Sample ACVFs / ACFs

Distribution & White Noise detection

STAT 464 / 864 | Fall 2024

Discrete Time Series Analysis

Skyepaphora Griffith, Queen's University

Sample ACVFs

Let $\{x_t\}$ be an **observed**, **N**-length, **stationary** time series.

- The standard **ACVF estimator** is $\hat{\gamma}(h) = \dfrac{1}{N} \sum\limits_{t=1}^{N-h} (x_t - \bar{x})(x_{t+h} - \bar{x})$

- Can only compute for $|h| < N$

- Don't trust this for lags **h** near **N** (or **-N**).

General Rule: Stick to lags less than $\log(\mathbf{N})$, or less than **N/4**
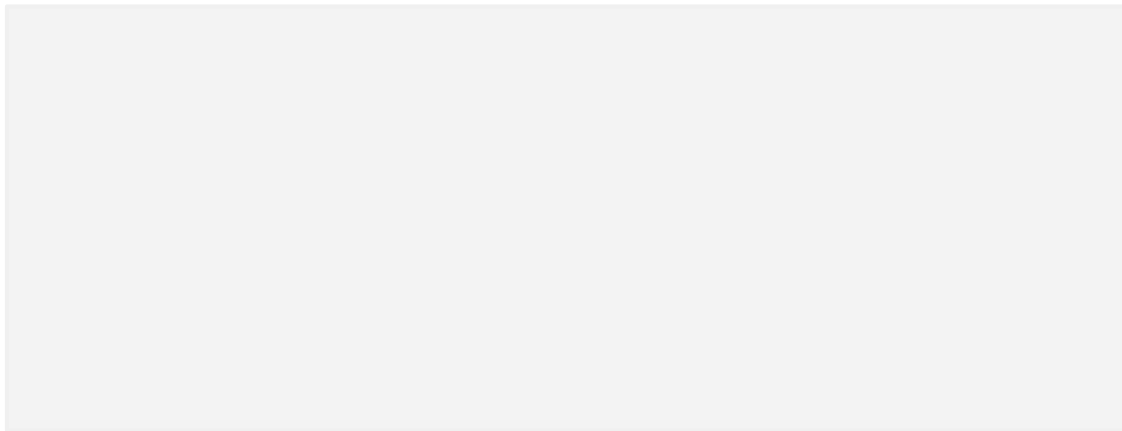
ACVFs and Seasonality

$$\hat{\gamma}(h) = \frac{1}{N} \sum_{t=1}^{N-h} (x_t - \bar{x})(x_{t+h} - \bar{x})$$

If $\{X_t\}$ has a cyclic component, $\hat{\gamma}_X$ should show a similar cycle.

Extreme example: $x(t) = \cos(2\pi t/d)$.   For $|h| \geq 0$,

$\hat{\gamma}(h + kd)$

will tend to be smaller than $\hat{\gamma}(h)$ as k increases

This is the original expression of $\boxed{\hat{\gamma}(h)}$, but with fewer terms

ACVFs and Seasonality:

$$\hat{\gamma}(h) = \frac{1}{N} \sum_{t=1}^{N-h} (x_t - \bar{x})(x_{t+h} - \bar{x})$$

Random component → $\hat{\gamma}(h)$ is approximate.

Too much random   → noise drowns out cyclic component

Suppose $\{X_t\}$ has the classical decomposition $(\star)$

Its **theoretical** ACVF is

X and Y have the same theoretical ACVF,
But when observed, the **sample** ACVFs can be very different!

Properties of the Sample ACVF

Also, wait, why do we divide by N?

1) Symmetry: $\gamma(h) = \gamma(-h)$	2) Still positive semidefinite

Let $y_t = x_t - \bar{x}$

Then define $A :=$

(i,j)th entry of AA^T is $\gamma_X(j - i)$

→ Divide by **N** $\quad \dfrac{1}{N}AA^T = \hat{\Gamma}_N$

So that's why we divide by **N** instead of **N-h**. Also these properties extend to $\hat{\rho}$.
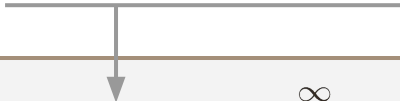
Distribution of $\hat{\rho}$

Let **m** be some maximum time-lag.

Let $\hat{\rho} = \big(\hat{\rho}(1), \dots, \hat{\rho}(m)\big)$, where $\hat{\rho}(h)$ is the sample ACF at lag **h**

Distribution (large **N**): $\quad \hat{\rho} \sim \mathcal{MVN}(\rho, \frac{1}{N}W)$

Covariance Matrix: $\qquad\qquad\qquad$ Mean vector: $\quad \rho = \big(\rho(1), \dots, \rho(m)\big)$

$$W[i,j] = \sum_{k=1}^{\infty} \big[\rho(k+i) + \rho(k-i) + \rho(i)\rho(k)\big] \times \big[\rho(k+j) + \rho(k-j) + \rho(j)\rho(k)\big]$$

Example W | White Noise Detection

Consider $X_t \sim \text{wn}(0, \sigma^2)$.

