rearranging
- $P(C|x)=P(C)P(x|C) / P(x)$

Law of Total Probability and Bayes Theorem

- Law of total probability states that if an event A can occur in m different ways: B1,B2,…,Bm (m subevents are mutually exclusive-cannot occur at the same time), the probability of y occuring is the sum of the probabilities of the subevents xi

$$P(y) = \sum_{x \in X} P(x \cap y)$$

- From definition of conditional probability P(A|B)
- (Bj∩A)=P(A|Bj)P(Bj)
- Then
- P(Bj|A)=P(A|Bj)P(Bj)/ΣP(A|Bj)P(Bj)
- Posterior = likelihood x prior / evidence

Example

- **Consider a clinical problem where we need to decide if a patient has a particular medical condition on the basis of an *imperfect* test:**
  - Someone with the condition may go undetected (*false-negative*)
  - Someone free of the condition may yield a positive result (*false-positive*)
- **Nomenclature**
  - The true-negative rate *P(NEG|¬COND)* of a test is called its SPECIFICITY
  - The true-positive rate *P(POS|COND)* of a test is called its SENSITIVITY

	TEST IS POSITIVE	TEST IS NEGATIVE	ROW TOTAL
HAS CONDITION	*True-positive* *P(POS|COND)*	*False-negative* *P(NEG|COND)*	
FREE OF CONDITION	*False-positive* *P(POS|¬COND)*	*True-negative* *P(NEG|¬COND)*	
COLUMN TOTAL			

- **PROBLEM**
  - Assume a population of **10,000** where **1** out of every 100 people has the condition
  - Assume that we design a test with **98%** specificity and **90%** sensitivity
  - Assume you are required to take the test, which then yields a POSITIVE result
  - **What is the probability that you have the condition**?
    - SOLUTION A: Fill in the joint frequency table above
    - SOLUTION B: Apply Bayes rule

Solution

- **Consider a clinical problem where we need to decide if a patient has a particular medical condition on the basis of an *imperfect* test:**
  - Someone with the condition may go undetected (*false-negative*)
  - Someone free of the condition may yield a positive result (*false-positive*)
- **Nomenclature**
  - The true-negative rate *P(NEG|¬COND)* of a test is called its SPECIFICITY
  - The true-positive rate *P(POS|COND)* of a test is called its SENSITIVITY

	TEST IS POSITIVE	TEST IS NEGATIVE	ROW TOTAL
HAS CONDITION	*True-positive* *P(POS|COND)* 100×0.90	*False-negative* *P(NEG|COND)* 100×(1-0.90)	100
FREE OF CONDITION	*False-positive* *P(POS|¬COND)* 9,900×(1-0.98)	*True-negative* *P(NEG|¬COND)* 9,900×0.98	9,900
COLUMN TOTAL	288	9,712	10,000

- **PROBLEM**
  - Assume a population of **10,000** where **1** out of every 100 people has the condition
  - Assume that we design a test with **98%** specificity and **90%** sensitivity
  - Assume you are required to take the test, which then yields a POSITIVE result
  - **What is the probability that you have the condition**?
    - SOLUTION A: Fill in the joint frequency table above
    - SOLUTION B: Apply Bayes rule

Solution using Bayes Theorem

- **SOLUTION B: Apply Bayes theorem**

$$P[COND \mid POS] =$$

$$= \frac{P[POS \mid COND] \cdot P[COND]}{P[POS]} =$$

$$= \frac{P[POS \mid COND] \cdot P[COND]}{P[POS \mid COND] \cdot P[COND] + P[POS \mid \neg COND] \cdot P[\neg COND]} =$$

$$= \frac{0.90 \cdot 0.01}{0.90 \cdot 0.01 + (1 - 0.98) \cdot 0.99} =$$

$$= 0.3125$$

Exercise in group (supplement problem)

- As an example in which two events are not independent, suppose that a company plans to test two new techniques for improving the extraction of oil from the ground. The first technique consists fo setting off an explosion at the bottom of a well to fracture the strata and then testing seismically to determine the extent of the fracturing. The second technique consists of injecting hot brine into the well to loosen the oil and then pumping to measure the oil recovery. Let E be the event that the explosion successfully fractures the strate within a radius of 100 meters, and let R be the event that oil can be recovered at a rate of 50 barrels per day after pumping in hot brine. In a certain region, $P(E)$ is estimated to be 0.8. If E occurs, the probability of R occuring is estimated to be $P(R|E)=0.9$, but if the explosion is not a success, the estimated probability of recovery isonly $P(R|\sim E)=0.3$. These 3 assumptions are sufficient to define the probabilities of any combination of outcomes, given any set of constraints.

- What is the probability that the explosion and the brine tests are both successful or both fail?
- What is the probability that the explosion test was successful, give the constraint that only one of the two tests was successful?
- What is the probability that the explosion was successful, given that recovery using brine was successful?
- If one or more of the tests was successful, what is the probability that the other one was successful?

Random Variables

- Example: A fair coin is tossed 3 times.

S={(TTT),(TTH),(THT),(HTT),(HHT),(HTH),(THH),(HHH)}

Random Variables

- Let X be the number of heads tossed in 3 tries.

$X(TTT)=$ $X(HHT)=$

$X(TTH)=$ $X(HTH)=$

$X(THT)=$ $X(THH)=$

$X(HTT)=$ $X(HHH)=$

- So $P(X=0)=$ $P(X=1)=$

$P(X=2)=$ $P(X=3)=$

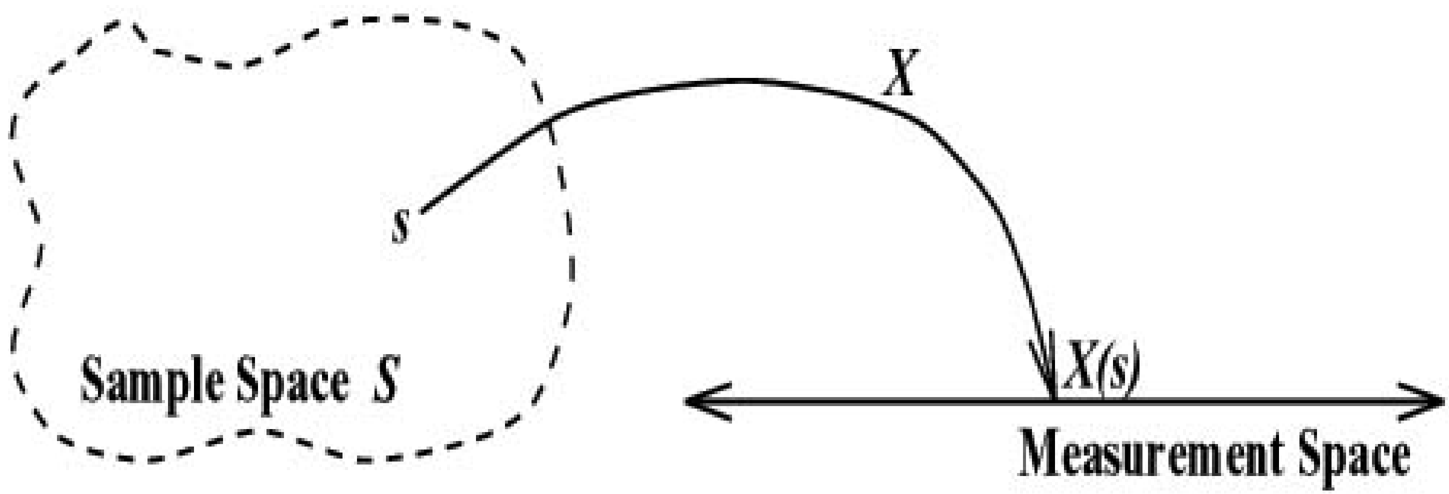
Random Variable as a Measurement

- A chemical reaction.
- A laser emitting photons.
- A packet arriving at a router.

