Computational Foundations of Cognitive Science

Lecture 21: Special Distributions and Densities

Frank Keller

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

February 23, 2010

1. Mean and Variance
2. Special Probability Distributions
   - Uniform Distribution
   - Binominal Distribution
   - Mid-lecture Problem
3. Special Probability Densities
   - Uniform Distribution
   - Exponential Distribution
   - Normal Distribution

Reading: Freund, Chs. 5.1, 5.2, 5.4, 6.1–6.3, 6.5.

Mean

The *mean* of a random variable X is its average value, normally denoted by μ.

Definition: Mean

If X is a discrete random variable and $f(x)$ is the value of its probability distribution at x, then the mean of X is:

$$\mu = \sum_x x \cdot f(x)$$

Mean and variance are *expectations* of a random variable. This will be discussed in more detail in subsequent lectures.

Variance

Definition: Variance

If X is a discrete random variable and $f(x)$ is the value of its probability distribution at x, and μ is its mean then:

$$\sigma^2 = \text{var}(X) = \sum_x (x - \mu)^2 f(x)$$

is the variance of X.

Intuitively, var(X) reflects the *spread* or *dispersion* of a distribution, i.e., how much it diverges from the mean.

σ is called the *standard deviation* of X.

Variance

Example

Let X be a discrete random variable with the distribution:

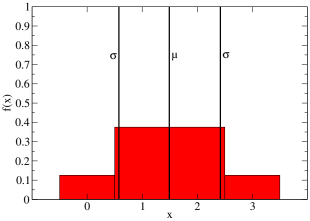
$$f(x) = \begin{cases} \frac{1}{8} & \text{for } x = 0 \\ \frac{3}{8} & \text{for } x = 1 \\ \frac{3}{8} & \text{for } x = 2 \\ \frac{1}{8} & \text{for } x = 3 \end{cases}$$

The mean and variance of X are:

$$\begin{aligned} \mu &= \sum_x x \cdot f(x) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} \\ \text{var}(X) &= \sum_x (x-\mu)^2 f(x) \\ &= (0-\frac{3}{2})^2 \cdot \frac{1}{8} + (1-\frac{3}{2})^2 \cdot \frac{3}{8} + (2-\frac{3}{2})^2 \cdot \frac{3}{8} + (3-\frac{3}{2})^2 \cdot \frac{1}{8} \\ &= 0.86 \end{aligned}$$

Variance

Histogram with mean and standard deviation for the previous example:



Uniform Distribution

Definition: Uniform Distribution

A random variable X has a discrete uniform distribution iff its probability distribution is given by:

$$f(x) = \frac{1}{k} \text{ for } x = x_1, x_2, \ldots, x_k$$

where $x_i \neq x_j$ when $i \neq j$.

The mean and variance of the uniform distribution are:

$$\mu = \sum_{i=1}^{k} x_i \cdot \frac{1}{k} \qquad \sigma^2 = \sum_{i=1}^{k} (x_i - \mu)^2 \frac{1}{k}$$

Binominal Distribution

Often we are interested in experiments with *repeated trials:*

- assume there is a fixed number of trials;
- each of the trial can have two outcomes (e.g., success and failure, head and tail);
- the probability of success and failure is the same for each trial: θ and $1-\theta$;
- the trials are all independent.

Then the probability of getting x successes in n trials in a given order is $\theta^x(1-\theta)^{n-x}$. And there are $\binom{n}{x}$ different orders, so the overall probability is $\binom{n}{x}\theta^x(1-\theta)^{n-x}$.

Binomial Distribution

Definition: Binomial Distribution

A random variable X has a binominal distribution iff its probability distribution is given by:

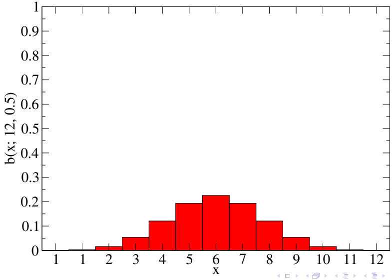
$$b(x; n, \theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \text{ for } x = 0, 1, 2, \ldots, n$$

