Simple Time Series Models: I

- definition: time series model for \( \{x_t\} \) is specification of joint distributions (or summaries thereof) of a sequence of random variables (RVs) \( \{X_t\} \), one of whose realizations is assumed to be \( \{x_t\} \)
  - sequence of RVs \( \{X_t\} \) called a stochastic process
  - realizations of \( \{X_t\} \) are coupled together as dictated by their joint distribution

- will use term ‘time series’ to refer to both
  - \( \{x_t\} \) (actual data or realization of stochastic process) and
  - \( \{X_t\} \) (stochastic process itself)
Simple Time Series Models: II

- complete model for \( \{x_t : t \in T\} \), \( T = \{1, 2, \ldots, n\} \), requires specification of joint distributions of RVs \( X_1, X_2, \ldots, X_n \), which is equivalent to specification of all probabilities

\[
P[X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n],
\]
where \(-\infty < a_t < \infty\) for all \( t \in T \)

- while fully general, too complicated, so will concentrate on summary of joint distributions afforded by first- and second-order moments, i.e.,

\[
E\{X_t\} \text{ and } E\{X_{t+h}X_t\}, \quad 1 \leq t + h \leq n
\]
or, equivalently, \( E\{X_t\} \) and \( \text{cov} \{X_{t+h}, X_t\} \), where

\[
\text{cov} \{U, V\} \equiv E\{(U - E\{U\})(V - E\{V\})\} = E\{UV\} - E\{U\}E\{V\}
\]
Simplest Time Series Model – IID Noise: I

• assume that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (IID) RVs so that

$$P[X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n] = P[X_1 \leq a_1] \times P[X_2 \leq a_2] \times \cdots \times P[X_n \leq a_n] = F(a_1) \times F(a_2) \times \cdots \times F(a_n),$$

where $F(\cdot)$ is cumulative probability distribution function of each of the IID RVs

• model says that there is no dependence between observations

• extend model to include future observations by requiring RVs $X_1, X_2, \ldots, X_n, X_{n+1}, \ldots, X_{n+h}$ to be IID for any $h \geq 1$
for extended model, conditioning on observed time series does not alter probabilistic description of $X_{n+h}$ since

$$P[X_{n+h} \leq a \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] = P[X_{n+h} \leq a]$$

implies that function $f$ minimizing the mean squared error

$$E\{[X_{n+h} - f(X_1, X_2, \ldots, X_n)]^2\}$$

is $f(X_1, X_2, \ldots, X_n) = \mu$, where $\mu$ is the common mean of all the $X_t$'s; i.e., best predictor of $X_{n+h}$ in the mean square sense is $\mu = E\{X_{n+h}\}$, which does not depend on $X_1, X_2, \ldots, X_n$

here are realizations of IID noise with $n = 100$ and with $F(\cdot)$ given by standard normal (Gaussian) distribution (i.e., zero mean and unit variance)
Gaussian IID Noise
Gaussian IID Noise
Gaussian IID Noise
Binary-valued IID Noise

- as second example of IID noise, suppose $X_t$ is binary-valued:
  $$P[X_t = 1] = p \quad \text{and} \quad P[X_t = -1] = 1 - p, \quad 0 \leq p \leq 1,$$
  for which $\mu = E\{X_t\} = p \times 1 + (1 - p) \times (-1) = 2p - 1$

- reconsider first all-star baseball game in each year, with 1 and $-1$ indicating, respectively, AL and NL victories (regard 2002 tie as a gap, along with actual gap in 1945)

- as of Jan 2016, have been 39 AL victories and 42 NL victories, so entertain hypothesis $p = \frac{39}{81} \doteq 0.481$

- compare actual time series with realizations of binary-valued IID noise with $p \doteq 0.481$
All-Star Baseball Games (First in Each Year)