- Sample space here is a set here with noncountable infinite number of elements

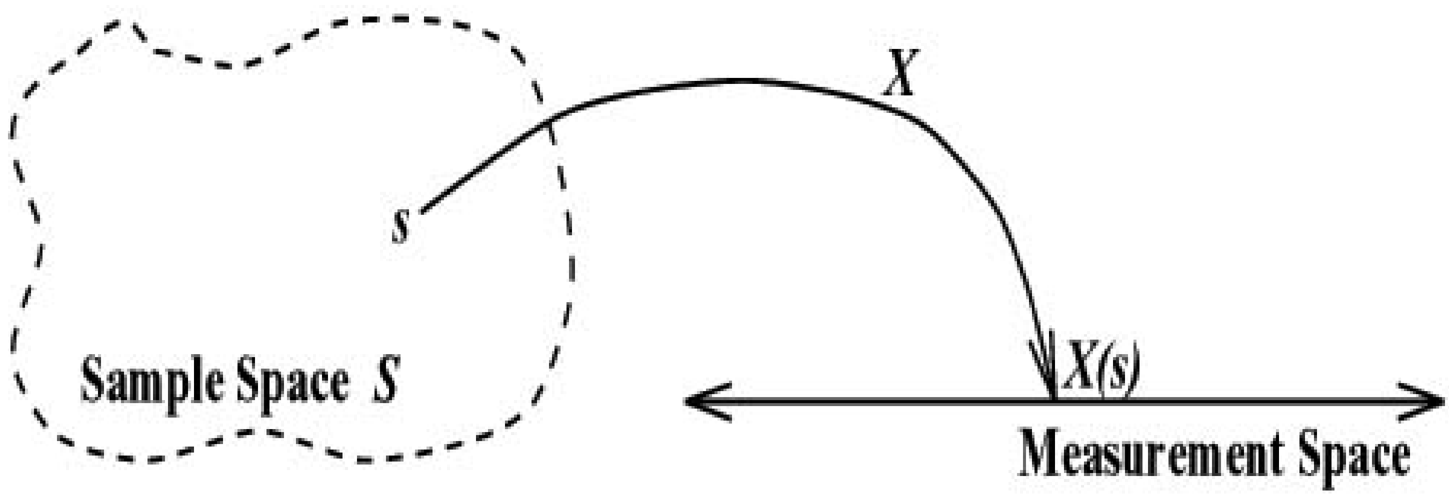
Random Variable as a Measurement

- Thus a random variable can be thought of as a measurement on an experiment



Random Variable as a Measurement

- Thus a random variable can be thought of as a measurement on an experiment



Probability Density Function (pdf)

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

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

Properties of the pdf

- Positivity

$$f_X(x) \geq 0$$

- Integral over all x (unit area)

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

More properties for the pdf

- $f_X(x)$ can be used to calculate probabilities

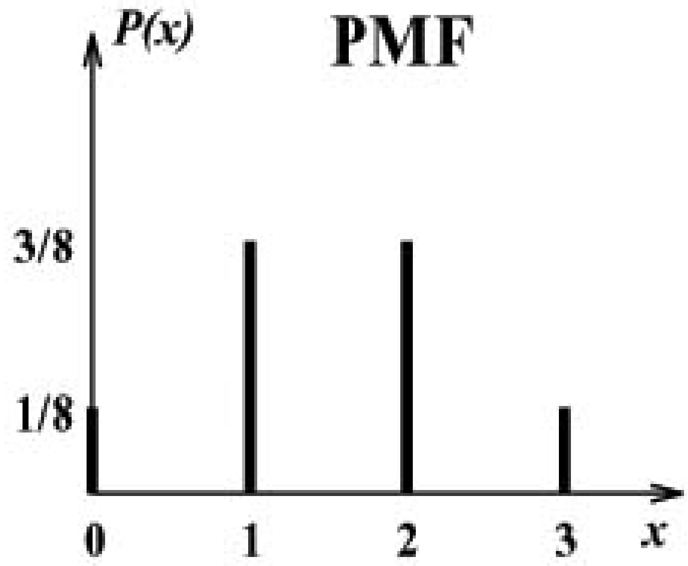
$$P\{x_1 < X \le x_2\} = \int_{x_1^+}^{x_2^+} f_X(x)\,dx$$

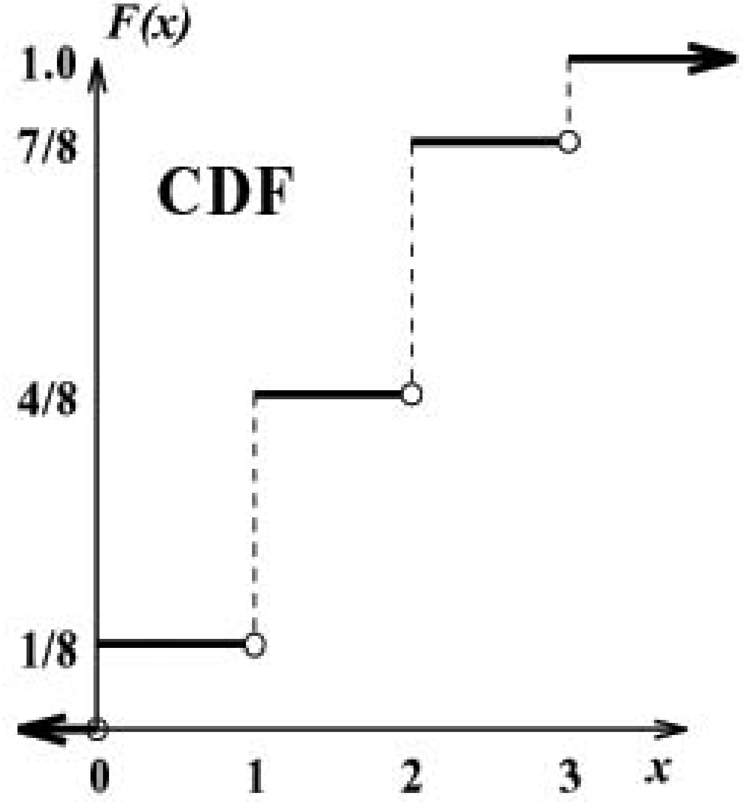
$$P\{x_1 \le X < x_2\} = \int_{x_1^-}^{x_2^-} f_X(x)\,dx$$

$$P\{x_1 \le X \le x_2\} = \int_{x_1^-}^{x_2^+} f_X(x)\,dx$$

$$P\{x_1 < X < x_2\} = \int_{x_1^+}^{x_2^-} f_X(x)\,dx$$

PMF and CDF for the 3 Coin Toss Example: X= # of heads





What would be the pdf for this example?

Types of Random Variables

- Discrete: pdf consist only of impulses. CDF has the shape of a staircase
- Continuous: CDF is a continuous function or, equivalently, pdf has no impulses.
- Mixed: neither continuous nor discrete

