Estimation of ACVF and ACF: I

- given a time series presumed to be a realization of a portion $X_1, X_2, \ldots, X_n$ of a stationary process, overheads II–64 and II–65 stated definitions for a sample ACVF and ACF that are realizations of the RVs

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \overline{X}_n)(X_t - \overline{X}_n) \text{ and } \hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

- here we look at these estimators of the ACVF $\gamma(h)$ and the ACF $\rho(h) = \gamma(h)/\gamma(0)$ in more detail
- prior to doing so, let's review three basic properties of the ACVF we've mentioned already, and then introduce a fourth

Four Basic Properties of ACVF $\{\gamma(h)\}$: I

1. $\gamma(0) \geq 0$ (since $\gamma(0) = \text{var}\,\{X_t\}$, restatement of $\text{var}\,\{X_t\} \geq 0$)
2. $|\gamma(h)| \leq \gamma(0)$ for all h (since $\rho(h) = \gamma(h)/\gamma(0)$ and $|\rho(h)| \leq 1$ because it is a correlation coefficient – see overhead II–16)
3. $\gamma(-h) = \gamma(h)$, i.e., $\gamma(h)$ is an even function (overhead II–15)
4. sequence $\{\gamma(h)\}$ is *nonnegative definite* – by definition this means that, for any positive integer n, if $t_1, t_2, \ldots, t_n$ are any n integers and if $a_1, a_2, \ldots, a_n$ are any n real-valued numbers, then we must have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(t_i - t_j) \geq 0$$

- note: same concept sometimes called *positive semidefinite*

Four Basic Properties of ACVF $\{\gamma(h)\}$: II

- to see that property 4 is true, consider $Y \stackrel{\text{def}}{=} \sum_{i=1}^{n} a_i X_{t_i}$, i.e., a linear combination of n RVs arbitrarily picked from $\{X_t\}$
- recalling that var $\{Y\}$ must be nonnegative, note that

$$\begin{aligned}
\operatorname{var}\{Y\} = \operatorname{cov}\{Y, Y\} &= \operatorname{cov}\left\{\sum_{i=1}^{n} a_i X_{t_i}, \sum_{j=1}^{n} a_j X_{t_j}\right\} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \operatorname{cov}\{X_{t_i}, X_{t_j}\} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(t_i - t_j),
\end{aligned}$$

thus showing that $\{\gamma(h)\}$ is nonnegative definite

Four Basic Properties of ACVF $\{\gamma(h)\}$: III

- let $\boldsymbol{a}$ be a column vector whose elements are a_1, a_2, $\ldots$, a_n, and let $\boldsymbol{a}'$ denote its transpose (an n-dimensional row vector)
- let Γ be an $n \times n$ matrix whose (i, j)th element is $\gamma(t_i - t_j)$
- nonnegative definiteness condition

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(t_i - t_j) \geq 0$$

can be restated as

$$\boldsymbol{a}' \Gamma \boldsymbol{a} \geq 0$$

Four Basic Properties of ACVF $\{\gamma(h)\}$: IV

- $\boldsymbol{a}'\Gamma\boldsymbol{a} \geq 0$ holds for arbitrary $\boldsymbol{a}$
- if $\boldsymbol{a}$ is an eigenvector for Γ, then $\Gamma\boldsymbol{a} = \lambda\boldsymbol{a}$, where λ is the eigenvalue corresponding to $\boldsymbol{a}$
- assuming the usual eigenvector normalization $\boldsymbol{a}'\boldsymbol{a} = 1$, we have

$$0 \leq \boldsymbol{a}'\Gamma\boldsymbol{a} = \boldsymbol{a}'(\lambda\boldsymbol{a}) = \lambda\boldsymbol{a}'\boldsymbol{a} = \lambda,$$

  which shows that nonnegative definiteness implies that the eigenvalues corresponding to the associated Γ must be nonnegative

Four Basic Properties of ACVF $\{\gamma(h)\}$: V

- theorem: a real-valued function defined on the integers is the ACVF for *some* stationary process if and only if it is even and nonnegative definite
- overhead VI–3 establishes one part of theorem (if $\{\gamma(h)\}$ is an ACVF, then it is nonnegative definite)
- second part (if $\{\gamma(h)\}$ is nonnegative definite, then there is a stationary process that has $\{\gamma(h)\}$ as its ACVF) is more difficult to establish (see B&D for details)
  - in fact, can show that, if $\{\gamma(h)\}$ is nonnegative definite, there exists a *Gaussian* stationary process having $\{\gamma(h)\}$ as its ACVF (i.e., any finite collection of RVs from $\{X_t\}$ obeys a multivariate normal distribution)

Four Basic Properties of ACVF $\{\gamma(h)\}$: VI

- given an arbitrary even function $\{\kappa(h)\}$ defined on the integers, it is usually quite difficult to show that it is nonnegative definite based directly on the definition of this concept (Example 2.1.1 of B&D gives a rare instance where this approach works)
- practical approaches for establishing nonnegative definiteness:
  1. find a stationary process that has $\{\kappa(h)\}$ as its ACVF, which means $\{\kappa(h)\}$ must be nonnegative definite (thus $\{\cos(ch)\}$ is such because it is the ACVF for stationary process $X_t = Z_2 \cos(ct) + Z_1 \sin(ct)$ considered in Problem 2(b))
  2. show that $\{\kappa(h)\}$ arises from an *integrated spectrum* and appeal to a theorem relating such spectra to nonnegative definite functions (theorem is a powerful result from spectral analysis)

Estimation of ACVF and ACF: II

- recall that, for $|h| \le n-1$,

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

- expression for $E\{\hat{\gamma}(h)\}$ is messy, so let's consider instead

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

for which

$$E\{\bar{\gamma}(h)\} = \frac{1}{n}\sum_{t=1}^{n-|h|}E\{(X_{t+|h|}-\mu)(X_t-\mu)\} = \frac{n-|h|}{n}\gamma(h) \neq \gamma(h)$$

