University of Central Florida
School of Electrical Engineering and Computer Science
EGN-3420 - Engineering Analysis.
Fall 2009 - dcm

Probability and Statistics Concepts

Random Variable: a rule that assigns a numerical value to each possible outcome of an experiment. All possible outcomes of the experiment constitute a sample space.

A random variable X on a sample space S is a function $X : S \mapsto \mathbb{R}$ which assigns a real number $X(s)$ to every sample point $s \in S$. This real number is called the probability of that outcome.

A discrete random variable maps events to values of a countable set e.g., the set of integers); each value in the range has a probability greater than or equal to zero.

Example 1. *the experiment is a coin toss; the outcome is either* 0 *(head) or* 1 *(tail).* If the coin is fair then $p_0 = p_1 = 0.5$; this means that in a large number of coin tosses we are likely to observe heads in about half of the cases and tails in the other half of the cases. Another example: when you throw throw a dice the outcome could be $1, 2, 3, 4, 5,$ or 6; for a fair dice $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 1/6$.

A continuous random variable maps events to values of an uncountable set (e.g., the real numbers).

Example 2. *the experiment is to measure the speed of cars passing through an intersection:* the speed could be any value between 15 and 80 miles/hour. the probability of observing cars with a speed of 19.1 miles/hour could be zero but the probability of observing cars with a speed from 15 to 19.1 miles/hour could be $P_{19.1} = 0.3$ which means that 30% of the cars we observed have a speed in the range we considered.

A discrete random variable X has an associated probability density function, (also called probability mass function) $p_X(x)$ defined as:

$$p_X(x) = \text{Prob}(X = x)$$

and a probability distribution function also called cumulative distribution function, $P_X(x)$ defined as:

$$P_X(t) = \text{Prob}(X \leq t) = \sum_{x \leq t} p_X(x)$$

Example 3. *You have a binary random variable X (the outcome is either* 0 *or* 1*) and:*

$$p_0 = \text{Prob}(X = 0) = q \qquad \text{and} \qquad p_1 = \text{Prob}(X = 1) = p, \quad \text{with} \quad p + q = 1.$$

Bernouli trials: call the outcome of 1 a "success" and ask the question what is the probability Y_n that in n Bernoulli trials we have k successes:

$$p_k = \text{Prob}(Y_n = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

The binomial cumulative distribution function is:

$$B(t:n,p) = \sum_{k=0}^{t} \binom{n}{k} p^k (1-p)^{n-k}$$

Example 4. *You have again Bernoulli trials and ask the question how many trails you need before the first "success".* If the first success occurs at the i-th trial then

$$p_Z(i) = q^{i-1}p$$

This is called a geometric distribution. It is easy to prove that:

$$\sum_{i=0}^{\infty} q^{i-1}p = \frac{p}{1-q} = 1.$$

A continuous random variable X has an associated probability density function, (also called probability mass function) $p_X(x)$ defined as:

$$f_X(x) = \text{Prob}(X = x)$$

and a probability distribution function also called cumulative distribution function, $F_X(x)$ defined as:

$$F_X(t) = \text{Prob}(X \leq t) = \int_{-\infty}^{t} f_X(x)dx$$

The expectation of random variable X: $E[X]$ is defined by

$$E[X] = \begin{cases} \sum_i x_i p_X(x_i) & \text{if X is discrete} \\ \int_{-\infty}^{+\infty} x f(x)dx & \text{if X is continuous} \end{cases}$$

The variance Var[X] and standard deviation, σ of random variable X are defined by:

$$\text{Var}[X] = \sigma^2 = \begin{cases} \sum_i (x_i - E[X])^2 p_X(x_i) & \text{if X is discrete} \\ \int_{-\infty}^{+\infty} (x - E[X])^2 f(x)dx & \text{if X is continuous} \end{cases}$$

The moment of order k of random variable X is defined as:

$$E[X^k] = \begin{cases} \sum_i x_i^k p_X(x_i) & \text{if X is discrete} \\ \int_{-\infty}^{+\infty} x^k f(x)dx & \text{if X is continuous} \end{cases}$$

The centered moment of order k of random variable X is defined as the k-th moment of the random variable $x - E[X]$:

$$\mu_k = E\left[(X - E[X])^k\right] = \begin{cases} \sum_i (x_i - E[X])^k p_X(x_i) & \text{if X is discrete} \\ \int_{-\infty}^{+\infty} (x - E[X])^k f(x)dx & \text{if X is continuous} \end{cases}$$