Types of Random Variables

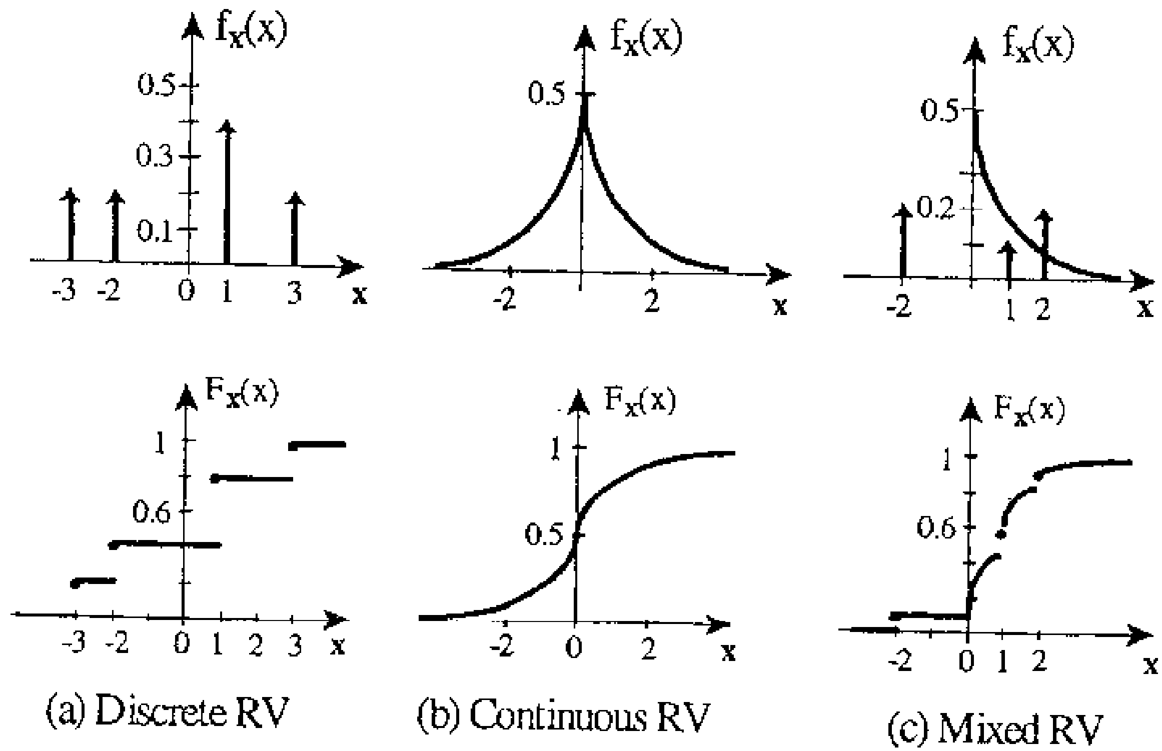
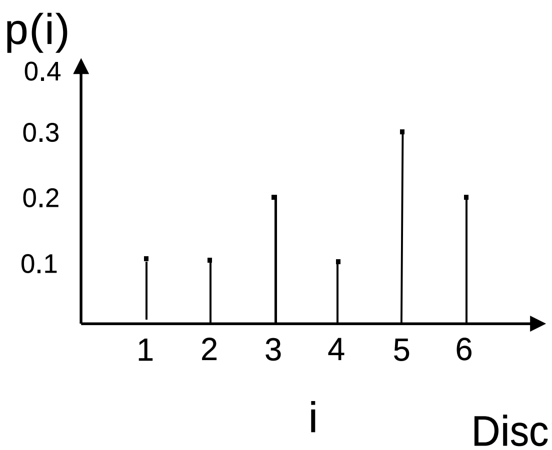
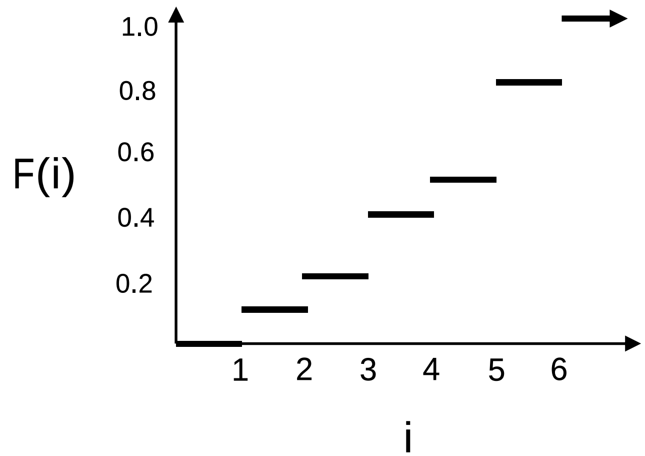


Figure 2.2 Examples of (a) discrete (b) continuous and (c) mixed random variables.

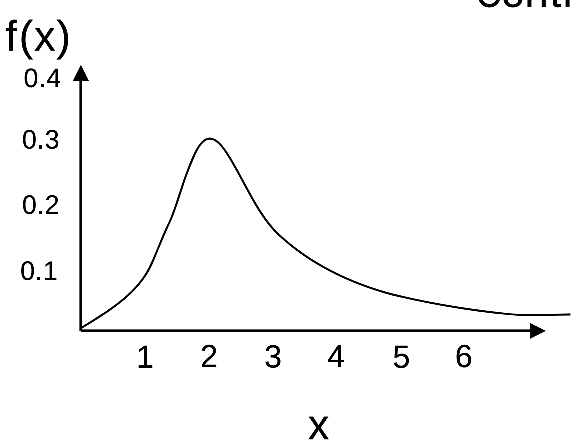
CMF, PMF, pdf

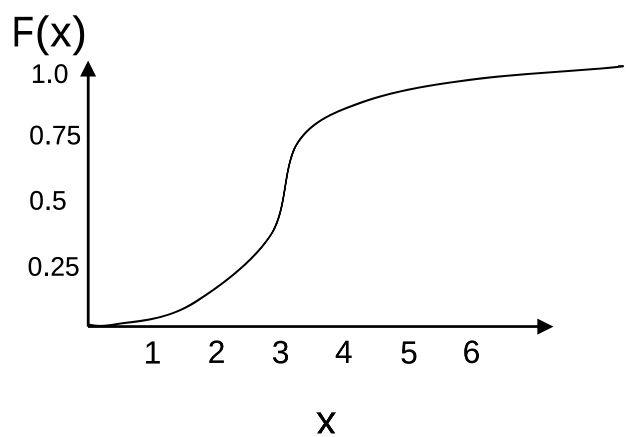




Discrete

Continuous





Statistical Characterizations

- Expectation (Mean Value, First Moment):

$$\eta_x = E\{X\} = \int_{-\infty}^{\infty} x\, f_x(x)dx$$

Statistical Characterizations

- Variance of X:

$$\sigma_x^2 = E\left\{\left(X - \eta_X\right)^2\right\} = \int_{-\infty}^{\infty} \left(x - \eta_X\right)^2 f_X\left(x\right) dx$$

$$= E\left\{X^2\right\} - \left(\eta_X\right)^2$$

$$\sigma_x = \sqrt{\sigma_x^2}$$

Higher Order Statistics

- k-th moment of

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

- k-th central moment of X:

$$\mu_k = E\{(X - \eta_X)^k\} = \int_{-\infty}^{\infty} (x - \eta_X)^k f_X(x)dx$$

Conditional Cumulative Functions

- Cumulative distribution function

$$F_X(x/C) = P\{X \le x/C\} = \frac{P\{X \le x, C\}}{P\{C\}}$$

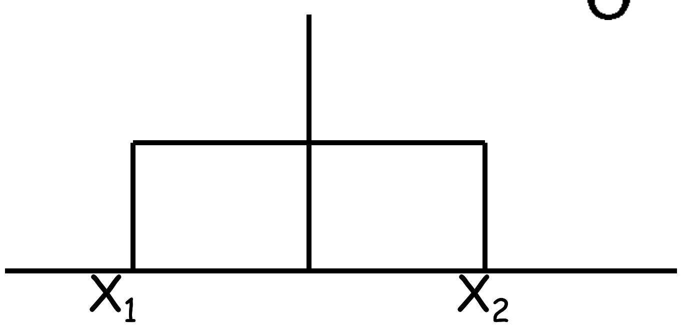
- Probability density function

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

Uniform pdf

- A R.V. X that is uniformly distributed between x_1 and x_2 has density function:

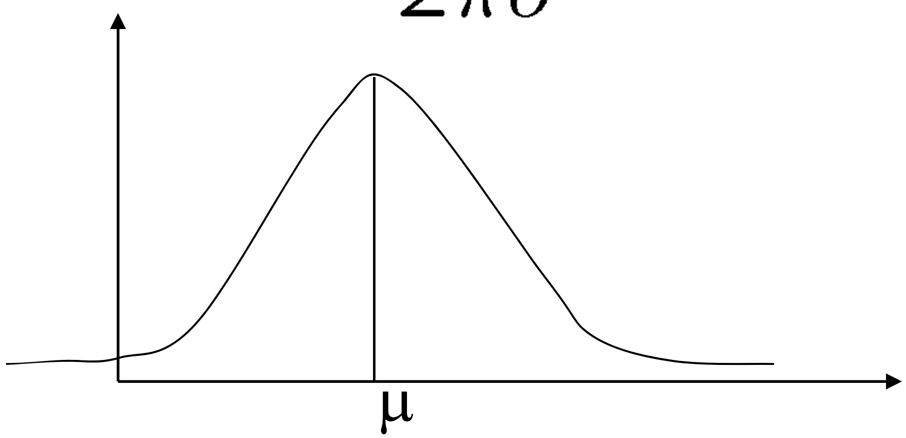
$$p_X(x) = \begin{cases} \frac{1}{x_2 - x_1} & x_1 \le x \le x_2 \\ 0 & otherwise \end{cases}$$



