XIV. Power Spectral Estimation
1. To use a digital computer for spectral analysis of the continuous EEG time series requires sampling at a finite sample interval.

2. If we know that the time series is band-limited to a certain frequency range, then we can avoid aliasing by sampling at a rate whose Nyquist frequency is above that range.
3. If the time series is not band-limited, or we are uncertain about its upper limit, or we are unable to sample fast enough, then low-pass filtering should be used to band-limit the time series prior to sampling.

4. In this way, a power spectrum for frequencies below the Nyquist frequency can be obtained that is free of aliasing errors, although information concerning activity at higher frequencies is necessarily lost.
5. Spectral analysis of the EEG involves the Fourier transform of a finite length time series. This means that the spectrum will consist of discrete components that depend on the sample length (are harmonically related to the fundamental frequency). Because the EEG time series will in general contain frequency components that are not harmonically related to the fundamental frequency, the spectrum will inevitably contain errors due to leakage.
We now consider the question of power spectral estimation. This question arises because, for reasons previously discussed, the EEG is an inherently noisy signal, and is modeled as a stochastic process.

For EEG analysis, therefore, the problem is one of estimating the power spectrum of a random process. This means that, in addition to the errors previously mentioned, there will be further errors, statistical in nature, in the computed spectrum, i.e., estimation errors.
In order to estimate the power spectrum of a stochastic process, we need to consider a collection of time series segments as forming a statistical sample of the stochastic process, i.e. a collection of realizations of the stochastic process.
The stochastic process is considered to have certain statistical parameters, and we will seek to compute estimators of those parameters. For example, a power spectral estimator will be defined for a statistical sample of time segments. The sample is considered to be a collection of realizations of the stochastic process, and the sample power spectral estimator to estimate the "true" power spectrum of the process.
In dealing with sample estimators, we will consider two types of estimation error.

The first type of estimation error is called **bias**. It is error caused by the mean of the estimator differing from the true mean of the stochastic process.

The second type of error is called **mean squared error**. It is caused by the variance of the estimator being large.
An estimator with small bias is **accurate**. An estimator with small mean squared error is **precise**.

A desirable property for an estimator is for both types of error to be small, or at least to get progressively smaller as the sample size increases. An estimator for which both types of error decrease with increasing sample size is called **consistent**.
First, remember that a stationary Gaussian interdependent stochastic process is adequately described by its mean, variance, and autocovariance function (acvf). The power spectrum is equivalent to the acvf.

In what follows, the sample acvf and the sample spectrum will be treated as estimators of the true acvf and true power spectrum of the stochastic process, and we will determine whether these two estimators are consistent.
In what follows:
1) We first define the true acvf and the sample acvf.
2) We next define the true power spectrum and the sample power spectrum.
3) We finally deal with the estimation errors that arise when we use the estimators to estimate the true acvf and true power spectrum.
The (true) autocovariance function (acvf) is defined as:

\[ \gamma_{xx}(\tau) = \text{Cov}[X(t), X(t+\tau)] \]

\[ = E\left[(X(t) - \mu)(X(t+\tau) - \mu)\right] \]
The Sample Autocovariance Function

To estimate the acvf of a stochastic process, we begin with a time series $X(t)$ that is regarded as a realization of the stochastic process.

A realization, $x(t)$, of $X(t)$ is one of many possible time series which might be observed. Since $x(t)$ is of finite length $T$, it is 0 when $t < 0$ and when $t > T$. 
Following Jenkins & Watts (p. 174), the sample acvf estimator is defined as:

\[
\hat{C}_{xx}(\tau) = \begin{cases} 
\frac{1}{T} \int_0^{\tau+T} (X(t) - \bar{X})(X(t + |\tau|) - \bar{X}) dt, & 0 \leq |\tau| \leq T \\
0, & |\tau| > T
\end{cases}
\]