in general when $h \neq 0$; i.e., $\bar{\gamma}(h)$ is a biased estimator

Estimation of ACVF and ACF: III

- rather than using $\bar{\gamma}(h)$, might seem more natural to consider

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

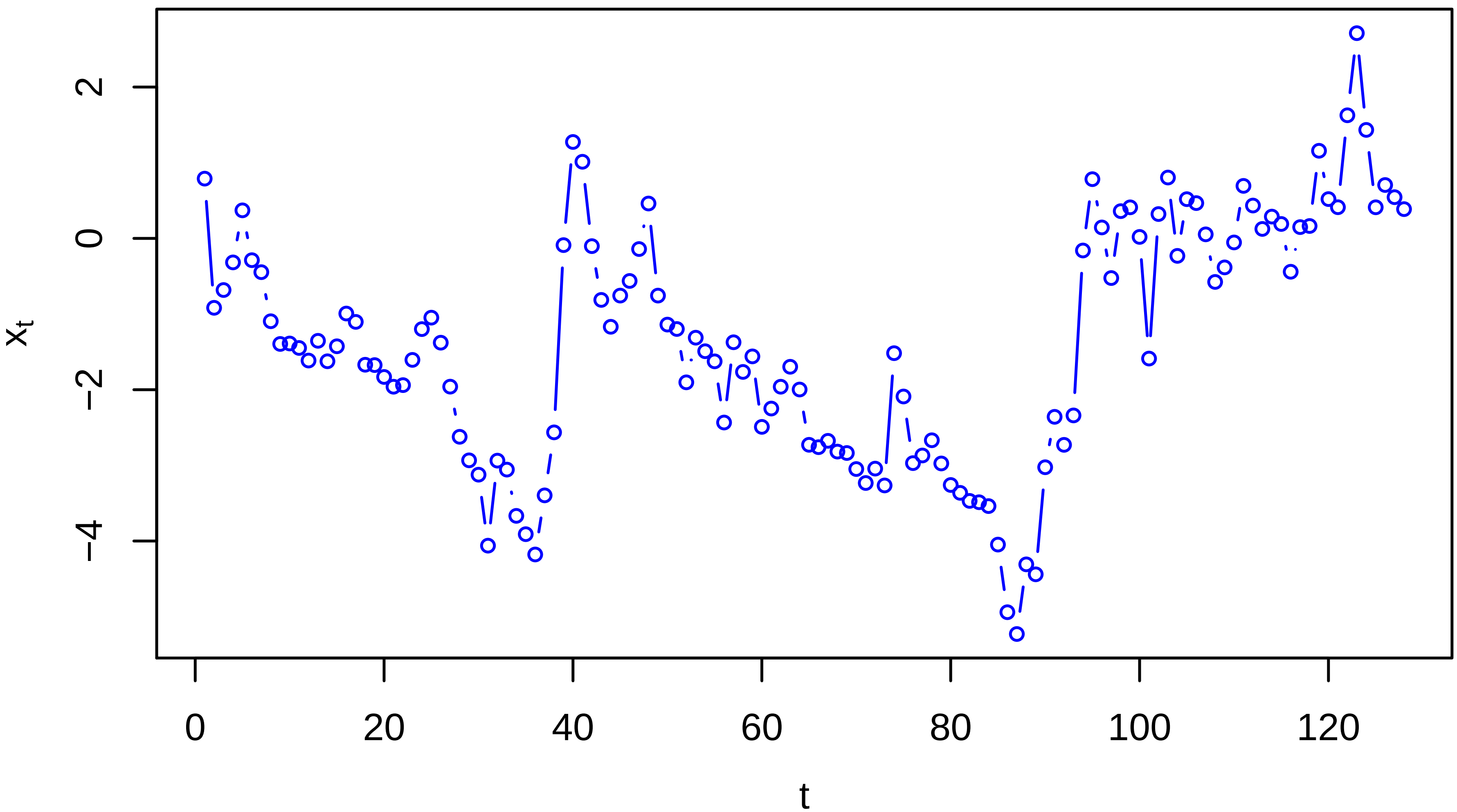
  which differs from $\bar{\gamma}(h)$ only in its divisor, and for which

$$E\{\tilde{\gamma}(h)\} = \frac{1}{n-|h|} \sum_{t=1}^{n-|h|} E\{(X_{t+|h|} - \mu)(X_t - \mu)\} = \gamma(h);$$

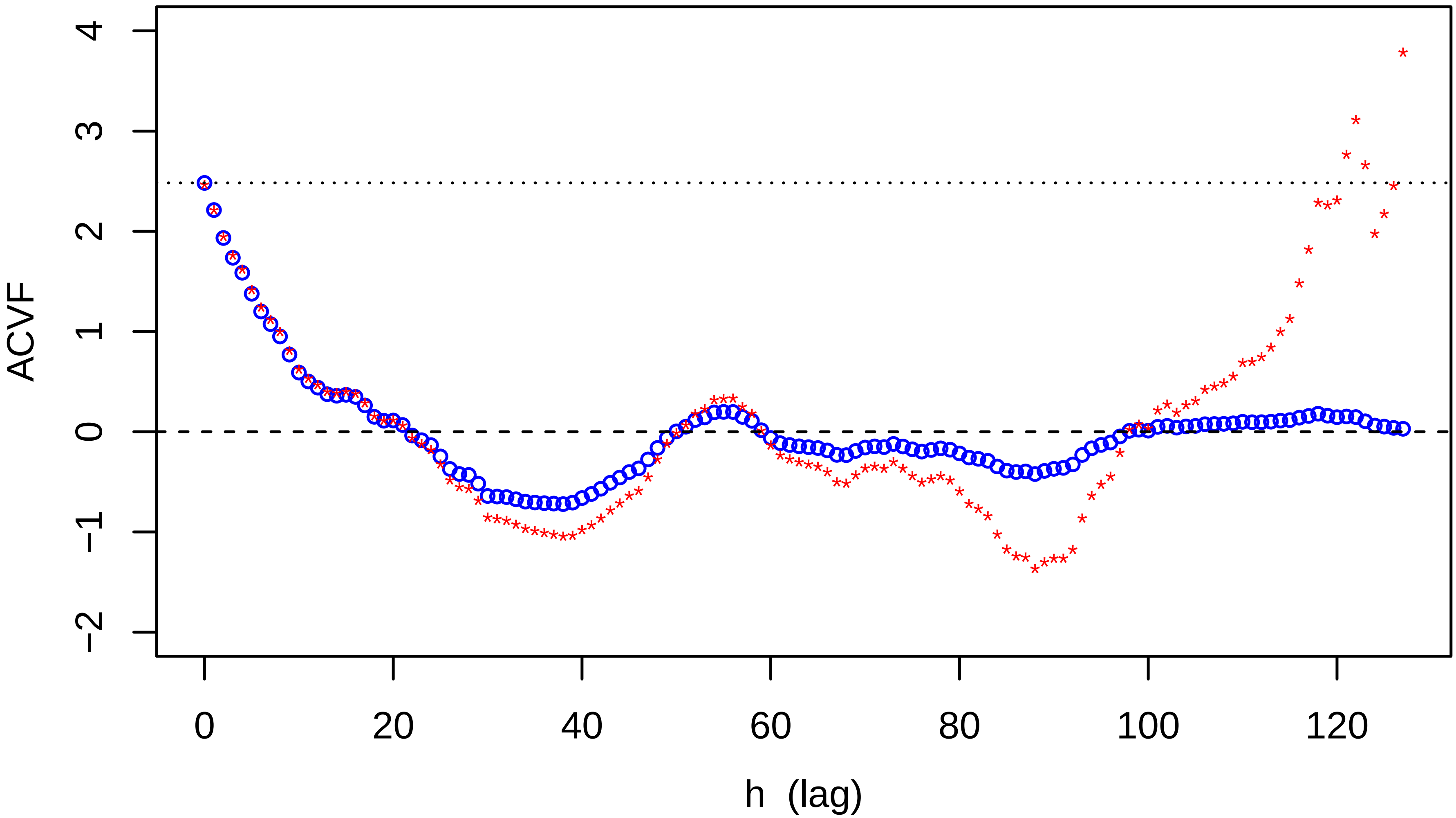
  i.e., $\tilde{\gamma}(h)$ is a unbiased estimator