Example

The probability of getting five heads and seven tail in 12 flips of a balanced coin is:

$$b(5; 12, \frac{1}{2}) = \binom{12}{5}(\frac{1}{2})^5(1-\frac{1}{2})^{12-5}$$

Binomial Distribution



Binomial Distribution

If we invert successes and failures (or heads and tails), then the probability stays the same. Therefore:

Theorem: Binomial Distribution

$$b(x; n, \theta) = b(n-x; n, 1-\theta)$$

Two other important properties of the binomial distribution are:

Theorem: Binomial Distribution

The mean and the variance of the binomial distribution are:

$$\mu = n\theta \quad \text{and} \quad \sigma^2 = n\theta(1-\theta)$$

Mid-lecture Problem

A study found that 80% of all people over 60 years of age wear glasses. If a random sample of 6 people over 60 years of age is selected, what's the probability that:

1. exactly 4 people will wear glasses;
2. at most 2 people will wear glasses;
3. at least 2 people will wear glasses.

Uniform Distribution

Definition: Uniform Distribution

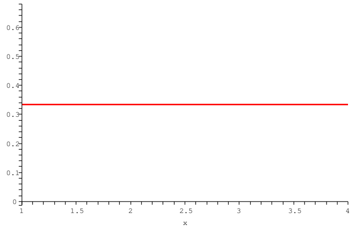
A random variable X has a continuous uniform distribution iff its probability density is given by:

$$u(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha} & \text{for } \alpha < x < \beta \\ 0 & \text{elsewhere} \end{cases}$$

The mean and variance of the uniform distribution are:

$$\mu = \frac{\alpha + \beta}{2} \qquad \sigma^2 = \frac{1}{12}(\alpha - \beta)^2$$

Uniform Distribution



Uniform distribution for $\alpha = 1$ and $\beta = 4$.

Exponential Distribution

Definition: Exponential Distribution

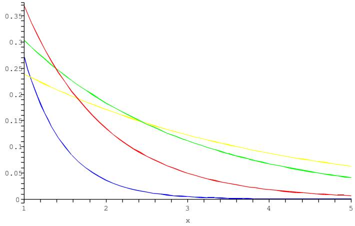
A random variable X has an exponential distribution iff its probability density is given by:

$$g(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

The mean and variance of the exponential distribution are:

$$\mu = \theta \qquad \sigma^2 = \theta^2$$

Exponential Distribution



Exponential distribution for $\theta = \{0.5, 1, 2, 3\}$.

Normal Distribution

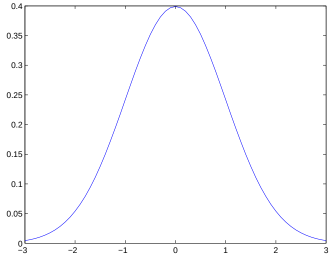
Definition: Normal Distribution

A random variable X has a normal distribution iff its probability density is given by:

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \text{ for } -\infty < x < \infty$$

- Normally distributed random variables are ubiquitous in probability theory;
- many measurements of physical, biological, or cognitive processes yield normally distributed data;
- such data can be modeled using a normal distributions (sometimes using mixtures of several normal distributions).

Standard Normal Distribution



Normal Distribution

Definition: Standard Normal Distribution

The normal distribution with $\mu = 0$ and $\sigma = 1$ is referred to as the standard normal distribution:

$$n(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Theorem: Standard Normal Distribution

If a random variable X has a normal distribution, then:

$$\begin{aligned} P(|x-\mu| < \sigma) &= 0.6826 \\ P(|x-\mu| < 2\sigma) &= 0.9544 \end{aligned}$$

This follows from Chebyshev's Theorem (later in this course).

Normal Distribution

Theorem: Z-Scores

If a random variable X has a normal distribution with the mean μ and the standard deviation σ then:

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

has the standard normal distribution.

This conversion is often used to make results obtained by different experiments comparable: convert the distributions to Z-scores.

Summary

- The uniform distribution assigns each value the same probability;
- The binomial distributions models an experiment with a fixed number of independent binary trials, each with the same probability;
- The normal distribution models the data generated by measurements of physical, biological, or cognitive processes;
- Z-scores can be used to convert a normal distribution into the standard normal distribution.