Without any loss of generality, we assume that \(X(t)\) is a zero-mean process. Then the sample acvf is:

\[
\hat{C}_{xx}(\tau) = \begin{cases} 
\frac{1}{T} \int_0^{\tau+T} X(t)X(t + |\tau|) dt, & 0 \leq |\tau| \leq T \\
0, & |\tau| > T
\end{cases}
\]

The sample acvf is an estimator of the true acvf of the stochastic process.
The True Power Spectrum

The true power spectrum of stochastic process \( X(t) \) is the Fourier Transform of the true acvf. Thus:

\[
\Gamma_{xx}(\omega) = \int_{-\infty}^{\infty} \gamma_{xx}(\tau) e^{-j\omega \tau} d\tau
\]
The Sample Power Spectrum

The sample power spectrum is an estimator of the true power spectrum of the stochastic process $X(t)$. There are two ways to define the sample spectrum of a stochastic process.
1) The sample power spectrum may simply be defined as the Fourier Transform of the sample acvf:

$$\hat{C}_{xx}(n) = \int_{-T}^{T} C_{xx}(\tau) e^{-jn\omega_0\tau} d\tau$$

The sample spectrum obtained in this way is called the correlogram. This method was commonly used prior to 1965. It is also referred to as the Blackman-Tukey spectral estimator (Kay, 1988).
2) The sample spectrum may also be derived by first computing the Fourier transform of the time sample, and then multiplying by the complex conjugate in the frequency domain, as follows:

\[
\hat{C}_{xx}(n) = T |Z(n)|^2 = T Z(n) Z^*(n)
\]

The sample spectrum obtained in this way is called the periodogram.
Historical Note

In 1965, a Fast Fourier Transform (FFT) algorithm was published. It made the periodogram method for computing the sample power spectrum substantially faster than the correlogram method. It gained its speed by using the fact that, when the length is a power of 2, a substantial part of the sums and products in the classical procedure are calculated more than once. By not repeating these duplicate operations, this algorithm made the periodogram approach the more rapid one. The widespread use of the FFT is the reason that EEG sampling is often performed at powers of 2 sampling rates.
We now demonstrate that the two definitions of the sample spectrum are equivalent, following the derivation of Jenkins & Watts (p. 213-217).

Using the periodogram definition of the sample spectrum, and the definition of $Z(n)$:

$$\left|TZ(n)\right|^2 = T \cdot \frac{1}{T} \int_{-T/2}^{T/2} X(t) e^{-jn\omega_0 t} dt \cdot \frac{1}{T} \int_{-T/2}^{T/2} X(t') e^{jn\omega_0 t'} dt'$$

$$= \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} X(t) X(t') e^{-jn\omega_0 (t-t')} dt' dt$$
We can transform the region of integration by using the transformation: $u = t - t'$, and $v = t'$. This transformation is displayed graphically in Fig. 6.3 of Jenkins & Watts (p. 216).
1) Replace $t - t'$ by $u$, and replace $t$ by $v + u$. Then:

$$T \left| Z(n) \right|^2 = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} X(v) X(v + u) e^{-j\omega_0 u} dv \, d(v + u)$$

2) Since the inner integral integrates over $v$, it is not necessary to have $v$ in the outer integral as well. Therefore, we can change the variable of the outer integral to $u$, provided that we adjust the limits of the outer integral to be correct for $u$. Both $v + u$ and $v$ range from $-T/2$ to $T/2$, so $u$ ranges from $-T$ to $T$. {if $v = -T/2$ & $v + u = T/2$, then $u = T$; and if $v = T/2$ & $v + u = -T/2$, then $u = -T$}. Therefore:

$$T \left| Z(n) \right|^2 = \frac{1}{T} \int_{-T}^{T} \int_{-T/2}^{T/2} X(v) X(v + u) e^{-j\omega_0 u} dv \, du$$
3) Notice that $v$ does not always range from $-T/2$ to $T/2$. Instead, the range of $v$ depends on the value of $u$. Therefore, the limits of the inner integral can be adjusted to reflect this fact. Specifically, the range of $v$ is restricted by $u$ such that:

(a) when $0 < u < T$, $v$ ranges from $-T/2$ to $(T/2 - u)$.
(b) when $0 > u > -T$, $v$ ranges from $(T/2 - u)$ to $T/2$.

