CHAPTER II

CALCULATION PROCEDURES

Let $\{X_n\}$, $\{Y_n\}$ be any two finite records,

$$\{X_n\} = \{X_n / X_n = X(t_o + n\Delta t),\ n = 1,2,\ldots,N, \text{ and}$$

t_o is arbitrary time not included in the record

Δt is the sampling time interval$\}$

$\{Y_n\}$ is defined similar to $\{X_n\}$

The transformed record $\{x_n\}$ is defined by

$$\{x_n\} = \{X_n - \overline{X} / \overline{X} = \frac{1}{N}\sum_{n=1}^{N} X_n\}$$

Calculation of the mean square value

The sample mean square value is given by $\overline{x_n^2}$

$$\overline{x_n^2} = \frac{1}{N}\sum_{n=1}^{N} x_n^2$$

The standard deviation s, can be calculated from the mean square value by the equation

$$s = \left\{\frac{N}{N-1}\overline{x^2}\right\}^{\frac{1}{2}}$$

Calculation of probability density and distribution function

Let $[a, b]$ be the interval for the value of $\{X_n\}$ which is interested

k be the class intervals in which $[a, b]$ is divided

c be the intervals

then $c = \dfrac{b-a}{k}$

Define $d_i = a+(i-1)c\,,\ i = 1,2,\ldots,k+1$

N_1 be the number of $X_n \ni X_n \leq d_1$

N_i be the number of $X_n \ni d_{i-1} < X_n \leq d_i\,,\ i = 2,3,\ldots,k+1$

N_{k+2} be the number of $X_n \ni X_n > d_{k+1}$

Note $d_1 = a$

$$N = \sum_{i=1}^{k+2} N_i$$

The value of N_i , $i = 1,2,\ldots,k+1$, can be found by examing each X_n , $n = 1,2,\ldots,N$ as follows,

1. If $X_n \leq d_1$, add the integer one to N_1

2. If $X_n > d_{k+1}$, add the integer one to N_{k+2}

3. If $d_1 < X_n \leq d_{k+1}$, compute

$$I = \frac{X_n - d_1}{c}$$

then add the integer one to N_i where i is the smallest integer that is greater than or equal to $I+1$

The sequence f_i of sample probability density function at the midpoint of the ith class interval in $[a, b]$ defined by

$$f_i = \frac{N_i}{cN} \qquad i = 2,\ldots,k+1$$

The sequence F_j of sample probability distribution of the jth class interval in $[a , b]$ defined by

$$F_j = \sum_{i=1}^{j} \frac{N_i}{N}$$

Auto-correlation and power spectra calculations

The auto-correlation function for random data describes the general dependence of values of the data at one time on the values at another time; eg., the discrete sample auto-correlation at lag r defined by $\hat{R}_X(r)$[1] described the correlation of values of random data taken at times differing from each other by r sampling time intervals.

$$\hat{R}_X(r) = \frac{1}{N-r}\sum_{n=1}^{N-r} X_n X_{n+r} \qquad r \geqslant 0$$

For a finite sample, the auto-correlation function of lag 0 up to lag m, m the maximum number of correlation lag, is calculated. The maximum lag depends on the bandwidth of the power spectrum and the time interval of sampling.

The power spectrum of stationary data is the Fourier transform of the auto-correlation. Since the auto-correlation is a real and even function: $\hat{R}_X(-r)$[2] $= \hat{R}_X(r)$[2], only the cosine Fourier transform of auto-correlation will give the power spectrum.

Let $\hat{P}_X(f)$ denote power spectrum estimate of series data $\{X_n\}$, then

$$\hat{P}_X(f) = 2\Delta t\left[\hat{R}_0 + 2\sum_{r=1}^{m-1}\hat{R}_r \operatorname{Cos}\left(\frac{\pi r f}{f_c}\right) + \hat{R}_m \operatorname{Cos}\left(\frac{\pi m f}{f_c}\right)\right]$$

f_c, the cutoff frequency, is defined so that $\frac{1}{f_c}$ is the smallest period in record.

