Estimation of Process Mean μ: I

- stationary process $\{X_t\}$ is characterized by its mean μ 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–64)
- will now regard these as realizations of associated RVs whose statistical properties we want to study:

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

- start by studying properties of sample mean $\overline{X}_n$ as estimator of process mean μ

Estimation of Process Mean μ: II

- sample mean is an *unbiased* estimator of μ 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 α, can measure how well it estimates α via its *mean square error*:

$$\text{mse}\,\{\hat{\alpha}\} \stackrel{\text{def}}{=} E\left\{(\hat{\alpha} - \alpha)^2\right\}$$

(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

$$\text{mse}\,\{\hat{\alpha}_1\} < \text{mse}\,\{\hat{\alpha}_2\}$$

BD: 50, CC: 28, SS: 27

Estimation of Process Mean μ: III

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

$$
\begin{aligned}
\operatorname{mse}\{\overline{X}_n\} &= E\{(\overline{X}_n - \mu)^2\} \\
&= \operatorname{var}\{\overline{X}_n\} \\
&= \operatorname{cov}\{\overline{X}_n, \overline{X}_n\} \\
&= \operatorname{cov}\left\{\frac{1}{n}\sum_{r=1}^{n} X_r, \frac{1}{n}\sum_{s=1}^{n} X_s\right\} \\
&= \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} \operatorname{cov}\{X_r, X_s\} = \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} \gamma(r, s)
\end{aligned}
$$

Estimation of Process Mean μ: 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 μ: V

- hence
$$\mathrm{var}\,\{\overline{X}_n\} = \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n}\gamma(r-s) = \frac{1}{n^2}\sum_{h=-(n-1)}^{n-1}(n-|h|)\gamma(h)$$
$$= \frac{1}{n}\sum_{h=-(n-1)}^{n-1}\left(1-\frac{|h|}{n}\right)\gamma(h)$$
- as $n \to \infty$, if sum converges finitely, then $\mathrm{var}\,\{\overline{X}_n\} \to 0$; i.e., $\overline{X}_n$ converges in mean square to μ (one form of *consistency*)
- in addition, as $n \to \infty$, if $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$, then
$$n\,\mathrm{var}\,\{\overline{X}_n\} \to \sum_{h=-\infty}^{\infty}\gamma(h)$$

Estimation of Process Mean μ: VI

- based upon $n \operatorname{var}\{\overline{X}_n\} \to \sum_h \gamma(h)$, have, for large n,