- returning now to $\hat{\gamma}(h)$ (i.e., we use $\overline{X}$ rather than μ), could consider using $\frac{n}{n-|h|}\hat{\gamma}(h)$ in view of above result
- $\hat{\gamma}(h)$ & $\frac{n}{n-|h|}\hat{\gamma}(h)$ called, respectively, *biased* & *unbiased* ACVF estimators (even though latter is actually biased in general!)

Wind Speed Time Series $\{x_t\}$



Biased & Unbiased Sample ACVF for Wind Speed



Estimation of ACVF and ACF: IV

- as sample ACVF for wind speed time series demonstrates, unbiased ACVF estimate $\left\{\frac{n}{n-|h|}\hat{\gamma}(h)\right\}$ need *not* satisfy requirements needed to be a theoretical ACVF
- by contrast, biased ACVF estimate $\{\hat{\gamma}(h)\}$ *always* satisfies the requirements (in particular, it is always nonnegative definite)
  - McLeod and Jiménez (1984, 1985) present a clever proof based upon the ACVF for moving average processes
- for time series models in common use, biased estimator typically has smaller mean square error than unbiased estimator, which provides additional rationale for preferring $\hat{\gamma}(h)$

Estimation of ACVF and ACF: V

- for time series models we will be considering later on,

$$\hat{\boldsymbol{\rho}}_k = [\hat{\rho}(1), \ldots, \hat{\rho}(k)]'$$

is approximately $\mathcal{N}(\boldsymbol{\rho}_k, W/n)$ for large n (k fixed & k/n small), where

$$\boldsymbol{\rho}_k = [\rho(1), \ldots, \rho(k)]',$$

and the (i, j)th element of the $k \times k$ matrix W is given by *Bartlett's formula* (here $i = 1, \ldots, k$, and $j = 1, \ldots, k$ also):

$$w_{i,j} = \sum_{l=1}^{\infty} [\rho(l+i) + \rho(l-i) - 2\rho(i)\rho(l)] \times [\rho(l+j) + \rho(l-j) - 2\rho(j)\rho(l)]$$

- for IID noise, $w_{i,j} = 1$ if $i = j$ and $= 0$ otherwise, leading to previously stated result that $\hat{\rho}(1)$, ..., $\hat{\rho}(k)$ are approximately IID $\mathcal{N}(0, 1/n)$ RVs (see overhead II–66)

Example – Bartlett's Formula for MA(1) Process: I

- for MA(1) model $X_t = Z_t + \theta Z_{t-1}$ with $Z_t \sim \text{WN}(0, \sigma^2)$, have

$$w_{h,h} = \begin{cases} 1 - 3\rho^2(1) + 4\rho^4(1), & h = 1, \\ 1 + 2\rho^2(1), & h > 1, \end{cases}$$

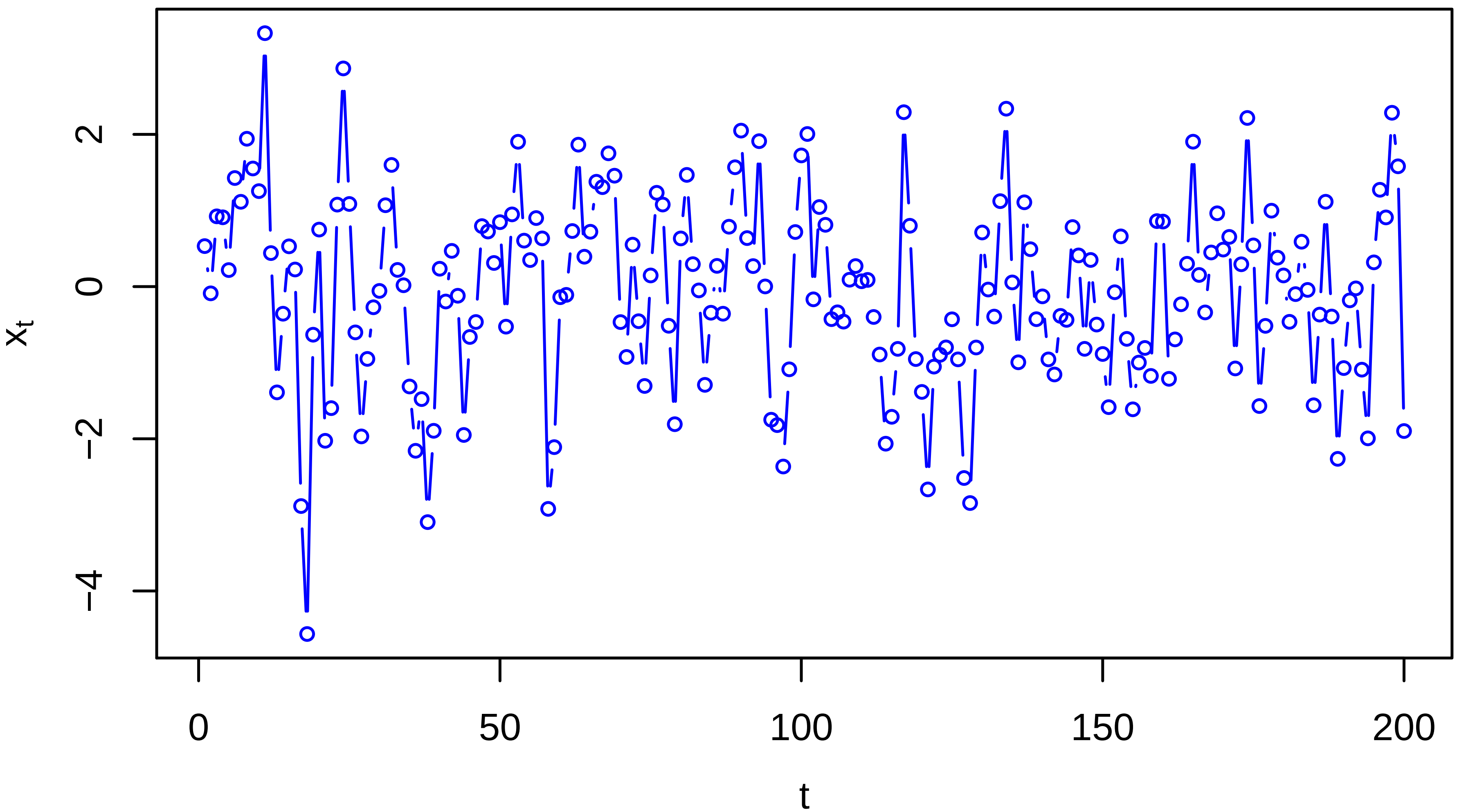
  so $\hat{\rho}(h)$ is approximately $\mathcal{N}(\rho(h), w_{h,h}/n)$ for large n (recall that $\rho(1) = \theta/(1+\theta^2)$ and $\rho(h) = 0$ for $h \geq 2$)

