3F1 Signals and Systems

(14) Review of probability and a glimpse of random signals

Timothy O'Leary

Michaelmas Term 2016

Review of Probability Concepts

The **probability density function** (pdf) of a continuous random variable X taking real values, x, is a non-negative function f_X such that:

$$Pr(X \leq a) = \int_{-\infty}^{a} f_X(x)dx := F_X(a)$$

We call F_X the cumulative density function (cdf).

Properties

1. $f_X(x) = \frac{d}{dx}F_X(x)$
2. f_X is normalised: $\int_{-\infty}^{\infty} f_X(x)dx = 1$
3. $Pr(b \leq X \leq a) = \int_b^a f_X(x)dx = F_X(a) - F_X(b)$
4. F_X is non-decreasing
5. $0 \leq F_X(x) \leq 1$
6. $Pr(X > a) = 1 - F_X(a)$

For discrete random variable, X, taking values in a set, $\{x_1, x_2, ...\}$ we can define the **probability mass function** (pmf), p_X:

$$p_X(x_i) = Pr(X = x_i)$$

Again, p_X is normalized: $\sum_i p_X(x_i) = 1$.

If A and B are two events, then the **joint probability** of A and B occurring together is denoted $Pr(A, B)$.

If $Pr(A, B) = Pr(A)Pr(B)$ we say that A and B are **independent**.

The **conditional probability** of A, given that B occurs, is denoted

$$Pr(A|B) = \frac{Pr(A, B)}{Pr(B)}$$

From this we can obtain **Bayes' Rule**:

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

Joint distributions

Given two continuous random variables, X and Y, the **joint cdf** is:

$$F_{X,Y}(x, y) = Pr\{(X \leq x) \text{ AND } (Y \leq y)\}$$

and the **conditional** cdf is:

$$F_{X|Y}(x|y) = Pr\{(X \leq x)|(Y \leq y)\} = \frac{F_{X,Y}(x, y)}{F_Y(y)}$$

Thus, Bayes' rule for cdfs is: $F_{X|Y}(x|y) = \frac{F_{Y|X}(y|x)F_X(x)}{F_Y(y)}$

The **joint pdf** is defined by $f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$. From this, it can be shown that:

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

The **law of total probability** (for pdfs) is:

$$
\begin{aligned}
\int f_{Y|X}(y|x) f_X(x) dx &= \int f_{Y,X}(y,x) dx \\
&= f_Y(y) \int f_{X|Y}(x|y) dx \\
&= f_Y(y)
\end{aligned}
$$

In this context, we call $f_Y(y)$ the **marginal** distribution for Y. The process of obtaining $f_Y(y)$ from the joint distribution, $f_{Y,X}(y,x)$, is called **marginalization**.

Functions of random variables

If a random variable, Y, is a monotonic function of another random variable, X, such that:

$$Y = g(X) \quad \text{and} \quad X = g^{-1}(Y)$$

Then we can relate the pdfs and cdfs of X and Y as follows. Noting that $y = g(x)$, the cdf of Y is:

$$F_Y(y) = Pr(Y \leq y) = Pr(g(X) \leq g(x)) = Pr(X \leq x) = F_X(x)$$

What about the pdf of Y?

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(x) = \left(\frac{d}{dx} F_X(x)\right) \frac{dx}{dy} = f_X(x) \frac{dx}{dy}$$

Thus, $f_Y(y) = \frac{f_X(x)}{g'(x)}$.

Expectation and Moments

The **Expectation** of a random variable, X, is defined as:

$$E[X] = \int x f_X(x) dx$$

For discrete X we have:

$$E[X] = \sum_i x_i p_X(x_i)$$

Expectation is a **linear operator**. For random variables, X and Y:

$$E[aX + bY] = aE[X] + bE[Y]$$

Often, we want to compute the expectation of a function of a random variable, e.g. $Y = g(X)$. In this case, using the relations of the previous section, we have $f_Y(y)dy = f_X(x)dx$, which gives:

$$E[g(X)] = E[Y] = \int y f_Y(y) dy = \int g(x) f_X(x) dx$$

Important examples of expectations are **moments of a pdf**. Putting $g(x) = x^n$ we define:

$$\text{nth order moment: } E[X^n] = \int x^n f_X(x) dx$$

case $n = 1$: $E(X) =$ Mean value

case $n = 2$: $E(X) =$ Mean-squared value (aka 'power' or 'energy')