Gaussian pdf

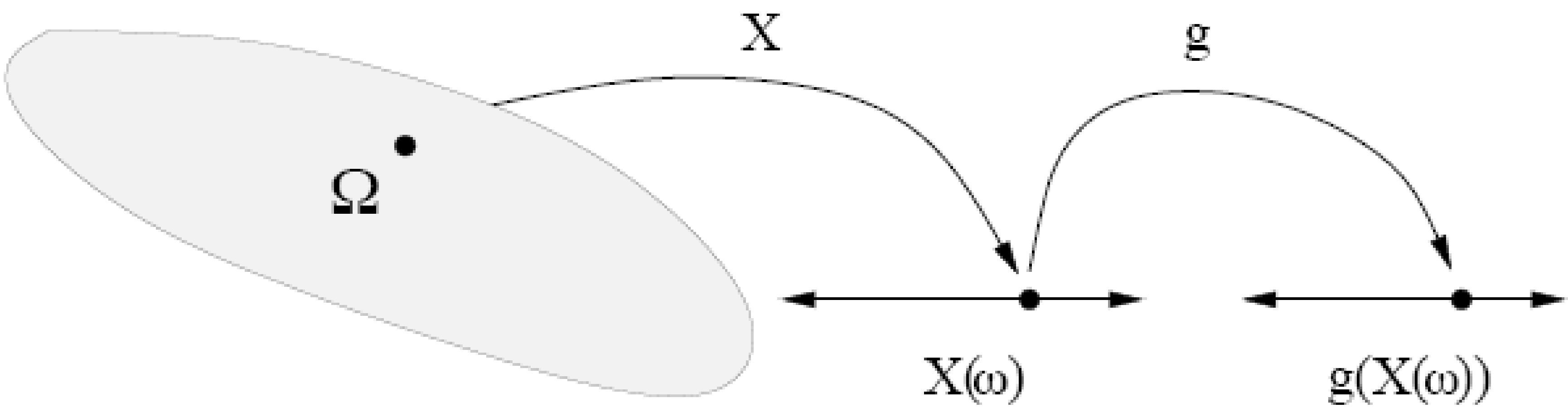
- A R.V. X that is normally distributed has density function:

$$p_X(x) = \frac{1}{2\pi\sigma} \exp -\frac{(x-\mu)^2}{2\sigma^2}$$



Transformations of Random Variables

$$Y = g(X)$$



Transformations of Random Variables

$$Y = g(X)$$

- The basic question is:
  - Knowing the statistical characterization of X, partial or total, what are the resulting characterizations of Y?

Two Random Variables

- Defined as real value functions, X (ζ) and Y(ζ), over the same sample space S

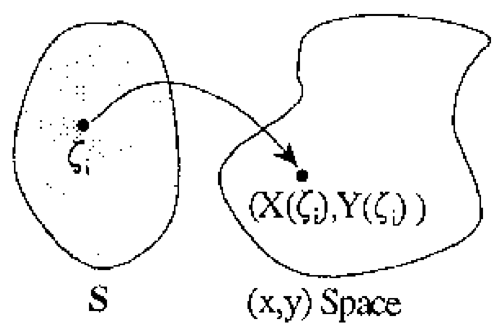


Figure 2.15 Random variables X and Y defined on the sample space S.

Total Characterization: Joint CDF

- The joint cumulative distribution function (JCDF) for a random variables *X and Y* is

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

Properties: Joint CDF

- Note that $F_{XY}(x,y)$ is bounded from the above and below

$$0 \leq F_{XY}(x,y) \leq 1$$

- Note that $F_{XY}(x,y)$ is non-decreasing in *x and y*, i.e.

$$F_{XY}(x_2, y) \geq F_{XY}(x_1, y) \quad \text{for all } x_2 > x_1, \text{ and all } y$$
$$F_{XY}(x, y_2) \geq F_{XY}(x, y_1) \quad \text{for all } y_2 > y_1, \text{ and all } x$$

Properties: Joint CDF

- $F_{XY}(x,y)$ is <u>continuous</u> from the right

$$\lim_{\substack{\varepsilon\to 0^+ \\ \delta\to 0^+}} F_{XY}(x+\varepsilon, y+\delta) = F_{XY}(x, y)$$

- $F_{XY}(x,y)$ can be used to <u>calculate probabilities</u>

$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = F_X(x_2, y_2) - F_X(x_1, y_1)$$

Properties of the Joint pdf

(1) **Positivity,**

$$f_{XY}(x, y) \geq 0 \quad \text{for all } x \text{ and } y \tag{2.83}$$

(2) **Integral over all x and y,**

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1 \tag{2.84}$$

(3) $f_{XY}(x, y)$ can be used to **calculate probability** of rectangular events as

$$P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = \int_{x_1^+}^{x_2^+} \int_{y_1^+}^{y_2^+} f_{XY}(x, y)\, dx\, dy \tag{2.85}$$

where x_1^+, x_2^+, y_1^+ and y_2^+ are limits from the positive sides,
or any event A as

$$P(\{(X, Y) \in A\}) = \int \int_A f_{XY}(x, y)\, dx\, dy \tag{2.86}$$

Partial Characterization

- Marginal densities
- Conditional densities
- Joint moments
  - Conditional means
  - Correlation, covariance
  - Higher order moments

Marginal Densities

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

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

Conditional pdf

$$f_X\left(x/y\right) = \frac{f_{XY}\left(x, y\right)}{f_Y\left(y\right)} = \frac{f_{XY}\left(x, y\right)}{\int_{-\infty}^{\infty} f_{XY}\left(x, y\right)dx}$$

$$f_Y\left(y/x\right) = \frac{f_{XY}\left(x, y\right)}{f_X\left(x\right)} = \frac{f_{XY}\left(x, y\right)}{\int_{-\infty}^{\infty} f_{XY}\left(x, y\right)dy}$$

$$f_{XY}\left(x, y\right) = f_Y\left(y/x\right)f_X\left(x\right) = f_X\left(x/y\right)f_Y\left(y\right)$$

