Estimation of Process Mean $\mu$: I

• stationary process $\{X_t\}$ is characterized by its mean $\mu$ and ACVF $\{\gamma(h)\}$, which, for a particular time series $x_1, \ldots, x_n$, are generally unknown and must be estimated.

• have already introduced sample mean and sample ACVF as appropriate estimators (see overhead II–62).

• will now regard these as realizations of associated RVs whose statistical properties we want to study:

$$
\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t \quad \text{and} \quad \hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n)
$$

• start by studying properties of sample mean $\bar{X}_n$ as estimator of process mean $\mu$. 

\[ V-1 \]
Estimation of Process Mean $\mu$: II

- sample mean is an *unbiased* estimator of $\mu$ since

$$E\{\overline{X}_n\} = \frac{1}{n} \sum_{t=1}^{n} E\{X_t\} = \frac{1}{n} \sum_{t=1}^{n} \mu = \mu$$

- given an estimator $\hat{\alpha}$ of some parameter $\alpha$, can measure how well it estimates $\alpha$ via its *mean square error*:

$$\text{mse}\{\hat{\alpha}\} \equiv E\{(\hat{\alpha} - \alpha)^2\}$$

(many other measures exist!)

- if $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are two competing estimators, would prefer $\hat{\alpha}_1$ over $\hat{\alpha}_2$ if

$$\text{mse}\{\hat{\alpha}_1\} < \text{mse}\{\hat{\alpha}_2\}$$
Estimation of Process Mean $\mu$: III

- because the sample mean is unbiased, its mean square error is just its variance:

$$\text{mse} \left\{ \overline{X}_n \right\} = E\{(\overline{X}_n - \mu)^2\}$$

$$= \text{var} \left\{ \overline{X}_n \right\}$$

$$= \text{cov} \left\{ \overline{X}_n, \overline{X}_n \right\}$$

$$= \text{cov} \left\{ \frac{1}{n} \sum_{r=1}^{n} X_r, \frac{1}{n} \sum_{s=1}^{n} X_s \right\}$$

$$= \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} \text{cov} \left\{ X_r, X_s \right\} = \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} \gamma(r, s)$$
Estimation of Process Mean $\mu$: IV

- can regard double sum as summing all elements of this matrix:

$$
\begin{bmatrix}
\gamma(1, 1) & \gamma(1, 2) & \gamma(1, 3) & \gamma(1, 4) & \cdots & \gamma(1, n) \\
\gamma(2, 1) & \gamma(2, 2) & \gamma(2, 3) & \gamma(2, 4) & \cdots & \gamma(2, n) \\
\gamma(3, 1) & \gamma(3, 2) & \gamma(3, 3) & \gamma(3, 4) & \cdots & \gamma(3, n) \\
\gamma(4, 1) & \gamma(4, 2) & \gamma(4, 3) & \gamma(4, 4) & \cdots & \gamma(4, n) \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma(n, 1) & \gamma(n, 2) & \gamma(n, 3) & \gamma(n, 4) & \cdots & \gamma(n, n)
\end{bmatrix},
$$

which can be reexpressed in terms of ACVF as

$$
\begin{bmatrix}
\gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) & \cdots & \gamma(n - 1) \\
\gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(n - 2) \\
\gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(n - 3) \\
\gamma(3) & \gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(n - 4) \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma(n - 1) & \gamma(n - 2) & \gamma(n - 3) & \gamma(n - 4) & \cdots & \gamma(0)
\end{bmatrix}
$$
Estimation of Process Mean $\mu$: V

• hence

$$\text{var}\{\overline{X}_n\} = \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} \gamma(r - s) = \frac{1}{n^2} \sum_{h=-(n-1)}^{n-1} (n - |h|) \gamma(h)$$

$$= \frac{1}{n} \sum_{h=-(n-1)}^{n-1} \left(1 - \frac{|h|}{n}\right) \gamma(h)$$

• as $n \to \infty$, if sum converges finitely, then $\text{var}\{\overline{X}_n\} \to 0$; i.e., $\overline{X}_n$ converges in mean square to $\mu$ (one form of consistency)

• in addition, as $n \to \infty$, if $\sum_{|h|=\infty}^{\infty} |\gamma(h)| < \infty$, then

$$n \text{var}\{\overline{X}_n\} \to \sum_{h=\infty}^{\infty} \gamma(h)$$
Estimation of Process Mean $\mu$: VI

• based upon $n \text{var} \{\overline{X}_n\} \rightarrow \sum_h \gamma(h)$, have, for large $n$,

$$\text{var} \{\overline{X}_n\} \approx \frac{v}{n}, \text{ where } v = \sum_{h=-\infty}^{\infty} \gamma(h),$$

which is a useful approximation if $v > 0$

• there are some processes for which $v = 0$ (e.g., $X_t = Z_t - Z_{t-1}$), in which case need to back up to exact expression for $\text{var} \{\overline{X}_n\}$