Central moments are obtained by subtracting the mean value from the distribution:

$$\text{nth order central moment: } E[(X - \bar{X})^n] = \int (x - \bar{X})^n f_X(x) dx$$

For example, **variance**, σ^2, is the second-order central moment:

$$\begin{aligned}
\sigma^2 = E[(X - \bar{X})^2] &= \int (x - \bar{X})^2 f_X(x) dx \\
&= \int x^2 f_X(x) dx - 2\bar{X} \int x f_X(x) dx + \bar{X}^2 \int f_X(x) dx \\
&= E[X^2] - 2\bar{X}^2 + \bar{X}^2 = E[X^2] - \bar{X}^2.
\end{aligned}$$

Sums of random variables

Consider a random variable

$$Y = X_1 + X_2$$

where X_1, X_2 are two independent random variables with pdfs f_1 and f_2. We can write the joint pdf for y and x_1 using the conditional probability formula:

$$f(y, x_1) = f(y|x_1)f_1(x_1)$$

Now since $X_2 = Y - X1$, the event $\{Y = y|X_1 = x_1\}$ is equivalent to $\{X_2 = y - x_1\}$. Hence

$$f(y|x_1) = f_2(y - x_1)$$

Now $f(y)$ can be obtained by marginalising, that is:

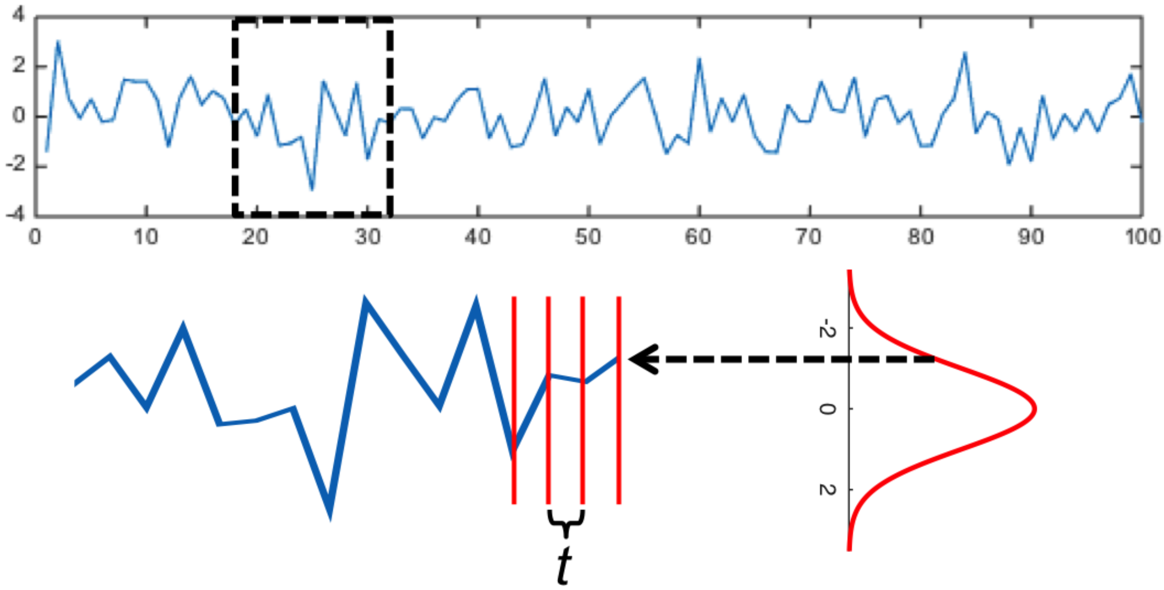
$$\begin{aligned} f(y) &= \int f(y|x_1)f_1(x_1)dx_1 \\ &= \int f_2(y - x_1)f_1(x_1)dx_1 \\ &= f_2 \star f_1 \end{aligned}$$

This argument can be extended to sums of three or more random variables. Since convolution is commutative, for a sum of n random variables we get a pdf:

$$f(y) = f_n \star (f_{n-1} \star ... \star f_1) = f_n \star f_{n-1} \star ... \star f_1$$

Random Signals (informal introduction)

Random signals are random variables that evolve in time (or in some other variable, but we will only consider time). In this course we will concentrate on **continuous time random signals**. Heuristically, we can think of a random signal as being generated by drawing a sample from a random variable at each point in time.



However, for this to make sense in continuous time, **we need to describe how quickly the signal can vary in a given time interval**. We will develop tools to handle this next.