1940 1960 1980 2000
NL win tie AL win

outcome of game

year

1940 1960 1980 2000
Binary-valued IID Noise ($p = 0.481$)
Binary-valued IID Noise ($p \hat{=} 0.481$)
Binary-valued IID Noise \((p \doteq 0.481)\)
Random Walk Process: I

• can use IID noise as building block for other processes of interest
• suppose that \( \{X_t\} \) is IID noise such that \( E\{X_t\} = 0 \)
• for \( t \geq 1 \), construct new process

\[
S_t = \sum_{u=1}^{t} X_u;
\]

thus

\[
S_1 = X_1 = S_0 + X_1 \quad \text{(if we define } S_0 \text{ to be 0)}
\]
\[
S_2 = X_1 + X_2 = S_1 + X_2
\]
\[
S_3 = X_1 + X_2 + X_3 = S_2 + X_3
\]
\[\vdots\]
\[
S_t = X_1 + X_2 + \cdots + X_{t-1} + X_t = S_{t-1} + X_t
\]
Random Walk Process: II

• for $t \geq 1$, can recover $X_t$ from first backward difference of $S_t$:

$$S_t - S_{t-1} = X_t$$

since

$$S_t = X_1 + X_2 + \cdots + X_{t-2} + X_{t-1} + X_t$$
$$S_{t-1} = X_1 + X_2 + \cdots + X_{t-2} + X_{t-1}$$

(works for $t = 1$ also since $S_0 = 0$ by definition)

• $\{S_t\}$ is called a zero-mean random walk process

• consider realizations of two examples of random walks, for which
  - $\{X_t\}$ is a standard Gaussian IID process
  - $\{X_t\}$ is a binary-valued IID process (1 or $-1$ with $p = 0.5$; Brockwell & Davis call this a simple symmetric random walk)
Gaussian Random Walk Process

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{gaussian_random_walk.png}
\caption{Gaussian Random Walk Process}
\end{figure}
Gaussian Random Walk Process

\[
\begin{align*}
\text{Figure: Gaussian Random Walk Process.}
\end{align*}
\]
Gaussian Random Walk Process
Simple Symmetric Random Walk Process
Simple Symmetric Random Walk Process
Simple Symmetric Random Walk Process
Stationary Models: I

- let \( \{X_t : t \in \mathbb{Z}\} \) be a time series (stochastic process) such that \( E\{X_t^2\} < \infty \) for all \( t \), where \( \mathbb{Z} \equiv \{\ldots, -1, 0, 1, \ldots\} \)

- mean function for \( \{X_t\} \) is defined to be
  \[ \mu_X(t) = E\{X_t\} \]

- covariance function for \( \{X_t\} \) is defined to be
  \[ \gamma_X(r, s) = \text{cov}\{X_r, X_s\} = E\{[X_r - \mu_X(r)][X_s - \mu_X(s)]\} \]
  for all integers \( r \) and \( s \)

- \( \{X_t\} \) is said to be (weakly) stationary if
  1. \( \mu_X(t) \) is independent of \( t \)
  2. \( \gamma_X(r, s) \) depends on just \( r - s \), i.e., the lag (spacing in time) between \( X_r \) & \( X_s \), but not on \( r \) & \( s \) (thus \( \gamma_X(t + h, t) \) is independent of \( t \) for each lag \( h \))
Stationary Models: II

- stationary says that first- and second-order properties of
  \[ X_t, X_{t+1}, \ldots, X_{t+m} \] \hspace{1cm} (*)
  are the same as those for
  \[ X_{t+h}, X_{t+h+1}, \ldots, X_{t+h+m} \] \hspace{1cm} (**) 
  for all \( t \in \mathbb{Z} \), all \( m \geq 0 \) and all \( h \in \mathbb{Z} \)

- weak stationarity is also called second-order stationarity, co-variance stationarity and wide-sense stationarity

- another notion is strict stationarity, which says that the RVs in (*) have the same joint distribution as the RVs in (**) 

- if \( \{X_t\} \) is strictly stationary and \( E\{X_t^2\} < \infty \), then \( \{X_t\} \) is also weakly stationary (converse not true in general)