- consider $n = 200$ observations from simulated Gaussian MA(1) process with $\theta = 0.8$ and $\sigma^2 = 1$

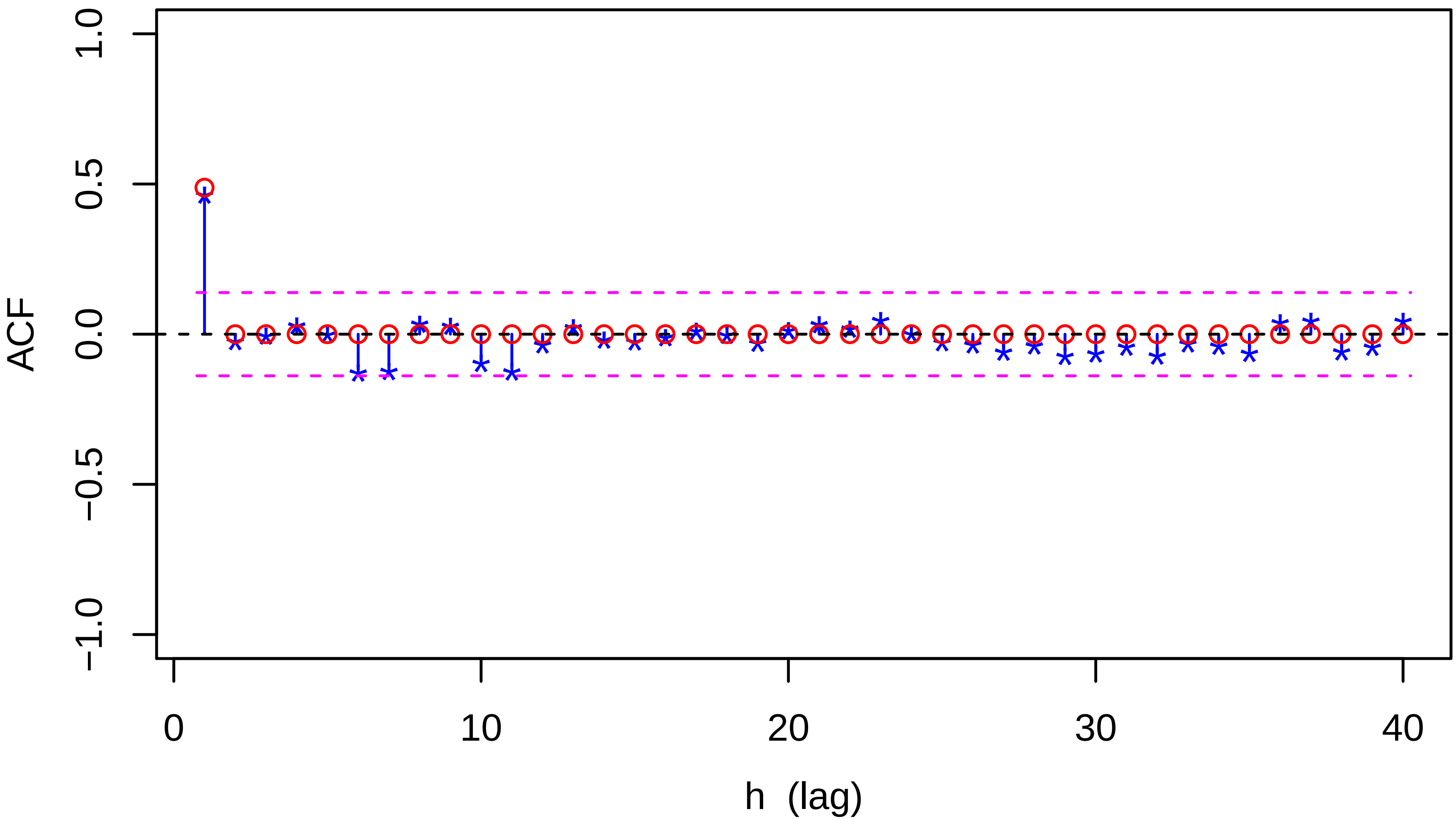
- true ACF is

$$\rho(h) = \begin{cases} 1, & h = 0, \\ 0.8/(1+0.8^2) \doteq 0.4878, & h = \pm 1, \\ 0, & \text{otherwise} \end{cases}$$