---

1 see References no. 1

2 see References no. 2

The values of the function $\hat{P}_X(f)$ can be calculated only at the m+1 special discrete frequency value

$$f = \frac{kf_c}{m} \qquad k = 0,1,\ldots,m$$

At these discrete frequency points,

$$\hat{P}_k = \hat{P}_X\left(\frac{kf_c}{m}\right)$$

$$= 2\Delta t\left[\hat{R}_0 + 2\sum_{r=1}^{m-1}\hat{R}_r \cos\left(\frac{\pi r k}{m}\right) + (-1)^k\hat{R}_m\right]$$

k is the index called the harmonic number

$\hat{P}_k$ is the estimate of the power spectral density function at harmonic k, corresponding to the frequency $f = \frac{kf_c}{m}$.

Cross-correlation and cross-power spectra calculation

The cross-correlation function for two sets of random data describes the general dependence of the value of one set of data on the other.

The estimates for the sample cross-correlation function at lag number r = 0,1,2,...,m are defined by $R_{XY}(r)$ and $R_{YX}(r)$ [3].

$$\hat{R}_{XY}(r) = \frac{1}{N-r}\sum_{n=1}^{N-r} X_n Y_{n-r} \qquad r \geqslant 0$$

$$\hat{R}_{YX}(r) = \frac{1}{N-r}\sum_{n=1}^{N-r} Y_n X_{n-r} \qquad r \geqslant 0$$

The cross-power spectrum is defined by

$$\hat{P}_{XY}(f) = \hat{C}_{XY}(f) + i\,\hat{Q}_{XY}(f)$$

---

3 see References 2

The real part of cross-spectrum is called co-spectrum /power spec.

The imaginary part of cross-spectrum is called quadrature /power spectrum

Let $\hat{A}_r = \hat{A}_{XY}(r)$

$$= \tfrac{1}{2}[\hat{R}_{XY}(r) + \hat{R}_{YX}(r)]$$

$$\hat{B}_r = \hat{B}_{XY}(r)$$

$$= \tfrac{1}{2}\{\hat{R}_{XY}(r) - \hat{R}_{YX}(r)\}$$

then $\hat{C}_{XY}(f)$ and $\hat{Q}_{XY}(f)$ can be calculated from

$$\hat{C}_{XY}(f) = 2\Delta t\left[\hat{A}_0 + 2\sum_{r=1}^{m-1}\hat{A}_r \mathrm{Cos}\left(\frac{\pi r f}{f_c}\right) + \hat{A}_m \mathrm{Cos}\left(\frac{\pi m f}{f_c}\right)\right]$$

$$\hat{Q}_{XY}(f) = 2\Delta t\left[2\sum_{r=1}^{m-1}\hat{B}_r \mathrm{Sin}\left(\frac{\pi r f}{f_c}\right) + \hat{B}_m \mathrm{Sin}\left(\frac{\pi m f}{f_c}\right)\right]$$

As power spectrum, these function may be calculated only at the m+1 special discrete frequency for harmonic number k, where

$$f = \frac{k f_c}{m} \qquad k = 0,1,2,\ldots,m$$

At these discrete frequency points

$$\hat{C}_k = \hat{C}_{XY}\left(\frac{k f_c}{m}\right)$$

$$= 2\Delta t\left[\hat{A}_0 + 2\sum_{r=1}^{m-1}\hat{A}_r \mathrm{Cos}\left(\frac{\pi r k}{m}\right) + (-1)^k\hat{A}_m\right]$$

$$\hat{Q}_k = \hat{Q}_{XY}\left(\frac{k f_c}{m}\right) = 4\Delta t\sum_{r=1}^{m-1}\hat{B}_r \mathrm{Sin}\left(\frac{\pi r k}{m}\right)$$

Notes 1. The auto-correlation, power spectrum, cross-correlation and cross-power spectrum calculations will be later referred to spectral analysis.

2. The spectrum calculation from the above procedure is called raw spectrum. The smooth spectrum S_k^*[4] can be calculated from

$$S_1^* = \tfrac{1}{2}(S_1 + S_2)$$

$$S_m^* = \tfrac{1}{2}(S_{m-1} + S_m)$$

$$S_k^* = 0.25\,S_{k-1} + 0.5\,S_k + 0.25\,S_{k+1}, \quad k=2,\ldots,m-1$$

where m is the maximum lag number ,

S_k is the raw spectrum

3. In spectral analysis, it is assumed that the mean value of each sample record is zero. If the mean value of any sample record is not zero, the sample record should be transformed before it is analysed.

---

4 see References no. 3