Statistical Independence

$$f_X(x/y) = \frac{f_{XY}(x, y)}{f_Y(y)} = f_X(x)$$

$$f_Y(y/x) = \frac{f_{XY}(x, y)}{f_X(x)} = f_Y(y)$$

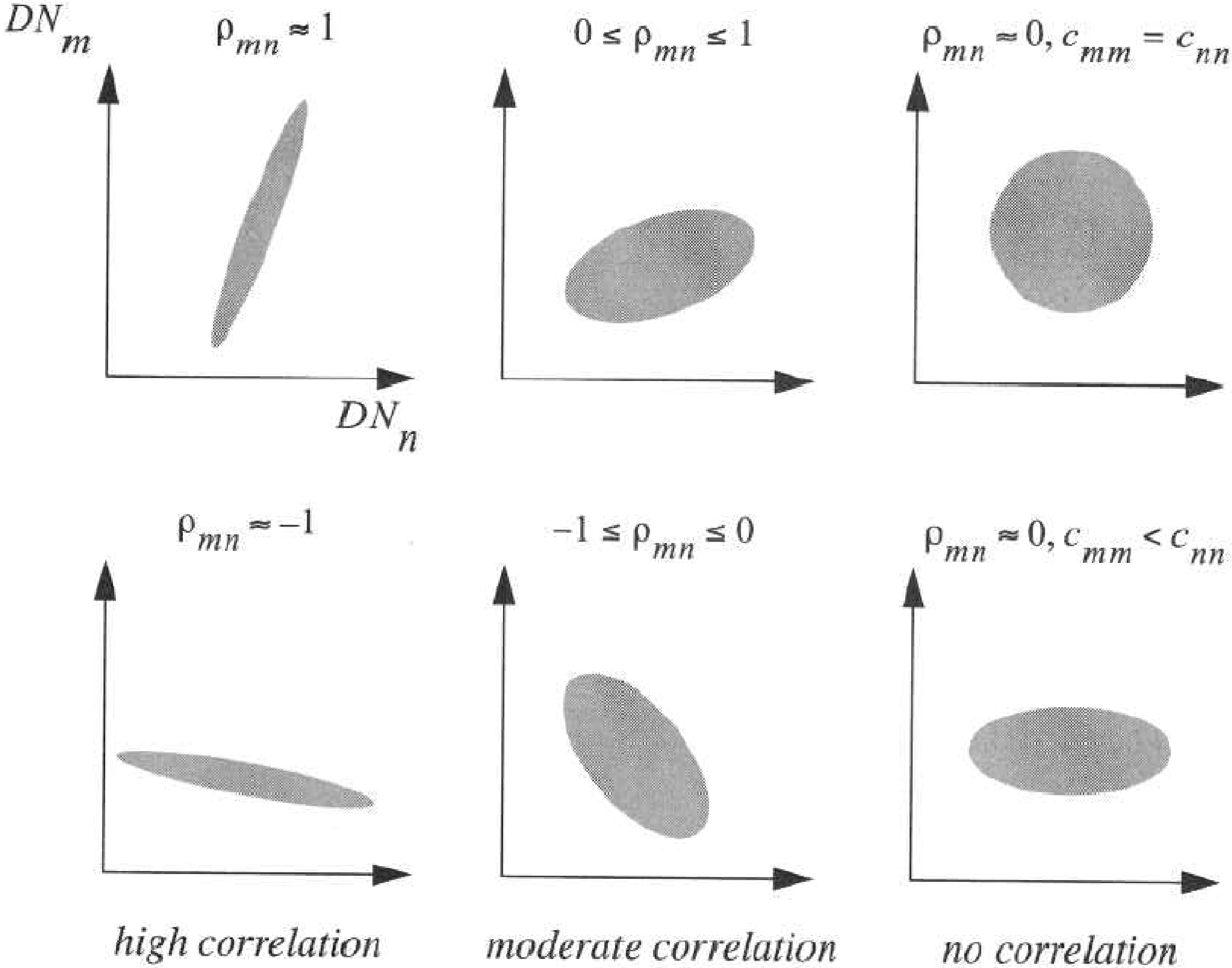
$$f_{XY}(x, y) = f_Y(y) f_X(x)$$

Correlation Coefficient ρ_{XY} (cont.)

$$-1 \leq \rho_{XY} \leq +1$$

- The closer ρ_{XY} is to -1 or +1 the more the random variables X and Y are said to be linearly related

Correlation Between 2 Variables



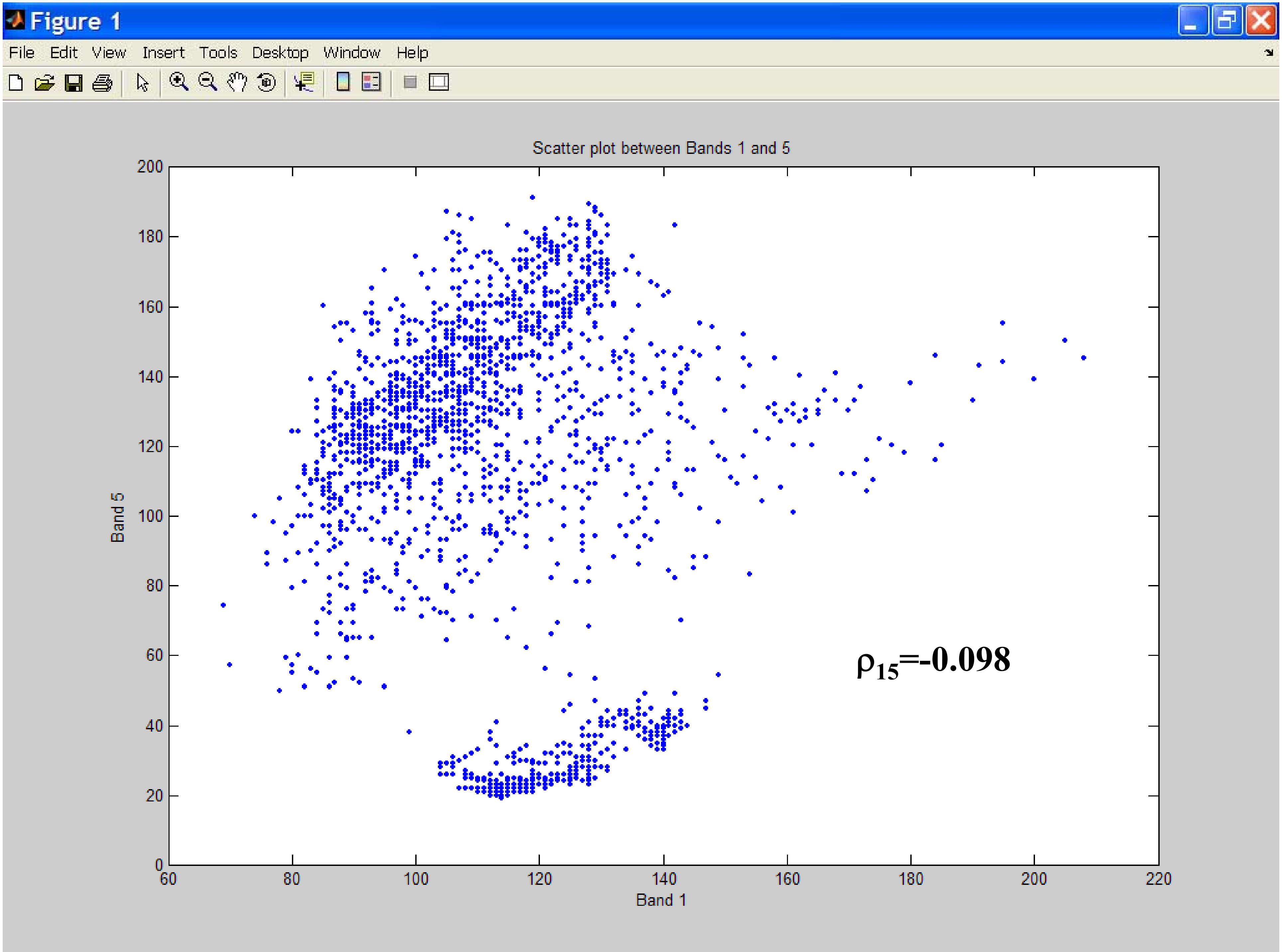


Figure 1
File Edit View Insert Tools Desktop Window Help
Scatter plot between Bands 1 and 5
Band 5
Band 1
200
180
160
140
120
100
80
60
40
20
0
60
80
100
120
140
160
180
200
220
ρ_{15}=-0.098