Therefore:

$$T \left| Z(n) \right|^2 = \int_{0}^{T} \left[ \frac{1}{T} \int_{-T/2}^{(T/2)-u} X(v) X(v + u) \, dv \right] e^{-jn_0 u} \, du$$

$$+ \int_{0}^{0} \left[ \frac{1}{T} \int_{-(T/2)-u}^{T/2} X(v) X(v + u) \, dv \right] e^{-jn_0 u} \, du$$

The expressions within the square brackets now are the positive and negative halves of the sample acvf.
Using the definition of the sample acvf, and letting $u = \tau$, this equation becomes:

$$\hat{C}_{xx}(n) = T \left| Z(n) \right|^2 = \int_{-T}^{T} \hat{C}_{xx}(\tau) e^{-j\omega_0 \tau} d\tau$$

This is the definition of the correlogram (i.e. the Fourier transform of the sample acvf). Therefore, we have shown that the periodogram and correlogram methods for determining the sample spectrum are equivalent. That is, the Fourier transform of the time series times its complex conjugate is equal to the Fourier transform of the sample autocovariance function.
Is the Sample Power Spectrum Consistent?

We will now determine whether the sample spectrum is a consistent estimator of the true spectrum.

Remember that both bias and mean squared error decrease with increasing sample size for a consistent estimator. We will determine whether the (finite) sample spectrum converges to the (infinite) true spectrum as $T$ tends to infinity.
Remember that the true power spectrum, $\Gamma_{XX}(\omega)$, is defined for a continuous range of $\tau$ from negative infinity to positive infinity, whereas the sample spectrum is defined over a finite range, $T$. The true power spectrum:

$$\Gamma_{xx}(\omega) = \int_{-\infty}^{\infty} \gamma_{xx}(\tau) e^{-j\omega \tau} d\tau$$

The correlogram spectral estimator:

$$\hat{C}_{xx}(\tau) = \int_{-T}^{T} \hat{C}_{xx}(\tau) e^{-j\omega \tau} d\tau$$
We will consider the bias of the sample acvf estimator, and then use the sample acvf bias to determine the bias of the sample spectrum.

But first, remember that the bias of any estimator $\hat{\phi}$ is the difference between its expected value and the true value $\phi$:

$$\text{Bias} \equiv \Delta \phi = E[\hat{\phi}] - \phi$$
The estimator is said to be unbiased if it has zero bias, i.e., if its expected value is equal to the true value.
Example of an unbiased estimator: the sample mean of a Gaussian random variable $X$

$$ m \equiv \mathbb{E}[X] $$

$$ \hat{m}_n \equiv \frac{1}{N} \sum_{i=1}^{N} X_i $$

$$ E[\hat{m}_N] = E\left[ \frac{1}{N} \sum_{i=1}^{N} X_i \right] = \frac{1}{N} \sum_{i=1}^{N} E[X_i] $$
Since the random variable is Gaussian:

\[ E[X_i] = E[X] \]

Therefore:

\[ E[\hat{m}_N] = E[X] = m \]

This shows that the expected value of the estimator of the mean is equal to the true mean. In other words, **the sample mean is an unbiased estimator.**
Bias of the Sample Autocovariance Function