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

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

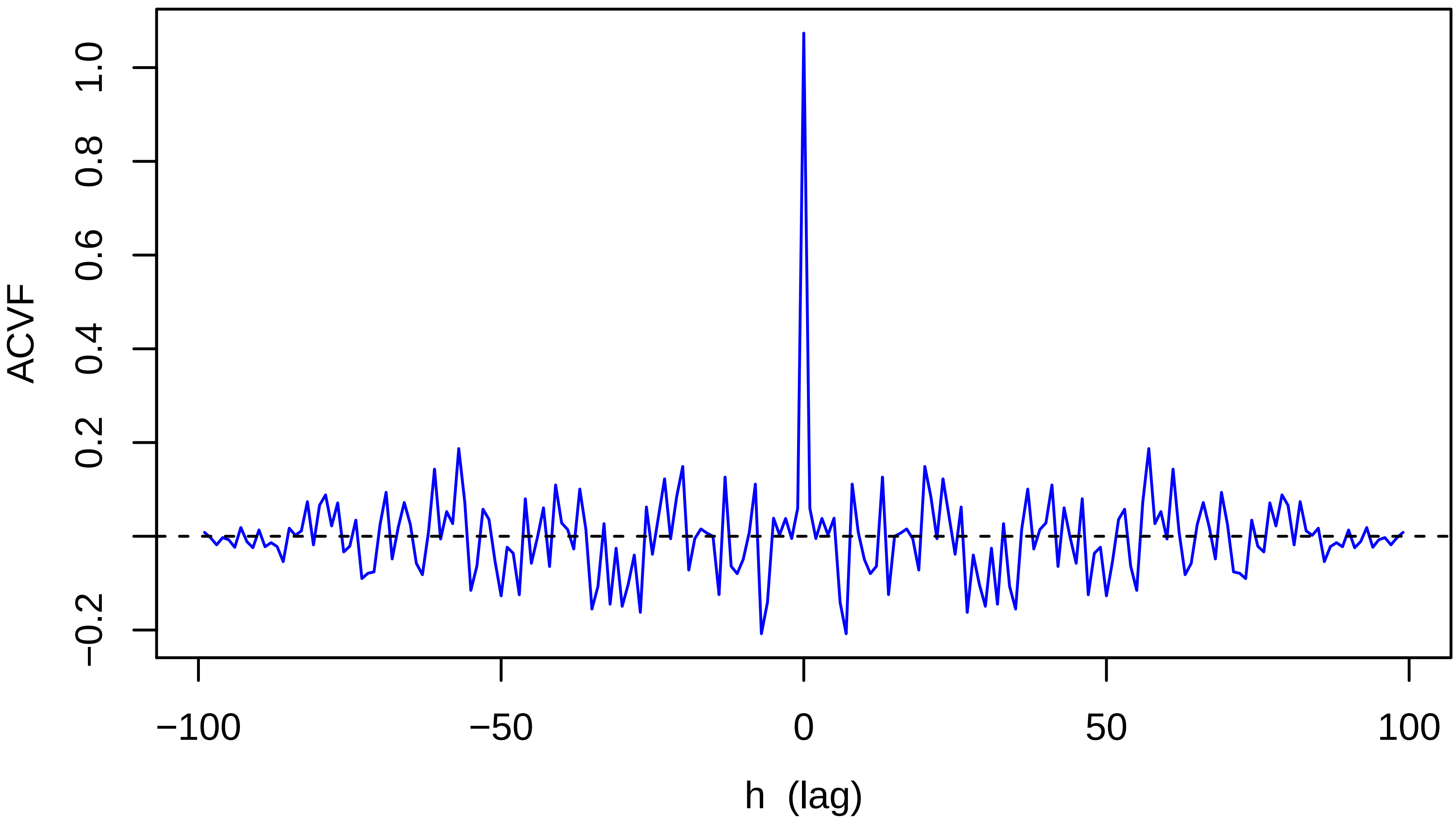
$$\left[\overline{X}_n - 1.96\frac{\sqrt{v}}{\sqrt{n}}, \overline{X}_n + 1.96\frac{\sqrt{v}}{\sqrt{n}}\right]$$

- v not generally known, so must be estimated

Estimation of Process Mean μ: 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.,
$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \overline{X}_n)(X_t - \overline{X}_n)$$
becomes, for $h = n-1$,
$$\hat{\gamma}(n-1) = \frac{(X_n - \overline{X}_n)(X_1 - \overline{X}_n)}{n}$$

Sample ACVF for Gaussian IID(0,1) $\{x_t\}$ **(**$n = 100$**)**



Estimation of Process Mean μ: VIII

- natural estimator for

$$v = \sum_{h=-\infty}^{\infty} \gamma(h) \text{ would seem to be } \hat{v} = \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h)$$

- disaster strikes: exercise says that $\hat{v} = 0$ always!
- to patch up disaster, can use something like

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

  where $[x]$ means 'x rounded to nearest integer'
  - B&D claim $\hat{v}$ is good approximation to v for large n
- 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 μ: IX

- recall that an AR(1) process $\{X_t\}$ with mean μ satisfies

$$X_t - \mu = \phi(X_{t-1} - \mu) + Z_t$$

with $|\phi| < 1$ & $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ (see overhead II–48)

- have argued that
  - $\rho(h) = \phi^{|h|}$ (overhead II–50)
  - $\gamma(0) = \sigma^2/(1 - \phi^2)$ (overhead II–51)
- 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 μ: X