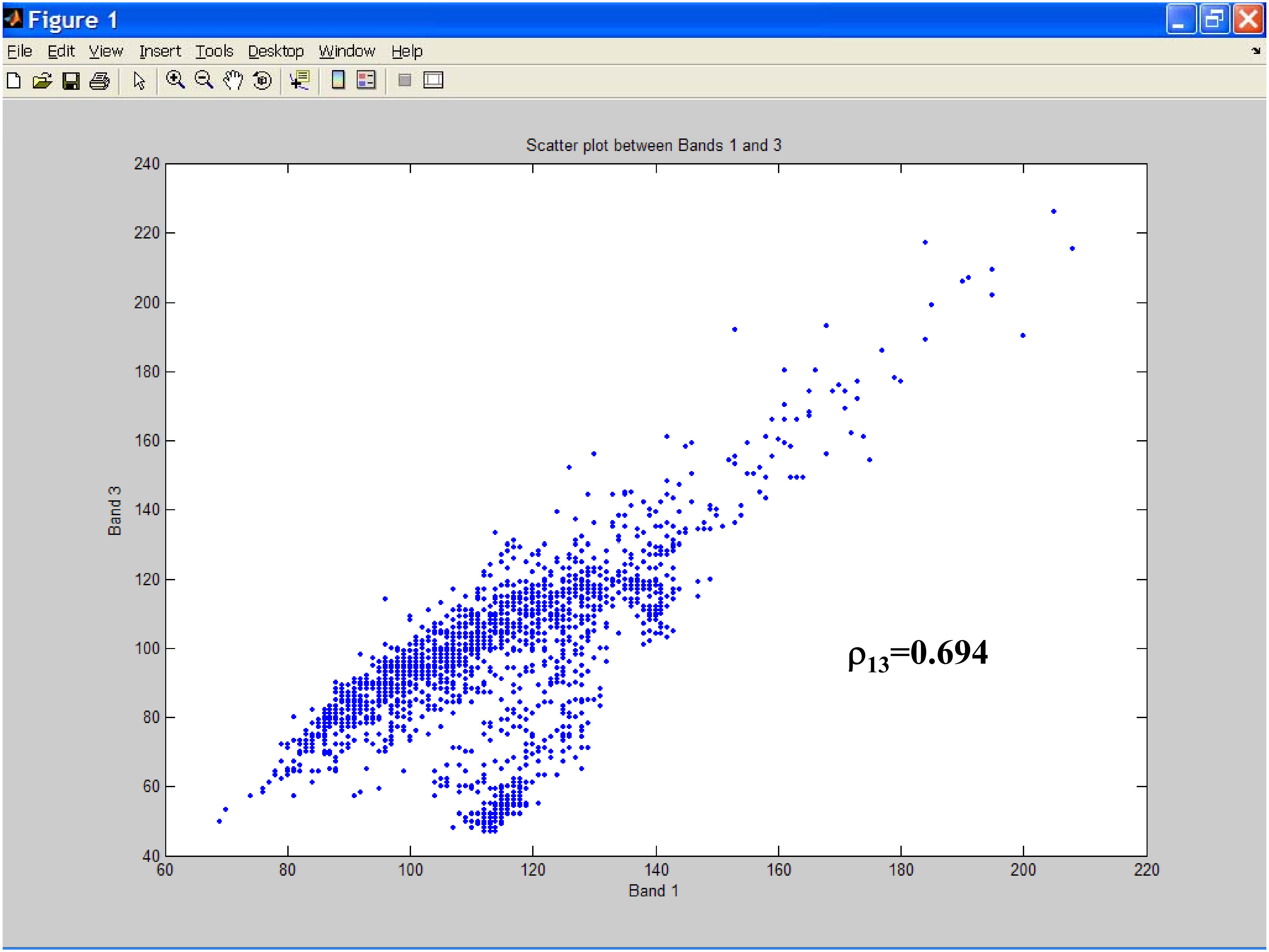


Figure 1
File Edit View Insert Tools Desktop Window Help
Scatter plot between Bands 1 and 3
Band 3
Band 1
240
220
200
180
160
140
120
100
80
60
40
60
80
100
120
140
160
180
200
220
ρ13=0.694

Correlation Coefficient ρ_{XY}

- Useful relations

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$$\sigma_{XY} = \rho_{XY} \sigma_X \sigma_Y$$

$$r_{XY} = \sqrt{1 - \rho_{XY}^2}$$

2.84 Given X and Y are random variables with joint probability density function $f_{XY}(x, y)$ as follows:

$$f(x, y) = \begin{cases} 1, & -y < x < y, 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

(a) Are X and Y independent random variables? Why or why not?

(b) Are X and Y uncorrelated random variables? Why or why not?

(c) Are X and Y orthogonal random variables? Why or why not?

Orthogonal, Uncorrelated, Independent

- Two random variables are defined as <u>uncorrelated</u> if

$$R_{XY} = E[X]\,E[Y] \text{ or } \sigma_{XY} = 0 \text{ or } \rho_{XY} = 0$$

- Notice that independence → uncorrelated (Why?)
- Two random variables are called <u>orthogonal</u> if $R_{XY}=0$ but it does not imply that $\sigma_{XY}=0$ or equivalently that $\rho_{XY}=0$
  - Need one mean =0 and the other finite.

Orthogonal, Uncorrelated, Independent

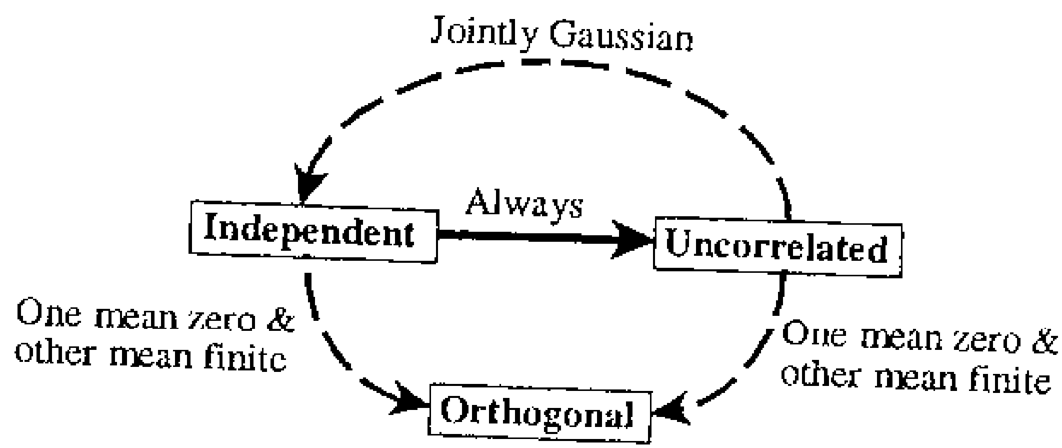


Figure 2.20 Relationship between the definitions of independent, uncorrelated, and orthogonal.

Jointly Gaussian Random Variables (JGRV)

- X and Y are jointly Gaussian if their joint pdf $f_{XY}(x,y)$ has the form:

$$f_{XY}(x,y) = \frac{1}{2\pi\left(\sigma_X \sigma_Y \sqrt{1-\rho_{XY}^2}\right)} \exp\left\{ \frac{-1}{2\left(1-\rho_{XY}^2\right)} \left[\frac{\left(x-\eta_X\right)^2}{\sigma_X^2} - \frac{2\rho_{XY}\left(x-\eta_X\right)\left(y-\eta_Y\right)}{\sigma_X \sigma_Y} + \frac{\left(y-\eta_Y\right)^2}{\sigma_Y^2} \right] \right\}$$

- A very important model in statistical signal processing, communications, and control.

Properties of JGRV

- Marginal distributions are Gaussian

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x-\eta_X)^2}{2\sigma_X^2}\right\}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left\{-\frac{(y-\eta_Y)^2}{2\sigma_Y^2}\right\}$$

(2.106)

- Conditional pdf are Gaussian (Show this!)

$$f_Y(y|x) = \frac{1}{\sqrt{2\pi}\sigma_{Y|x}} \exp\left\{-\frac{(y-\eta_{Y|x})^2}{2\sigma_{Y|x}^2}\right\}, \qquad \eta_{Y|x} = \eta_Y + \frac{\rho_{XY}\sigma_Y}{\sigma_X}(x-\eta_X)$$

$$\sigma_{Y|x}^2 = \sigma_Y^2(1-\rho_{XY}^2)$$

$$f_X(x|y) = \frac{1}{\sqrt{2\pi}\sigma_{X|y}} \exp\left\{-\frac{(x-\eta_{X|y})^2}{2\sigma_{X|y}^2}\right\}, \qquad \eta_{X|y} = \eta_X + \frac{\rho_{XY}\sigma_X}{\sigma_Y}(y-\eta_Y)$$

$$\sigma_{X|y}^2 = \sigma_X^2(1-\rho_{XY}^2)$$

(2.107)