Our goal is to determine whether the sample spectrum (derived from realizations of a stochastic process) is an unbiased estimator of the true power spectrum.
Determining the bias of the sample spectrum depends on first determining the bias in the estimation of the autocovariance function (acvf). Therefore, we will first determine whether the acvf estimator is unbiased.
The first moment of the sample acvf is:

\[
\text{E}\left[ \hat{C}_{xx}(\tau) \right] = \text{E}\left[ \frac{1}{T} \int_{0}^{T} X(t) X(t + |\tau|) dt \right]
\]

\[
= \frac{1}{T} \int_{0}^{T} \text{E}\left[ X(t) X(t + |\tau|) \right] dt = \frac{1}{T} \int_{0}^{T} \gamma_{XX}(\tau) dt
\]

where \( \gamma_{XX}(\tau) \) is the covariance between \( X(t) \) and \( X(t+\tau) \) [i.e. the true acvf of \( x(t) \)].
Hence:

$$E[\hat{C}_{xx}(\tau)] = \begin{cases} 
\gamma_{xx}(\tau) \left(1 - \frac{|\tau|}{T}\right), & 0 \leq |\tau| \leq T \\
0, & |\tau| > T 
\end{cases}$$
Conclusions

1. $\hat{C}_{xx}(\tau)$ is a **biased estimator**, since its expected value is NOT equal to the true acvf, $\gamma_{xx}(\tau)$.

2. However, $\hat{C}_{xx}(\tau)$ is an **asymptotically unbiased estimator**, since it approaches the true acvf, $\gamma_{xx}(\tau)$, as the sample length $T$ goes to infinity. This means that the estimator becomes more accurate as $T$ increases.
Bias of the Sample Spectrum

Taking the expected value of the sample spectrum, using the correlogram definition, and rearranging the order of expectation and integration:

\[
E\left[\hat{C}_{xx}(n)\right] = \int_{-T}^{T} E\left[\hat{C}_{xx}(\tau)\right] e^{-j\omega\tau} d\tau
\]
Substituting:

\[
[\hat{\mathcal{C}}_{xx}(\tau)] = \begin{cases} 
\gamma_{xx}(\tau) \left(1 - \frac{|\tau|}{T}\right), & 0 \leq |\tau| \leq T \\
0, & |\tau| > T
\end{cases}
\]

into:

\[
E[\hat{\mathcal{C}}_{xx}(n)] = \int_{-T}^{T} E[\hat{\mathcal{C}}_{xx}(\tau)] e^{-j\omega_0 \tau} d\tau
\]

gives:

\[
E[\hat{\mathcal{C}}_{xx}(n)] = \int_{-T}^{T} \gamma_{xx}(\tau) \left(1 - \frac{|\tau|}{T}\right) e^{-j\omega_0 \tau} d\tau
\]
The expected value of the sample power spectrum does NOT equal the true power spectrum of the stochastic process.

Thus, we see that, like the sample acvf, the sample spectrum is biased.

\[
E\left[ \hat{C}_{xx}(n) \right] = \int_{-T}^{T} \gamma_{xx}(\tau) \left( 1 - \left| \frac{\tau}{T} \right| \right) e^{-j\omega_0 \tau} d\tau
\]
The sample power spectrum is biased. However, in the limit, as the sample length $T$ increases to infinity, the expected value of the sample spectrum approaches the Fourier transform of the true acvf, which is the true power spectrum of the stochastic process:

$$\lim_{T \to \infty} E\left[ \hat{C}_{xx}(n) \right] = \int_{-\infty}^{\infty} \gamma_{xx}(\tau) e^{-j\omega \tau} d\tau = \Gamma(\omega)$$
Conclusions

1. The sample spectrum is a biased estimator of the true power spectrum.

2. The bias approaches zero as the sample size $T$ goes to infinity. Therefore, the sample spectrum is an asymptotically unbiased estimator.
Variance and Consistency

A consistent estimator is unbiased (its expected value approaches the true value of the quantity being estimated), and its variance also decreases to 0 as the sample size increases to infinity.

If the variance decreases with increasing sample size, we can have greater confidence in the estimator as more points are used to form the estimate. Greater confidence allows expression of the estimator value with greater precision.
The variance of an estimator is defined as:

$$\text{Var}(\hat{\phi}) = \mathbb{E} \left[ (\hat{\phi} - \mathbb{E}[\hat{\phi}])^2 \right]$$

An unbiased estimator whose variance also decreases as the sample size $N$ increases:

$$\lim_{N \to \infty} \text{Var}(\hat{\phi}) = 0$$

is consistent.