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

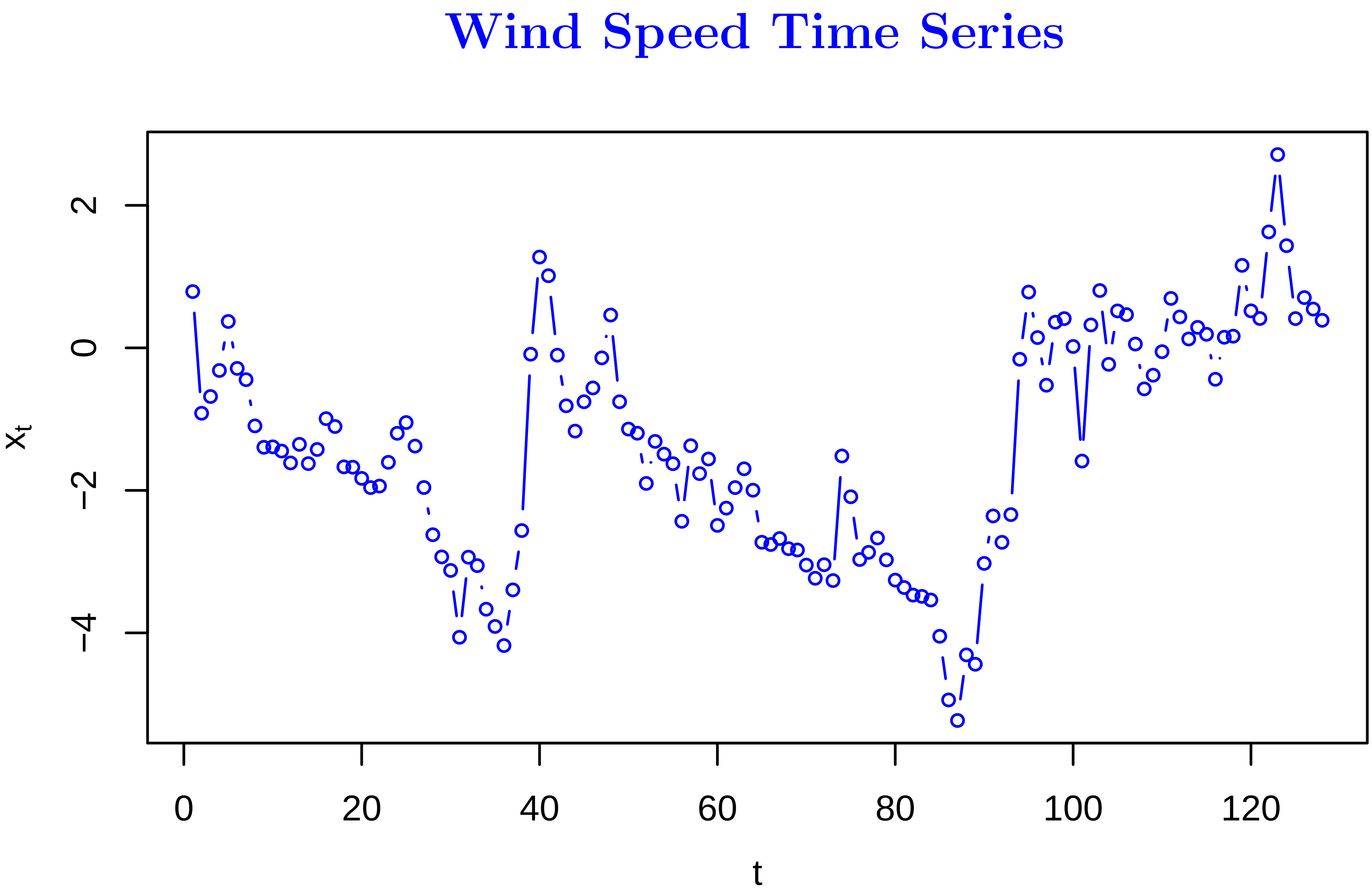
$$v = \sum_{h=-\infty}^{\infty} \gamma(h) = \frac{\sigma^2}{1-\phi^2}\left(1 + 2\sum_{h=1}^{\infty} \phi^h\right) = \frac{\sigma^2}{(1-\phi)^2}$$