Simulated MA(1) Time Series



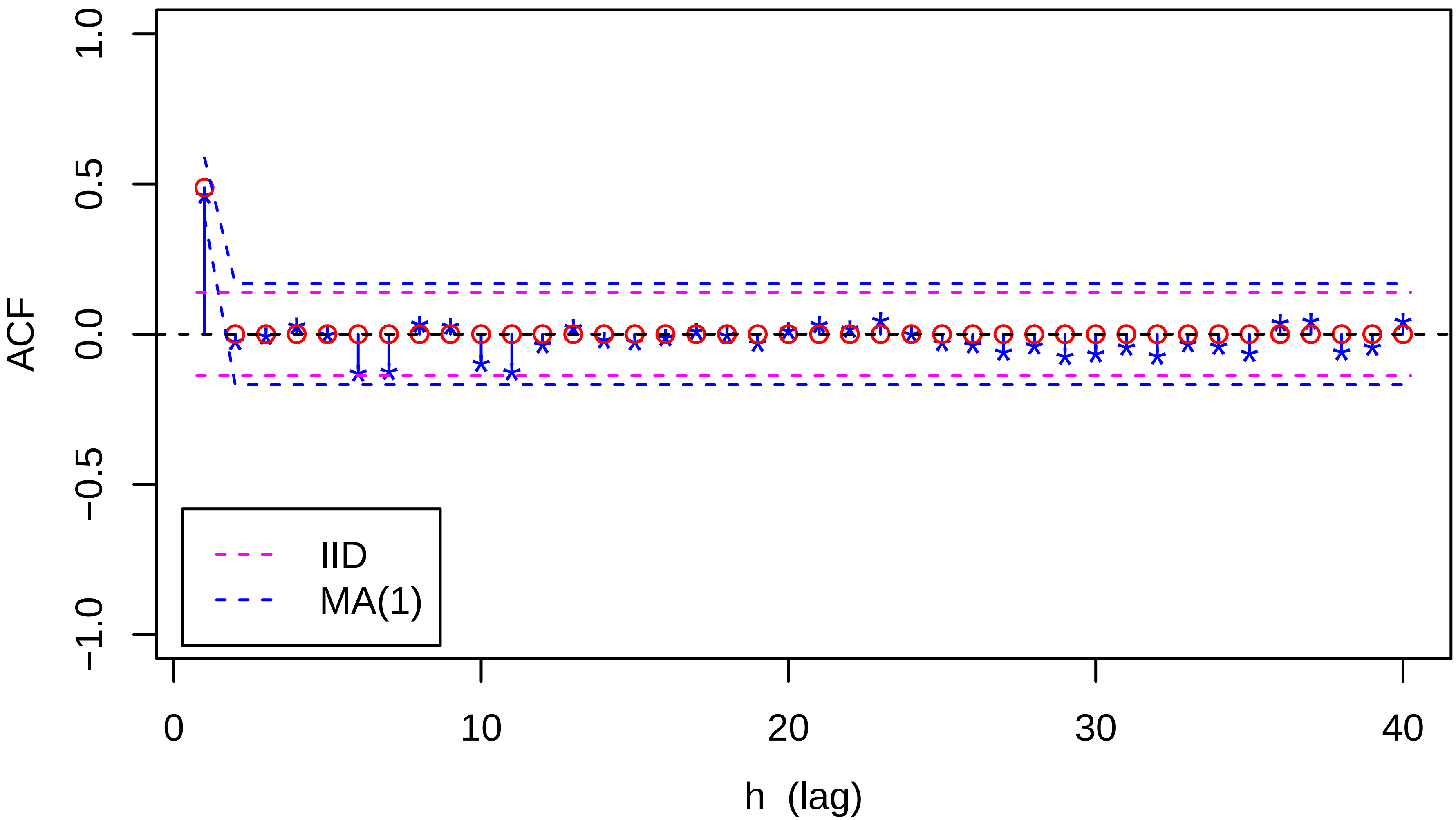
True & Sample ACFs & 95% IID Confidence Bounds



Examining Hypothesis of IID Noise

- suppose we did not know that observations came from MA(1) process and we entertain (incorrect) null hypothesis of IID noise
- theory then says to compare sample ACF to limits $\pm 1.96/\sqrt{n}$ and see if approximately 95% of values fall within these limits
- if we consider ACF at lags $1, \ldots, 40$ and note that 95% of 40 is 38, might expect to see at most a small number (e.g., 1, 2 or 3) of sample ACF values outside of limits
- from this viewpoint, plot on previous overhead seems to be consistent with IID hypothesis – just $\hat{\rho}(1)$ falls outside of $\pm 1.96/\sqrt{n}$ limits; however, p-value for $\hat{\rho}(1)$ is 2×10^{-11}, which is strong evidence *against* the (incorrect) IID hypothesis

True & Sample ACFs & 95% Confidence Bounds: I



Example – Bartlett's Formula for MA(1) Process: II

- 95% confidence bounds for IID noise are $\pm 1.96/\sqrt{n}$, which just depend just on sample size n
- MA(1)-based bounds shown on previous overhead are $\rho(h) \pm 1.96\sqrt{(w_{h,h}/n)}$, which depend on n and true ACF since

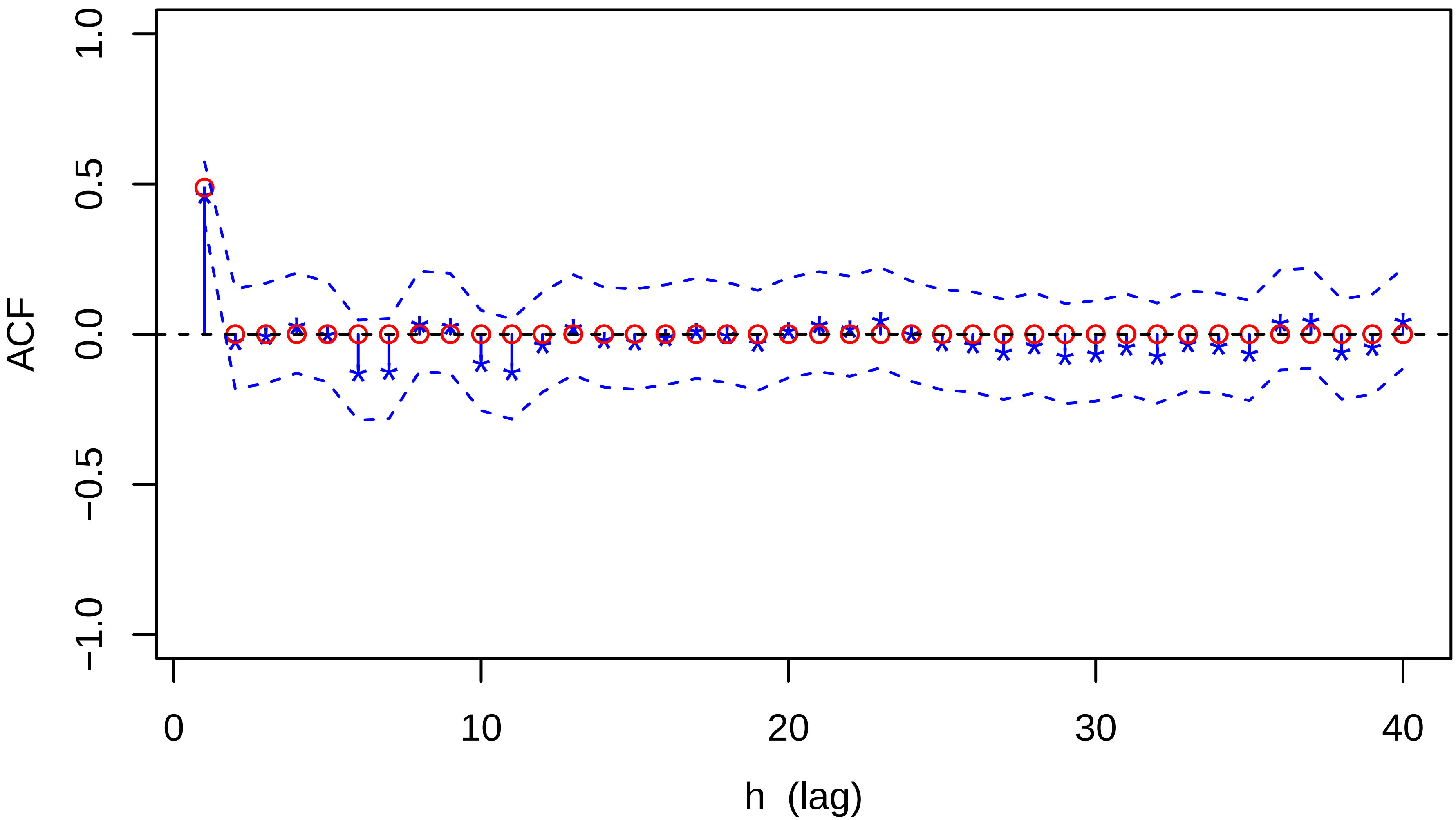
$$w_{h,h} = \begin{cases} 1 - 3\rho^2(1) + 4\rho^4(1), & h = 1, \\ 1 + 2\rho^2(1), & h > 1 \end{cases}$$