Since the bias and variance of a consistent estimator both approach zero as the sample size $N$ approaches infinity, the accuracy and precision of the estimator increase without limit (J&W, p. 99). This means that the sampling distribution tends to be centered at $\phi$ and to have a small dispersion around $\phi$. 
Example of a consistent estimator: the sample mean of a Gaussian random variable

We have already established that the sample mean of a Gaussian distribution is unbiased. Now, consider its variance:

$$\text{Var}(\hat{m}_N) = E\left[\left(\hat{m}_N - E[\hat{m}_N]\right)^2\right]$$

$$= E\left[\left(\frac{1}{N}\sum_{i=1}^{N}X_i - E[X]\right)^2\right] = E\left[\left(\frac{1}{N}\sum_{i=1}^{N}(X_i - E[X])\right)^2\right]$$
\[
\begin{align*}
\mathbb{E} & \left[ \frac{1}{N^2} \left( \sum_{i=1}^{N} (X_i - \mathbb{E}[X]) \right)^2 \right] \\
& = \frac{1}{N^2} \sum_{i=1}^{N} \mathbb{E} \left[ (X_i - \mathbb{E}[X])^2 \right] = \frac{1}{N^2} \sum_{i=1}^{N} \text{Var}(X) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2 \\
& = \frac{1}{N^2} N \sigma^2 = \frac{\sigma^2}{N} \\
\therefore \quad \lim_{N \to \infty} \text{Var} \left( \hat{m}_N \right) &= 0
\end{align*}
\]

Note that we have used the previous result showing that the Gaussian sample mean is unbiased, meaning that the expected value of the Gaussian sample mean equals the true mean.
We conclude that the Gaussian sample mean is a consistent estimator, since it is unbiased and its variance goes to zero as the sample size $N$ goes to infinity.
Variance and Consistency of the Sample Autocovariance Function

We have established that the sample acvf is asymptotically unbiased. To determine whether it is consistent, we must also determine its variance. Following Jenkins & Watts (pp. 175-179), we first derive its covariance, and then show that the variance follows directly from the covariance.
We begin by defining the covariance between
\( \hat{C}_{xx}(\tau_1) \) and \( \hat{C}_{xx}(\tau_2) \), autocovariance estimates at two different lags \( \tau_1 \) and \( \tau_2 \), where \( \tau_2 \geq \tau_1 \geq 0 \) [in Jenkins & Watts, p. 412, it is shown that this condition is not restrictive in any case):

\[
\text{Cov}\left[ \hat{C}_{xx}(\tau_1), \hat{C}_{xx}(\tau_2) \right] = \text{Cov}\left[ \frac{1}{T} \int_0^T X(t)X(t+\tau_1)\,dt, \frac{1}{T} \int_0^T X(v)X(v+\tau_2)\,dv \right] = \frac{1}{T} \int \int \text{Cov}\left[ X(t)X(t+\tau_1), X(v)X(v+\tau_2) \right] dv dt
\]
We now use the property of a stationary stochastic process that:

\[
\text{Cov} \left[ X(t)X(t+\tau_1), X(v)X(v+\tau_2) \right]
\]

\[
= \gamma_{xx}(v-t)\gamma_{xx}(v-t+\tau_2-\tau_1) + \gamma_{xx}(v-t+\tau_2)\gamma_{xx}(v-t-\tau_1)
+ K_4(v-t,\tau_1\tau_2)
\]

where \( \gamma_{xx}(\tau) = \text{Cov}[ x(t), x(t+\tau)] \).