- henceforth will take ‘stationary’ to mean ‘weakly stationary’
Autocovariance and Autocorrelation Functions

- since $\gamma_X(t + h, t)$ depends on $h$ but not $t$, covariance function for a stationary process can be taken to be a function of one variable (the lag $h$)
- accordingly define the autocovariance function $\gamma_X(\cdot)$ (ACVF) via
  \[ \gamma_X(h) = \gamma_X(t + h, t) = \text{cov} \{ X_{t+h}, X_t \} \]
- note: $\gamma_X(0) = \text{cov} \{ X_t, X_t \} = \text{var} \{ X_t \}$ (the process variance)
- when $\gamma_X(0) > 0$, define the corresponding autocorrelation function $\rho_X(\cdot)$ (ACF) via
  \[ \rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{\text{cov} \{ X_{t+h}, X_t \}}{\text{var} \{ X_t \}} = \frac{\text{cov} \{ X_{t+h}, X_t \}}{\sqrt{\text{var} \{ X_{t+h} \} \text{var} \{ X_t \}}} \]
Example – IID Noise: I

• if \( \{X_t\} \) is IID noise such that \( E\{X_t^2\} < \infty \), then \( E\{X_t\} = \mu \) (a constant independent of \( t \)) and \( \text{cov}\{X_{t+h}, X_t\} = 0 \) for all \( h \neq 0 \)

• hence IID noise with a finite variance, say \( \sigma^2 \), is a stationary process with

\[
\gamma_X(h) = \begin{cases} 
\sigma^2, & h = 0, \\
0, & \text{otherwise}, 
\end{cases}
\]

and, when \( \sigma^2 > 0 \), ACF

\[
\rho_X(h) = \begin{cases} 
1, & h = 0, \\
0, & \text{otherwise} 
\end{cases}
\]

• use notation \( \{X_t\} \sim \text{IID}(\mu, \sigma^2) \) to denote IID noise
ACF for IID Noise

ACF

h (lag)
Example – IID Noise: II

• following four plots show examples of IID(0,1) noise with these distributions:

1. standard normal (Gaussian)
2. uniform over interval $[-\sqrt{3}, \sqrt{3}]$
3. double exponential with probability density function (PDF) given by $f(x) = \exp(-|x|\sqrt{2})/\sqrt{2}$
4. discrete distribution that assumes values $-5, 0$ and $5$ with probabilities $0.02, 0.96$ and $0.02$, respectively
IID(0,1) Noise from Gaussian Distribution
IID(0,1) Noise from Uniform Distribution
IID(0,1) Noise from Double Exponential Distribution
IID(0,1) Noise from Discrete Distribution
Example – White Noise: I

• by definition \( \{X_t\} \) is a white noise process if its RVs are uncorrelated (i.e., \( \text{cov}\{X_r, X_s\} = 0 \) as long as \( r \neq s \)) and have the same mean \( \mu \) and the same variance \( \sigma^2 \) (assumed to be finite)

• \( \{X_t\} \sim \text{WN}(\mu, \sigma^2) \) denotes a white noise process

• a \( \text{WN}(\mu, \sigma^2) \) process is a stationary process with ACVF

\[
\gamma_X(h) = \begin{cases} 
\sigma^2, & h = 0, \\
0, & \text{otherwise},
\end{cases}
\]

and, when \( \sigma^2 > 0 \), ACF

\[
\rho_X(h) = \begin{cases} 
1, & h = 0, \\
0, & \text{otherwise}
\end{cases}
\]

(same ACVF and ACF as for IID noise)
Example – White Noise: II

- every IID($\mu, \sigma^2$) process is also a WN($\mu, \sigma^2$) process (hence examples of Gaussian, uniform, double exponential and discrete IID(0,1) noise also serve as examples of WN(0,1) processes)