Multidimensional Gaussian pdf

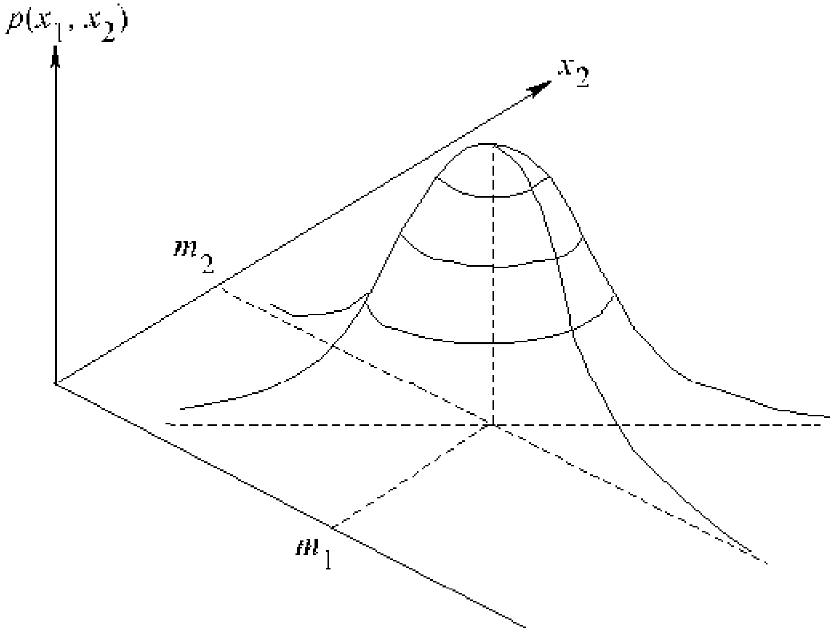
$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\left|\mathbf{K}_{\mathbf{X}}\right|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\eta}_{\mathbf{X}})^{T}\mathbf{K}_{\mathbf{X}}^{-1}(\mathbf{x}-\boldsymbol{\eta}_{\mathbf{X}})\right]$$

$$f_{\mathbf{X}}(\mathbf{x}) \sim N(\boldsymbol{\eta}_{\mathbf{X}}, \mathbf{K}_{\mathbf{X}})$$

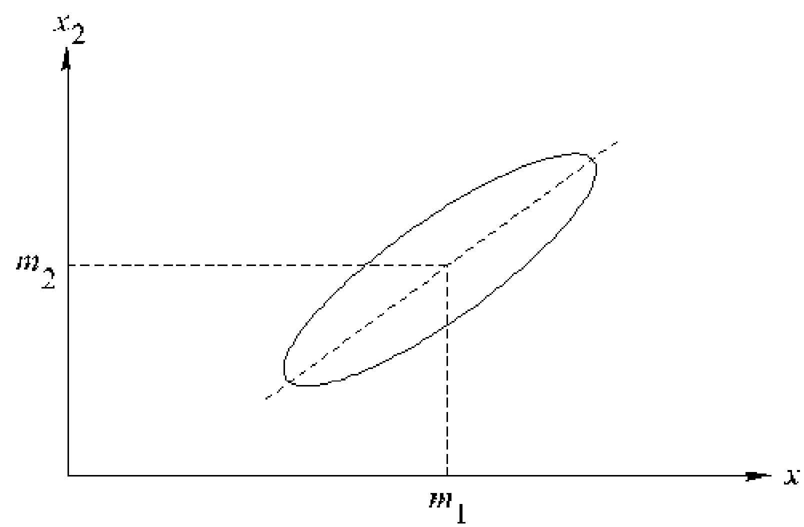
$$\boldsymbol{\eta}_{\mathbf{X}} = E[\mathbf{X}] = \begin{bmatrix} E[X_1] & E[X_2] & \cdots & E[X_n] \end{bmatrix}^{T}$$

$$\mathbf{K}_{\mathbf{X}} = E\left[(\mathbf{X}-\boldsymbol{\eta}_{\mathbf{X}})(\mathbf{X}-\boldsymbol{\eta}_{\mathbf{X}})^{T}\right]$$

Multidimensional Gaussian Density



(a)



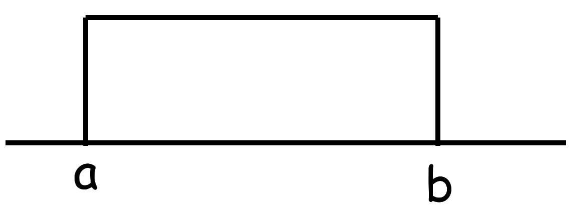
(b)

Properties of the Multivariate Gaussian pdf

- All marginal functions are Gaussian
- Joint pdf of any subset of components of X are JGRV
- Conditional pdfs of a subset of variables on a different subset are Gaussian
- Any linear transformation of X, Z=AX+b, the RV Z has a multidimensional Gaussian pdf

Uniform pdf

- A R.V. X that is uniformly distributed between x_1 and x_2 has density function:

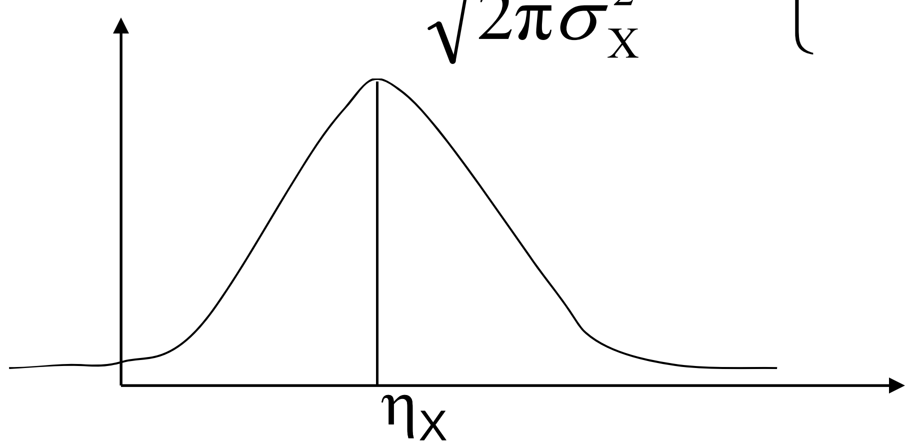


$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

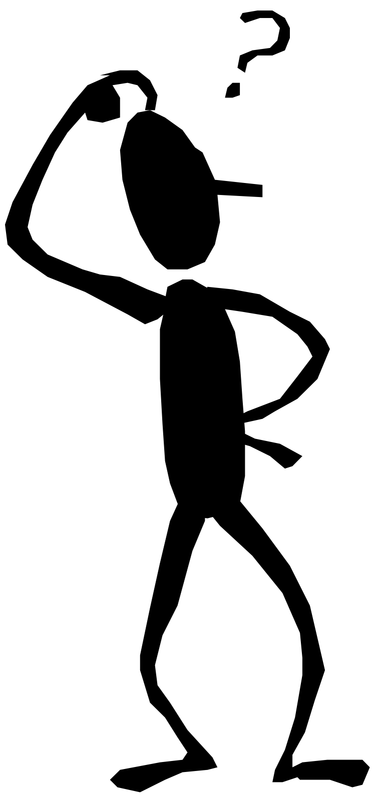
Gaussian pdf

- A R.V. X that is normally distributed has density function:

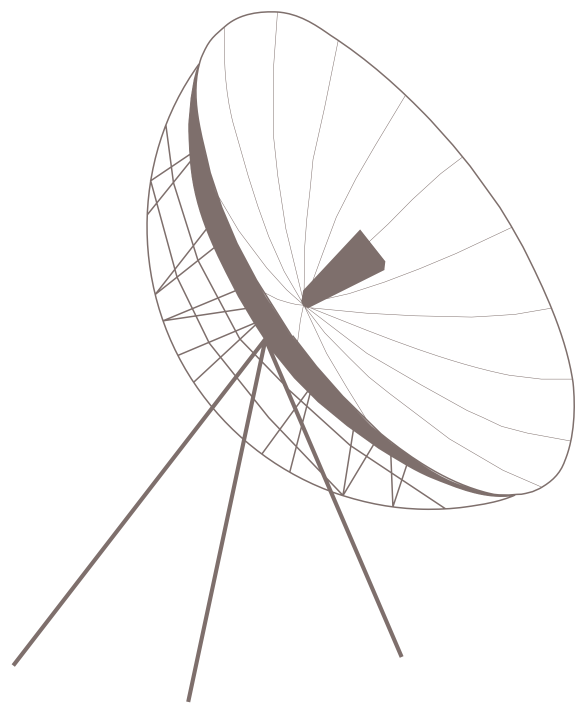
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left\{ -\frac{(x-\eta_X)^2}{2\sigma_X^2} \right\}$$