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

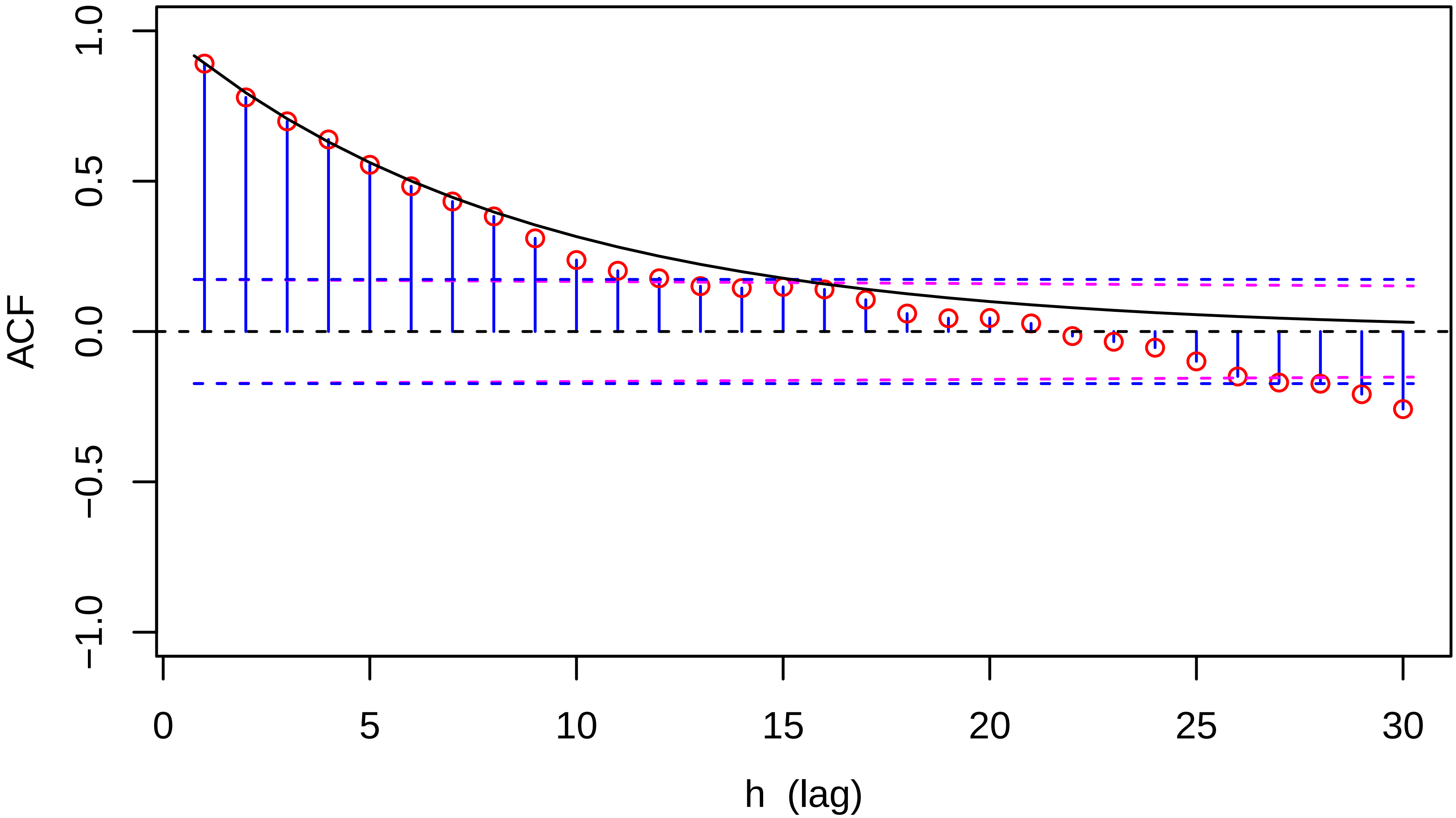
- leads to 95% CI of form

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

- let's see how nonparametric and parametric approaches to forming CIs for μ work on a wind speed time series


Wind Speed Time Series
x_t
2
0
−2
−4
0
20
40
60
80
100
120
t

Sample ACF for Wind Speed Series



Estimation of Process Mean μ: XI

- sample mean for wind speed series is $\bar{x}_n \doteq -1.37$
- if we assume Gaussian IID(μ, σ_X^2), 95% CI for μ is given by

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

- 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: $\hat{\phi} = \hat{\rho}(1) \doteq 0.891$

	lower bound	upper bound	CI width	ratio to AR
IID	-1.65	-1.10	0.55	0.25
nonparametric	-2.12	-0.63	1.49	0.69
AR(1)	-2.45	-0.29	2.16	1.00