We can drop the higher order term \( K_4 \) - the fourth joint cumulant of the stochastic process \( X(t) \) - by assuming that \( X(t) \) is Gaussian.
By substitution, we arrive at:

\[
\text{Cov}\left[\hat{C}_{xx}(\tau_1), \hat{C}_{xx}(\tau_2)\right] =
\]

\[
\frac{1}{T^2} \int_0^{T-\tau_1} \int_0^{T-\tau_2} \left\{ \gamma_{xx}(v-t)\gamma_{xx}(v-t+\tau_2-\tau_1) + \gamma_{xx}(v-t+\tau_2)\gamma_{xx}(v-t-\tau_1) \right\} \, dv \, dt
\]
We now use the transformation:
\[ r = v - t, \quad s = t. \]

Since:
- \( v \) ranges from 0 to \( T - \tau_2 \)
- \( t \) ranges from 0 to \( T - \tau_1 \),
- \( r \) ranges from \((-T+\tau_1)\) to \((T-\tau_2)\).

{if \( v = 0 \) & \( t = T - \tau_1 \), then
  \( r = (-T+\tau_1); \) and
if \( v = T - \tau_2 \) & \( t = 0 \), then
  \( r = (T - \tau_2)\)
The previous equation then becomes:

$$\text{Cov}\left[\hat{C}_{xx}(\tau_1), \hat{C}_{xx}(\tau_2)\right] =$$

$$\frac{1}{T^2} \int_{-T+\tau_2}^{T+\tau_2} \left\{ \gamma_{xx}(r) \gamma_{xx}(r + \tau_2 - \tau_1) + \gamma_{xx}(r + \tau_2) \gamma_{xx}(r - \tau_1) \right\} dr \int ds$$

Since the integrand in the integral over $s$ is independent of $s$, integration with respect to $s$ yields a line integral with a length that depends on $r$. 
The length of the line integral $\phi_s(r)$ is the horizontal length (over $s$) from the left-most extent to the right-most extent of the transformed integration region for any value of $r$, as given by:

$$\phi_s(r) = \begin{cases} 
T - \tau_2 - r, & r \geq 0 \\
T - \tau_2, & -(\tau_2 - \tau_1) \leq r \leq 0 \\
T - \tau_1 + r, & -(T - \tau_1) \leq r \leq -(\tau_2 - \tau_1)
\end{cases}$$

We see that the line integral length is defined differently for 3 different sections of the integration region, marked as red, blue, and green.
The previous two equations lead to:

\[
\text{Cov}\left[ \hat{C}_{xx}(\tau_1), \hat{C}_{xx}(\tau_2) \right] = \\
\frac{1}{T^2} \int_{T-\tau_2}^{T-\tau_1} \phi_s(r) \left\{ \gamma_{xx}(r) \gamma_{xx}(r + \tau_2 - \tau_1) + \gamma_{xx}(r + \tau_2) \gamma_{xx}(r - \tau_1) \right\} dr
\]

This result shows that, in general, adjacent acvf estimates (i.e. successive lag values) have strong correlations. Consequently, the autocovariance function may fail to damp out with increasing lag.
When $\tau_1 = \tau_2$, we obtain an expression for the variance of the acvf estimator:

$$
\text{Var}\left[\hat{C}_{xx}(\tau)\right] = \frac{1}{T^2} \int_{-\tau}^{\tau} \left( T - \tau - |r| \right) \{ \gamma_{xx}(r) + \gamma_{xx}(r + \tau) \gamma_{xx}(r - \tau) \} \, dr
$$
Conclusions

1. It can be demonstrated (J & W, pp. 178-179), that for large T, the variance of the acvf estimator is proportional to $1/T$. Thus, as T goes to infinity, the variance goes to zero.

2. Since the acvf estimator is asymptotically unbiased and its variance approaches zero as N approaches infinity, we conclude that it is asymptotically consistent, i.e., its expected value approaches the true value, and its variance approaches zero, as the number of points goes to infinity.
Note: Jenkins & Watts (pp. 174-180) also consider an unbiased estimator of the autocovariance function. However, they conclude that the biased estimator that we have considered is a more satisfactory estimator since the variance of the unbiased estimator tends to infinity as $\tau$ tends to $T$ (see Fig. 5.12 on p. 180).
Variance and Consistency of the Sample Spectrum

To compute the variance of the sample spectrum of any arbitrary stochastic process, we first consider that of a complex Gaussian white noise process with variance $\sigma^2$.