- in practical applications, true ACF unknown, so common procedure is to use $\hat{\rho}(h) \pm 1.96\sqrt{(\hat{w}_{h,h}/n)}$, where

$$\hat{w}_{h,h} = \begin{cases} 1 - 3\hat{\rho}^2(1) + 4\hat{\rho}^4(1), & h = 1, \\ 1 + 2\hat{\rho}^2(1), & h > 1 \end{cases}$$

- next overhead shows MA(1)-based 95% bounds using $\hat{\rho}(h)$ and $\hat{w}_{h,h}$ (as $n \to \infty$, these should get closer to $\rho(h)$ and $w_{h,h}$)

True & Sample ACFs & 95% Confidence Bounds: II



Example – Bartlett's Formula for MA(1) Process: III

- since ACF for MA(1) process $X_t = Z_t + \theta Z_{t-1}$ takes the form

$$\rho(h) = \begin{cases} \theta/(1+\theta^2), & h = 1, \\ 0, & h \geq 2, \end{cases}$$

  evidence against an MA(1) process would be either

  * $\hat{\rho}(1) \pm 1.96\sqrt{(\hat{w}_{1,1}/n)}$ trapping 0 or
  * $\hat{\rho}(h) \pm 1.96\sqrt{(\hat{w}_{h,h}/n)}$ for $h \geq 2$ failing to trap 0 for some h (a few near misses might be OK, but one $\hat{\rho}(h)$ with a high p-value would flag MA(1) as an untenable null hypothesis)

- no evidence against MA(1) in previous overhead (not surprising: time series is a realization of an MA(1) process!)

Example – Bartlett's Formula for AR(1) Process

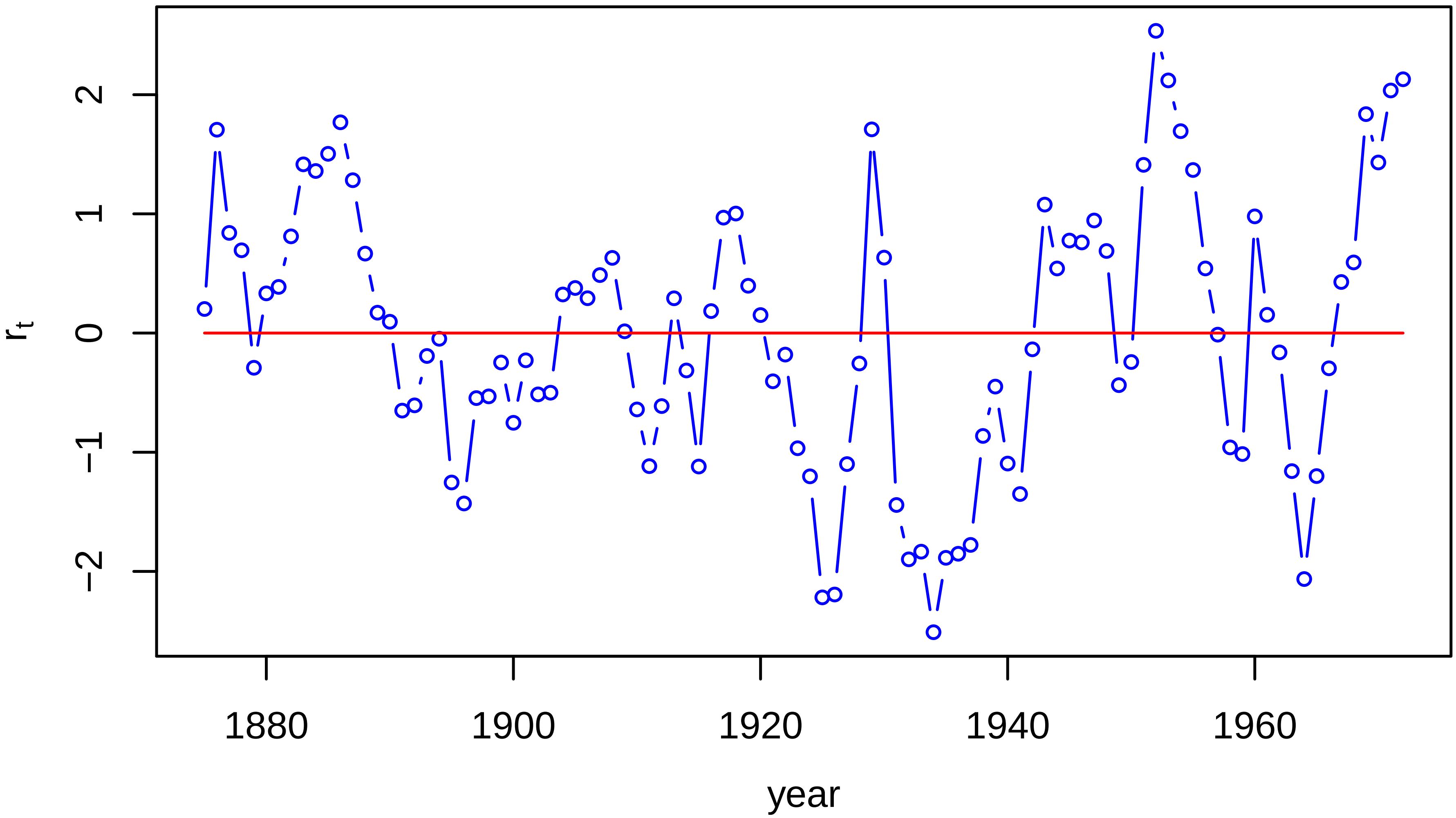
- for AR(1) model $X_t = \phi X_{t-1} + Z_t$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ and $|\phi| < 1$, have

$$w_{h,h} = \frac{(1 - \phi^{2h})(1 + \phi^2)}{1 - \phi^2} - 2h\phi^{2h}, \quad h = 1, 2, \ldots$$

