Introduction

Today[1]:

- Parametric estimation of a p.d.f.
- Definitions: Events, joint probability etc.
- Bayes theorem.

Introduction Statistics , by Prem S. Mann, fifth edition, Wiley 2004.

Probability and Random Processes with Applications to Signal Processing , by H. Stark & J.W. Woods, third edition, Prentice Hall 2002.

[1] January 16, 2008

Parametric estimation of a p.d.f. I

Lets assume that we have a set of samples or observations $\{x_i\}_{i=1,\cdots,n}$ of the random variable X. One approach to probability density function estimation is parametric i.e. we assume that the underlying p.d.f. $f(x)$ of the observations $\{x_i\}_{i=1,\cdots,n}$, belongs to a known parametric family of distributions, for instance the normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

the estimate of the p.d.f., $\hat{f}$ is then expressed as:

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp\left(-\frac{(x-\overline{x})^2}{2\hat{\sigma}^2}\right)$$

where $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$ is an estimate of μ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})^2$ is the plug-in estimate of the standard deviation σ.

Parametric estimation of a p.d.f. II

Now consider the observations $\{\mathbf{x}_i\}_{i=1,\cdots,n}$ of a random vector $\mathbf{x} \in \mathbb{R}^d$. In that case the multivariate normal p.d.f. is defined as:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\mathrm{C}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathrm{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

where C is the matrix of covariance, $|\mathrm{C}| = \det(\mathrm{C})$ and $\boldsymbol{\mu}$ the mean.
The parametric estimate $\hat{f}(\mathbf{x})$ is then computed using estimates $\overline{\mathbf{x}}$ and $\hat{\mathrm{C}}$.

Dependence and Independence I

An experiment is a process that, when performed, results in one and only one of many observations. These observations are called the outcomes of the experiment. The collection of all outcomes for an experiment is called a sample space Ω. An event is a collection of one or more outcomes of an experiment (i.e. a subset of Ω).

Definition (Marginal Probability)

Marginal probability $P(A)$ is the probability of a single event A without consideration of any other event.

Dependence and Independence II

Definition (Conditional probability)

Conditional probability $P(A|B)$ (read as *the probability of A given that B has already occurred*) is the probability that one event A will occur given that another event B has already occurred.

Definition (Mutually exclusive events)

Mutually Exclusive Events A and B are events that cannot occur together.

Dependence and Independence III

Definition (Independent events)

Two events A and B are said to be independent if the occurrence of one does not affect the probability of the occurrence of the other:

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B)$$

otherwise the A and B are said dependent.

Definition (Complementary events)

The complement of event A, denoted $\overline{A}$ is the event that includes all the outcomes for an experiment that are not in A.

$$P(A) + P(\overline{A}) = 1$$

Dependence and Independence IV

Definition (Joint probability)

The joint probability of two events $P(A,B)$ is the probability that the events A and B occur together, and we have:

$$P(A,B) = P(B|A)P(A) = P(A|B)P(B) \quad \text{multiplication rule to find joint probability}$$

Theorem (Bayes's theorem)

Bayes' theorem relates the conditional and marginal probabilities of stochastic events A and B:

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

Exercise I

Consider the following events for one roll of a die

$A = \{2,4,6\}$ an even number is observed
$B = \{1,3,5\}$ an odd number is observed
$C = \{1,2,3,4\}$ a number less than 5 is observed

1. Are events A and B mutually exclusive? Are events A and C mutually exclusive?
2. Compute the sample space Ω.
3. Compute $P(A)$, $P(B)$ and $P(C)$.
4. Compute $P(A,B)$ and $P(A,C)$.
5. Compute $P(A|B)$, $P(A|C)$ and $P(C|A)$.

Exercise II

Consider the random variable x in $\Omega = \mathbb{R}$ and the event $x \in A = [0;1]$ with the p.d.f.

$$f(x) = \begin{cases} 1 & |x| < 1/2 \\ 0 & \text{otherwise} \end{cases}$$

1. Compute $P(\Omega)$.
2. Compute $P(A)$.

Bayes' theorem for probability densities

Theorem (Bayes' theorem for probability densities)

Consider two random variables X and Y and their p.d.fs. The Bayes rules is expressed as:

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

where

- *$f_{X|Y}(x|y)$ is the posterior distribution of X given Y.*
- *$f_{Y|X}(y|x)$ is (as a function of x) the likelihood function.*
- *$f_X(x)$ and $f_Y(y)$ are the marginals.*

Exercise

Let

$$f_{XY}(x,y) = \begin{cases} A(x+y) & 0 < x \leq 1, \quad 0 < y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

1. Find A.
2. What are the marginals?