Then $\hat{\rho}(1), \ldots, \hat{\rho}(m)$ are approximately independent $\mathcal{N}(0, 1/N)$ RVs

To test if $\{x_t\}$ is white noise, we can test $\hat{\rho}$

For a given (nonzero) lag h,

$$\text{Null Hypothesis:} \quad H_0 : \rho(h) = 0$$

$$\text{Alternative Hypothesis:} \quad H_A : \rho(h) \neq 0$$

Example W | White Noise Detection

🕐 Under the **wn($0, \sigma^2$)** hypothesis, $\hat{\rho}(i) \overset{\cdot}{\sim} N(0, 1/N)$

🕐 Reject H$_A$ if $|\hat{\rho}(h)|$ is too large, say, **> c**

🕐 Probability of type 1 error (under the null):

$$H_0 : \rho(h) = 0$$
$$H_A : \rho(h) \neq 0$$

where $Z \sim \mathcal{N}(0, 1)$

Example W | White Noise Detection

🕐 Now we want a **threshold** for P(type I error): some $\alpha \in (0,1)$

Typical choices: [α = 0.05] or [α = 0.01]

🕐 Set $\dfrac{c}{1/\sqrt{N}}$ equal to $\left(1 - \dfrac{\alpha}{2}\right)$ quantile of the **N(0,1)** distribution

For $\alpha = 0.05$, we want $\dfrac{c}{1/\sqrt{N}}$

to = the 0.975-quantile (**≈1.96**)

$\rightarrow \quad c = \dfrac{1.96}{\sqrt{N}}$
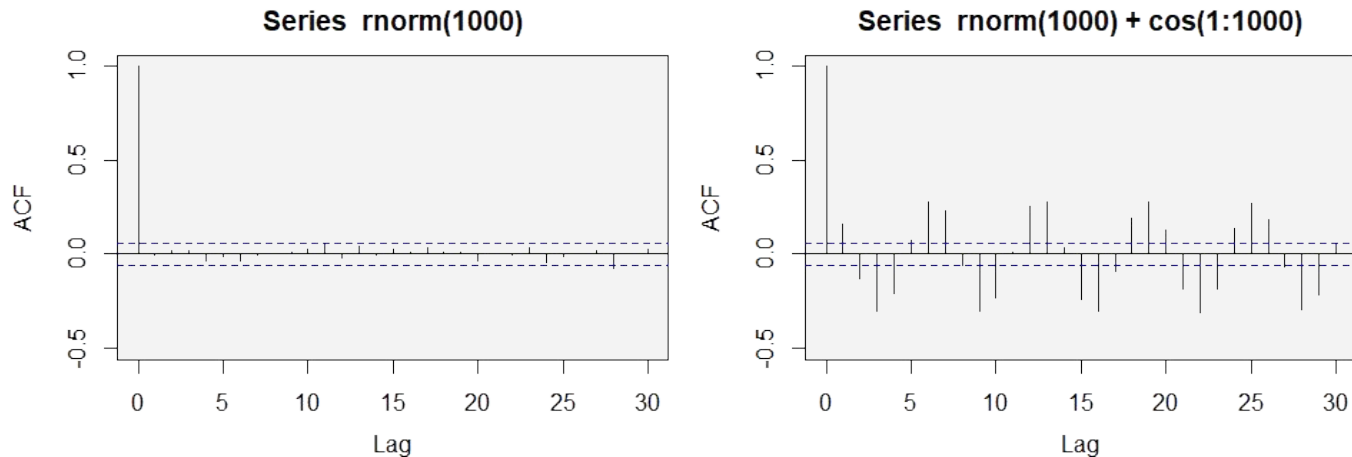
Example W | White Noise Detection

Conclusion:

🕐 To test if {X$_t$} is **wn(0, σ^2)**, at the $\alpha = 0.05$ level,

compute $\hat{\rho}(h)$ for some set of lags {h = **0, 1, ..., m**}

🕐 $|\hat{\rho}(h)| > 1.96/\sqrt{N}$ → Reject the null hypothesis {X$_t$} ~ **wn(0, σ^2)**
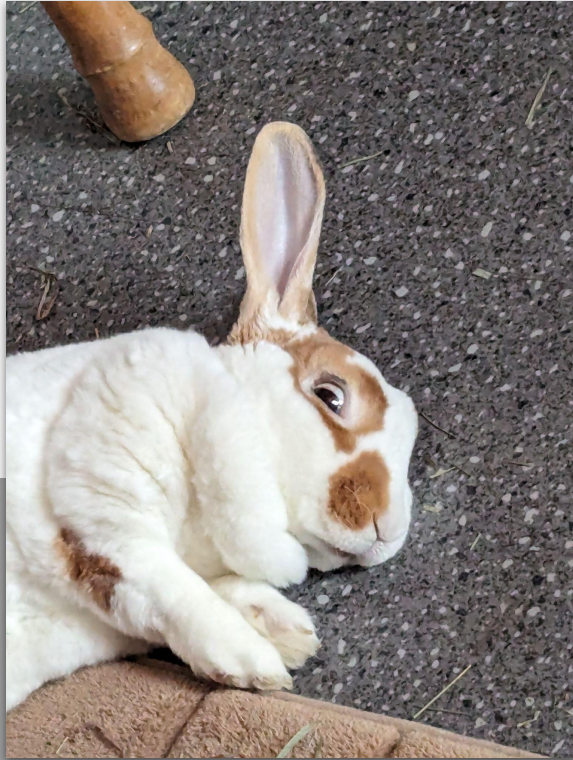
ACFs in R

🕐 Function `acf()` plots 2 horizontal dashed blue lines at $\pm 1.96/\sqrt{N}$

🕐 $\hat{\rho}(h)$ that fall outside blue lines = evidence $\{X_t\}$ is not $\mathbf{wn(0, \sigma^2)}$

Interpretation: If we compute $\hat{\rho}$ for white noise at **m** different lags,

We expect $\hat{\rho}$ to fall outside the blue lines M x (alpha) times

We learned something today, in Time Series 🕛

What do we tell quin??