• for many time series of interest, $\overline{X}_n$ is approximately $\mathcal{N}(\mu, v/n)$, so an approximate 95% confidence interval (CI) for $\mu$ is

$$\left[ \overline{X}_n - 1.96 \frac{\sqrt{v}}{\sqrt{n}}, \overline{X}_n + 1.96 \frac{\sqrt{v}}{\sqrt{n}} \right]$$

• $v$ not generally known, so must be estimated
Estimation of Process Mean $\mu$: VII

- to estimate

$$v = \sum_{h=-\infty}^{\infty} \gamma(h),$$

can entertain using $\hat{\gamma}(h)$ to estimate $\gamma(h)$ for $|h| \leq n - 1$

- $\gamma(h)$ problematic for $|h| \geq n$

- usual practice is to assume these are zero, which is in keeping with $\hat{\gamma}(h)$ for $h$ close to $n - 1$; e.g.,

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n)$$

becomes, for $h = n - 1$,

$$\hat{\gamma}(n - 1) = \frac{(X_n - \bar{X}_n)(X_1 - \bar{X}_n)}{n}$$
Sample ACVF for Gaussian IID(0,1) \( \{x_t\} \ (n = 100) \)
Estimation of Process Mean $\mu$: VIII

- natural estimator for

$$v = \sum_{h=-\infty}^{\infty} \gamma(h)$$

would seem to be

$$\hat{v} = \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h)$$

- disaster strikes: homework exercise says that $\hat{v} = 0$ always!

- to patch up disaster, can use something like

$$\hat{v} = \sum_{h=-\lceil\sqrt{n}\rceil}^{\lceil\sqrt{n}\rceil} \left(1 - \frac{|h|}{\sqrt{n}}\right) \hat{\gamma}(h),$$

where $[x]$ means ‘$x$ rounded to nearest integer’

- some justification for above provided in Stat/EE 520

- rather than taking this nonparametric approach, can assume, e.g., an AR(1) model and estimate $v$ as dictated by model
Estimation of Process Mean $\mu$: IX

• recall that an AR(1) process $\{X_t\}$ with mean $\mu$ satisfies
\[ X_t - \mu = \phi(X_{t-1} - \mu) + Z_t \]
with $|\phi| < 1$ & $\{Z_t\} \sim \text{WN}(0, \sigma^2)$ (see overhead II–47)

• have argued that
  - $\rho(h) = \phi|h|$ (overhead II–49)
  - $\gamma(0) = \sigma^2/(1 - \phi^2)$ (overhead II–50)

• since $\rho(h) = \gamma(h)/\gamma(0)$ have
\[ \gamma(h) = \rho(h)\gamma(0) = \frac{\phi|h|\sigma^2}{1 - \phi^2} \]
Estimation of Process Mean $\mu$: X

- using $\gamma(h) = \phi|h|\sigma^2/(1 - \phi^2)$ leads to

$$v = \sum_{h=-\infty}^{\infty} \gamma(h) = \frac{\sigma^2}{1 - \phi^2} \left(1 + 2 \sum_{h=1}^{\infty} \phi^h\right) = \frac{\sigma^2}{(1 - \phi)^2}$$

after using $\sum_{h=1}^{\infty} \phi^h = \phi/(1 - \phi)$ and doing some algebra

- leads to 95% CI of form

$$\left[\bar{X}_n - 1.96\frac{\sigma}{(1 - \phi)\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{(1 - \phi)\sqrt{n}}\right]$$

- let’s see how nonparametric and parametric approaches to forming CIs for $\mu$ work on a wind speed time series
Wind Speed Time Series
Sample ACF for Wind Speed Series
Estimation of Process Mean $\mu$: XI

- sample mean for wind speed series is $\bar{x}_n = -1.37$
- if we assume Gaussian IID($\mu, \sigma^2_X$), 95% CI for $\mu$ is given by
  $$
  \left[ \bar{x}_n - 1.96 \frac{\sigma_X}{\sqrt{n}}, \bar{x}_n + 1.96 \frac{\sigma_X}{\sqrt{n}} \right]
  $$
- table below compares this CI with ones based on approaches just discussed (with suitable estimates substituted for unknown parameters, thus rendering all approximate 95% CIs)
- note: $\hat{\phi} = \hat{\rho}(1) = 0.891$

<table>
<thead>
<tr>
<th></th>
<th>lower bound</th>
<th>upper bound</th>
<th>CI width</th>
<th>ratio to AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IID</td>
<td>$-1.65$</td>
<td>$-1.10$</td>
<td>$0.55$</td>
<td>$0.25$</td>
</tr>
<tr>
<td>nonparametric</td>
<td>$-2.12$</td>
<td>$-0.63$</td>
<td>$1.49$</td>
<td>$0.69$</td>
</tr>
<tr>
<td>AR(1)</td>
<td>$-2.45$</td>
<td>$-0.29$</td>
<td>$2.16$</td>
<td>$1.00$</td>
</tr>
</tbody>
</table>