Examples of common distributions

1. Uniform distribution in the interval [a,b]: see Figure 1.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if} \quad a \leq x \leq b \\ 0 & \text{if} \quad x < a \text{ or } x > b \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \frac{x-a}{b-a} & \text{if} \quad a \leq x \leq b \\ 1 & \text{if} \quad x > b \end{cases}$$

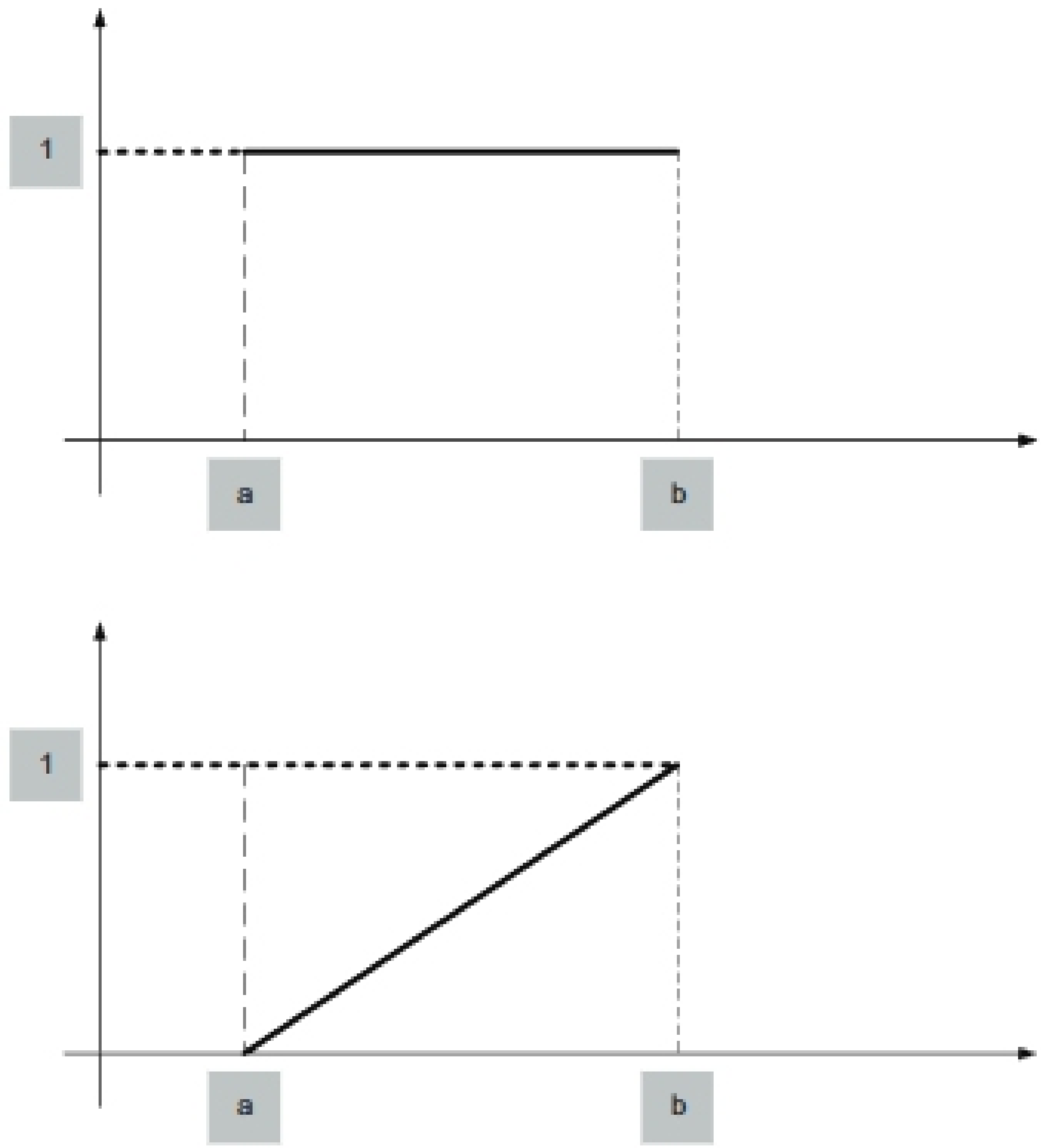

Figure 1: Probability density function (PDF) and cumulative distribution function of a uniform distribution

2. Standard normal distribution:

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

3. Normal distribution with mean μ and standard deviation σ:

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

The probability density function (PDF) and the cumulative distribution function (CDF) of a normal distribution are displayed in Figures 2 and 3.