We evaluate the covariance between samples of the sample spectrum at two frequencies $\omega_1$ and $\omega_2$. We then derive the variance by moving $\omega_1 \rightarrow \omega_2$. 
The variance of the white noise sample spectrum can be shown to be:

\[ \text{Var}\left[\hat{C}_w(\omega)\right] \approx \sigma^4 \]

This variance is dominated by a constant term, and so does NOT decrease with increasing sample length T. Therefore, the sample spectrum estimator is NOT consistent.
For an *arbitrary Gaussian stochastic process*, the properties of the variance are generally similar to those of Gaussian white noise. The general process can be treated as the output of a linear filter, $h(t)$, driven by white noise, $w(t)$. In the time domain, the output is expressed by the convolution of the filter coefficients and the white noise input.
It can be shown that the variance of the sample spectrum of an arbitrary Gaussian stochastic process is:  
\[ \text{Var} \left[ \hat{C}_{XX}(\omega) \right] \approx \Gamma_{XX}^2(\omega) \]

**Conclusion**

As the sample length \( T \) increases, the variance of the sample spectrum estimator is approximately the same size as the power spectrum that we are estimating. Therefore, the sample spectrum is NOT a consistent estimator: as \( T \) increases, the variance of the estimator does not become small, but rather is always of the same order of magnitude as the power spectrum. We do not gain confidence in the sample spectrum estimator by increasing \( T \). Increasing \( T \) increases frequency resolution and decreases the bias, but does not decrease the variance.
Improving the Spectral Estimate: 
The Bartlett Procedure

We want a way to reduce the variance associated with the sample spectrum and obtain a smooth spectral estimate without random variations resulting from the estimation process. One procedure for accomplishing this goal was provided by Bartlett. He showed that the average of multiple independent sample spectral estimates has a reduced variance [J & W, p. 239].
The Bartlett smoothing procedure is to divide the available time series into a series of blocks, compute a sample spectrum for each, and then average the sample spectra. The result is known as the Bartlett estimator. If there are $K$ blocks, then the (smoothed) Bartlett estimator is:

$$\bar{C}_{xx}(\omega) = \frac{1}{K} \sum_{k=1}^{K} \hat{C}_{xx}^k(\omega)$$
The expected value of the Bartlett estimator equals the expected value of the individual sample spectra.

\[
E[\bar{\mathcal{C}}_{xx}(\omega)] = \frac{1}{K} \sum_{k=1}^{K} E[\hat{\mathcal{C}}_{xx}(\omega)] = E[\hat{\mathcal{C}}_{xx}(\omega)] \frac{1}{K} \sum_{k=1}^{K} 1 = E[\hat{\mathcal{C}}_{xx}(\omega)]
\]

Here, we have used the fact that the expected value of the sample spectrum is the same for all \(K\) to move it outside of the summation.
More importantly, since the Bartlett estimator is the sample mean of the sample spectrum, we can assume that the sample distribution is Gaussian, and then use the result of our previous derivation showing that the variance of the Gaussian sample mean of a random variable is related to the variance of the random variable by:

$$\text{Var}[\bar{C}_{xx}(\omega)] = \frac{\text{Var}[\hat{C}_{xx}(\omega)]}{K}$$
As a consequence of this result, the variance of the Bartlett estimator is less than that of the sample spectral estimator. In the limit, as the number of blocks $K$ goes to infinity, the variance of the Bartlett estimator goes to zero, so the Bartlett estimator is an asymptotically consistent estimator of the power spectrum.
Therefore, a procedure for obtaining a consistent spectral estimator is:

1. Create K data blocks each of length M from the original time series of length N.
2. Apply the DFT to each block and compute its sample (power) spectrum.
3. By the Bartlett smoothing procedure, average the sample spectra.