Detection of a Signal

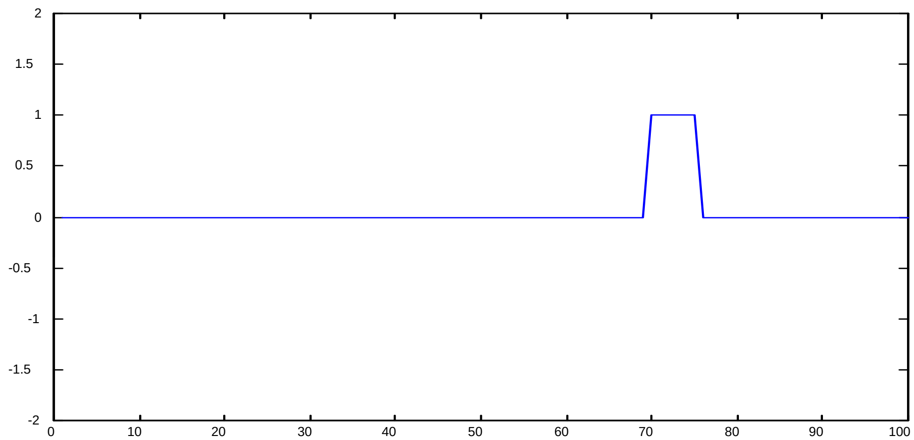


Is there some signal around?

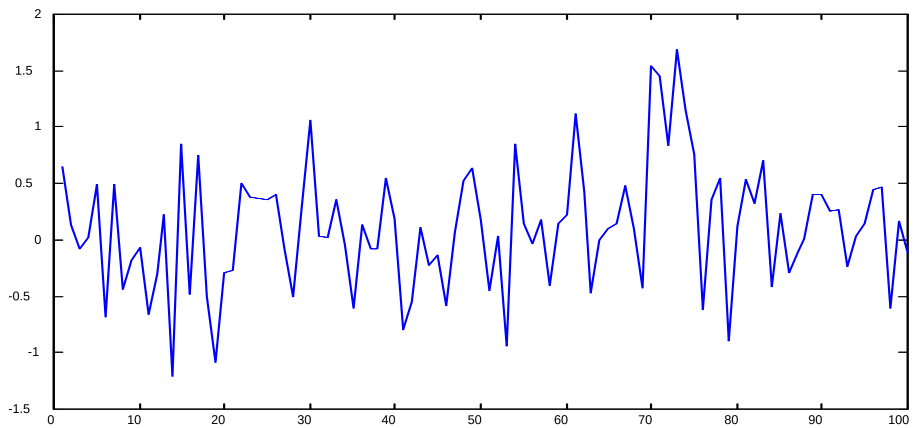


Detection Concept

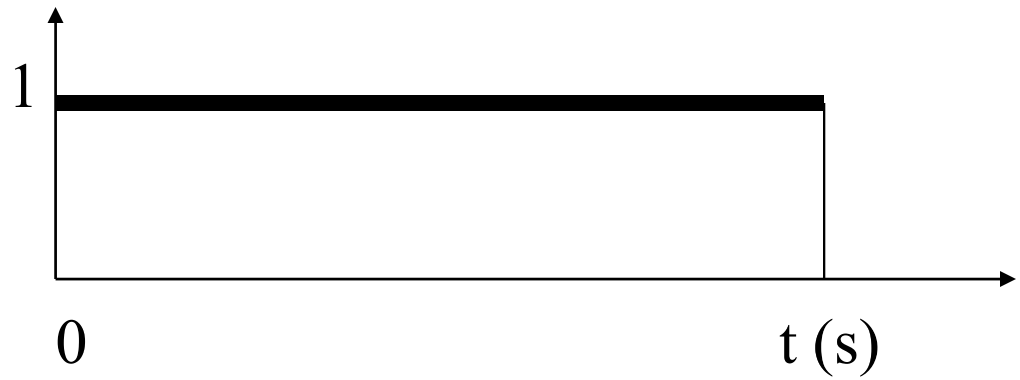
Signal



Signal
+
Noise



Bit 1 or 0



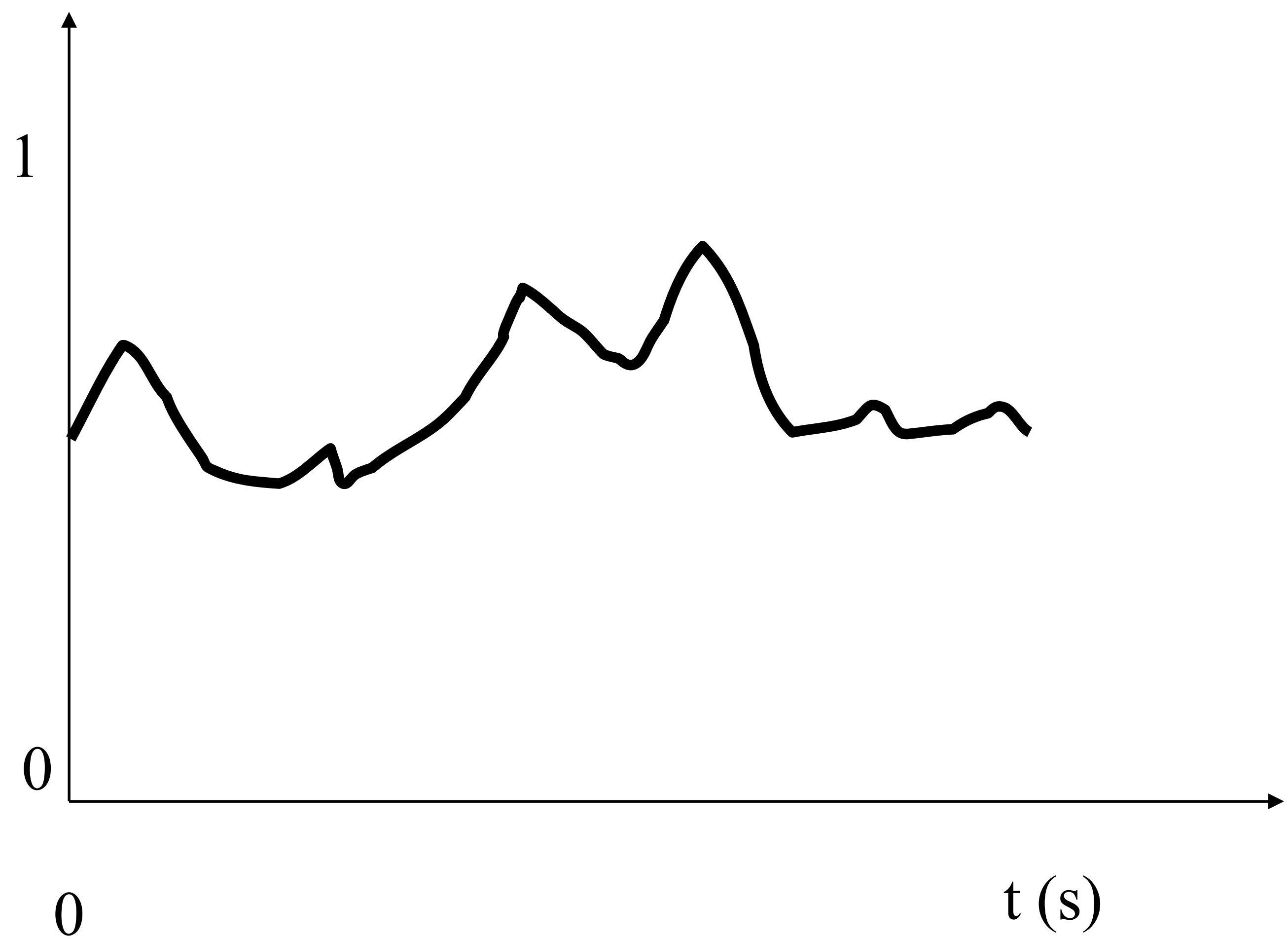
Two Signals + Noise

$$X = S_i + n$$

$$S_0 = \text{Bit } 0$$

$$S_1 = \text{Bit } 1$$

What is this?



Need a Threshold & Sampling