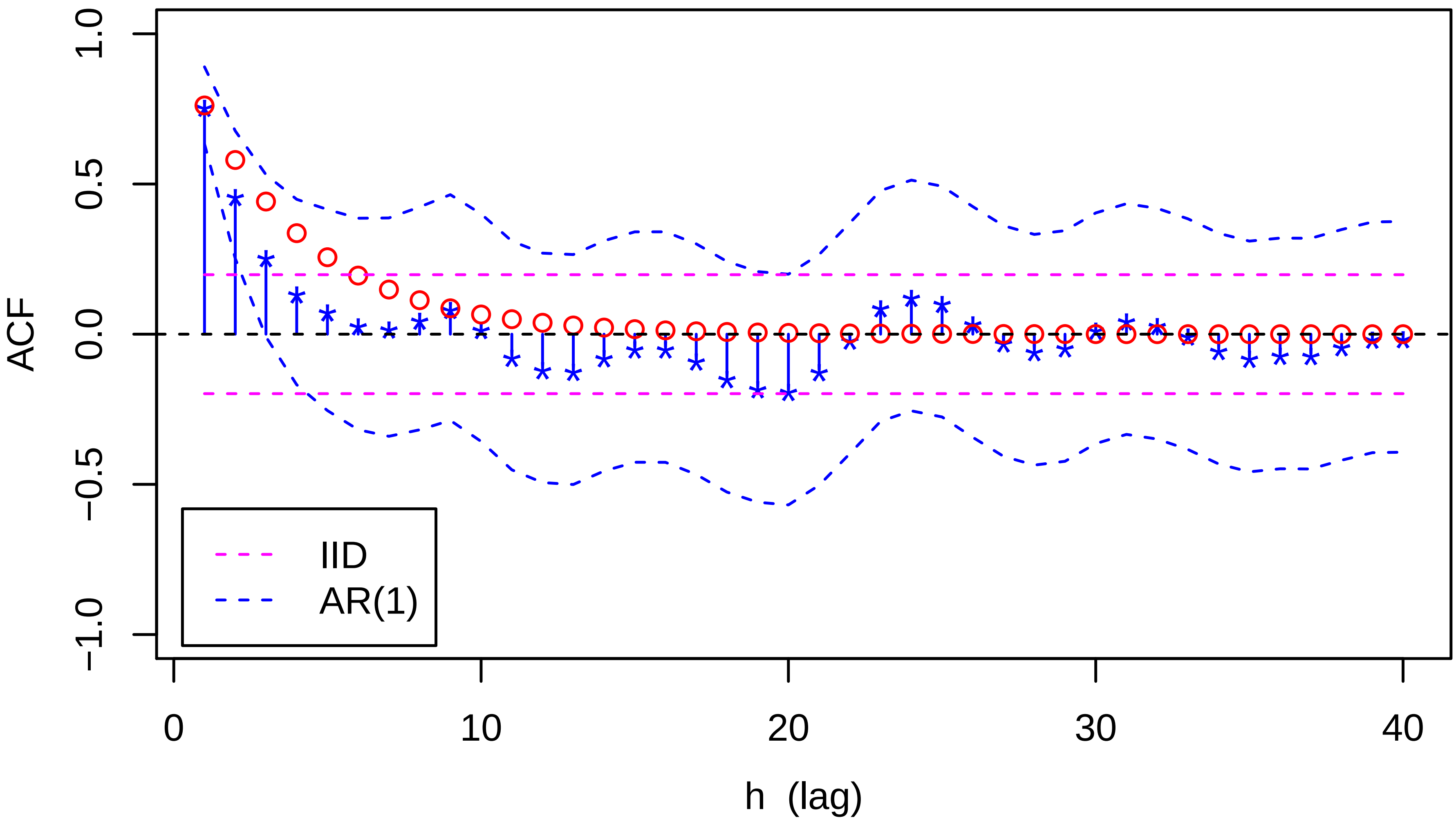
  so $\hat{\rho}(h)$ is approximately $\mathcal{N}(\rho(h), w_{h,h}/n)$ for large n (recall that $\rho(h) = \phi^h$ for $h \geq 1$)

- consider two examples
  - residuals $\{r_t\}$ from least squares line fit to Lake Huron levels
  - wind speed series $\{x_t\}$
- for both examples, will compare sample ACF with AR(1) model based on setting ϕ to $\hat{\rho}(1)$
- yields $\hat{\phi} \doteq 0.762$ for $\{r_t\}$ and $\hat{\phi} \doteq 0.891$ for $\{x_t\}$

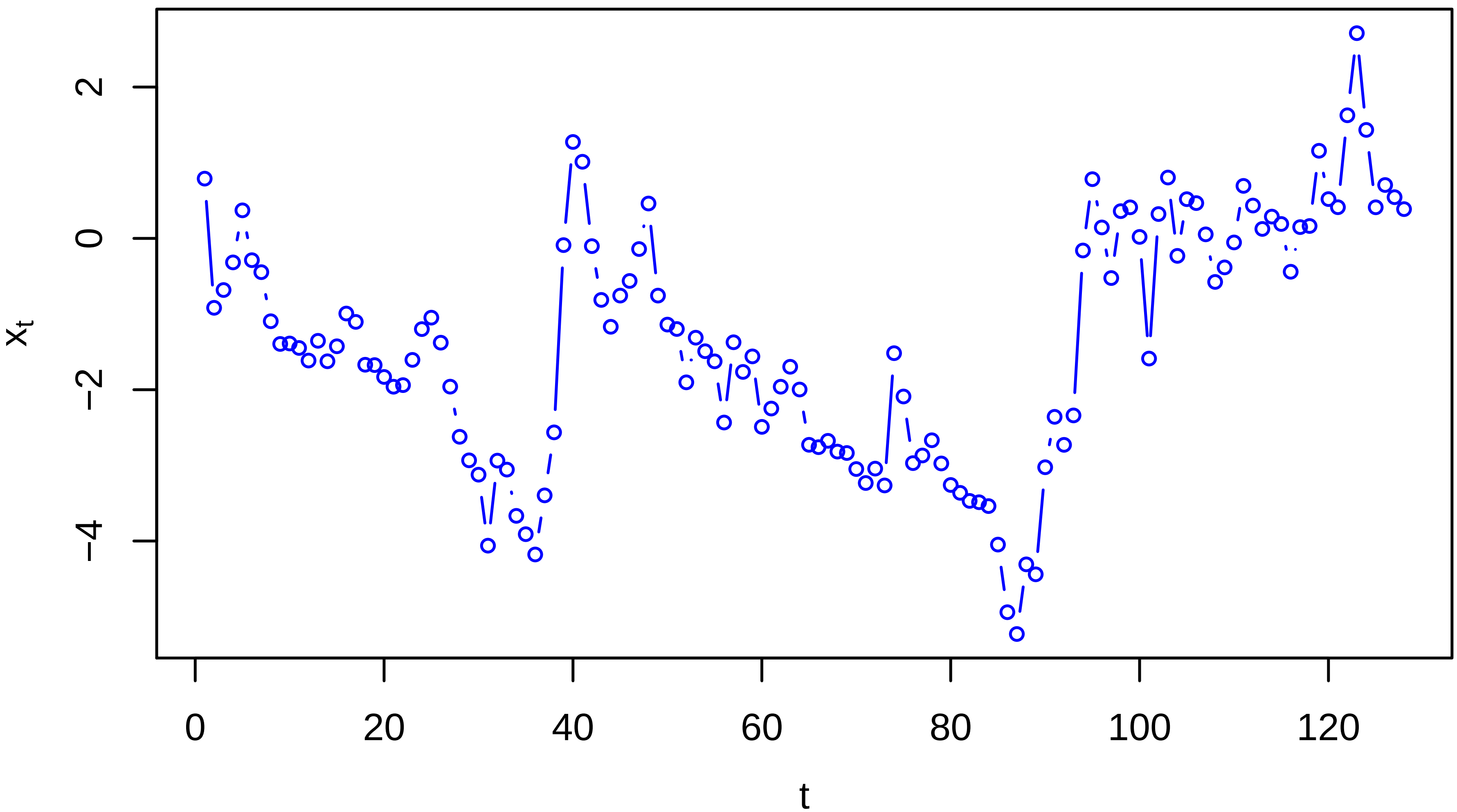
Residuals $r_t = x_t - \hat{c}_0 - \hat{c}_1 t$ from Least Squares Fit



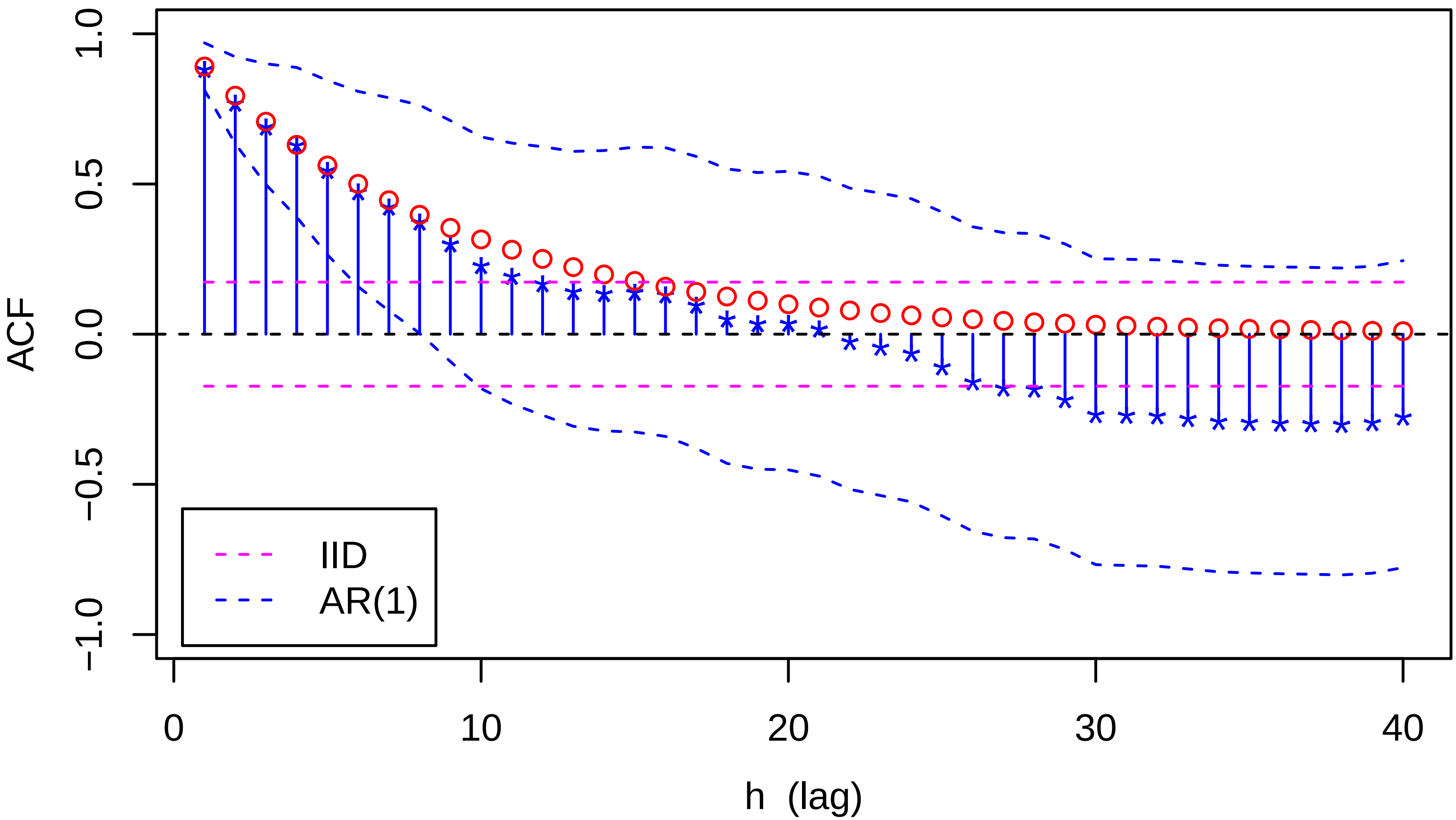
Model & Sample ACFs & 95% Confidence Bounds



Wind Speed Time Series $\{x_t\}$



Model & Sample ACFs & 95% Confidence Bounds



References

- A. I. McLeod and C. Jiménez (1984), 'Nonnegative Definiteness of the Sample Autocovariance Function,' *The American Statistician*, **38**, pp. 297–8
- A. I. McLeod and C. Jiménez (1985), 'Reply to Discussion by Arcese and Newton,' *The American Statistician*, **39**, pp. 237–8