- converse is not true, as the following two examples show:
  1. mishmash noise: at each time $t$, choose at random from amongst Gaussian, uniform, double exponential and discrete RVs (recall that each distribution has zero mean and unit variance)
  2. blocky noise: paste together blocks from Gaussian, uniform, double exponential and discrete IID(0,1) time series
Blocky White Noise
Example – Random Walk: I

• suppose that \( \{X_t\} \) is IID(0,\( \sigma^2 \)) noise, and construct corresponding random walk process:

\[
S_t = \sum_{u=1}^{t} X_u, \quad t \geq 1
\]

• since

\[
E\{S_t\} = \sum_{u=1}^{t} E\{X_u\} = 0, \quad t \geq 1
\]

first of two conditions for stationarity holds
Example – Random Walk: II

• in order for second condition to hold, the variance of \( \{ S_t \} \) cannot depend on \( t \); however, due to independence,

\[
\text{var} \{ S_t \} = \text{var} \left\{ \sum_{u=1}^{t} X_u \right\} = \sum_{u=1}^{t} \text{var} \{ X_u \} = t \sigma^2,
\]

which \textit{does} depend on \( t \) unless \( \sigma^2 = 0 \)

• conclusion: \( \{ S_t \} \) need \textit{not} be a stationary process, but it is sometimes called intrinsically stationary of unit order because its first-order backward differences \( S_t - S_{t-1} = X_t \) form a stationary process
Example – Random Walk: III

• recall the linearity property of covariances:

\[
\text{cov} \{aX + bY + c, Z\} = a \text{cov} \{X, Z\} + b \text{cov} \{Y, Z\},
\]

where \(X, Y\) & \(Z\) are RVs such that \(E\{X^2\}, E\{Y^2\} \text{ & } E\{Z^2\}\) are all finite, and \(a, b \text{ & } c\) are arbitrary real-valued constants

• covariance function for \(\{S_t\}\) is such that, for \(t \geq 1\) and \(h \geq 1\),

\[
\gamma_S(t + h, t) = \text{cov} \{S_{t+h}, S_t\}
\]

\[
= \text{cov} \{S_t + X_{t+1} + \cdots + X_{t+h}, S_t\}
\]

\[
= \text{cov} \{S_t, S_t\} + \sum_{l=1}^{h} \text{cov} \{X_{t+l}, S_t\}
\]

\[
= t\sigma^2 + \sum_{l=1}^{h} \text{cov} \{X_{t+l}, X_1 + \cdots + X_t\} = t\sigma^2
\]
Example – First-Order Moving Average Process: I

• suppose that \( \{Z_t : t \in \mathbb{Z}\} \sim \text{WN}(0, \sigma^2) \), and define
  \[
  X_t = Z_t + \theta Z_{t-1}, \quad t \in \mathbb{Z},
  \]
  where \( \theta \) is a real-valued constant

• since
  \[
  E\{X_t\} = E\{Z_t\} + \theta E\{Z_{t-1}\} = 0,
  \]
  first of two conditions for stationarity holds

• to establish stationarity, we need to show the second condition holds, namely, that \( \gamma_X(t + h, t) \) does not depend on \( t \)
Example – First-Order Moving Average Process: II

- now

\[ \gamma_X(t + h, t) = \text{cov} \{ Z_{t+h} + \theta Z_{t+h-1}, Z_t + \theta Z_{t-1} \} \]

\[ = \text{cov} \{ Z_{t+h}, Z_t \} + \theta \text{cov} \{ Z_{t+h}, Z_{t-1} \} + \theta \text{cov} \{ Z_{t+h-1}, Z_t \} + \theta^2 \text{cov} \{ Z_{t+h-1}, Z_{t-1} \} \]

- when \( h = 0 \), the 1st and 4th cov’s are equal to \( \sigma^2 \), while the 2nd and 3rd are zero, yielding \( \gamma_X(t, t) = \text{var} \{ X_t \} = \sigma^2(1+\theta^2) \)

- when \( h = 1 \), the 3rd cov is equal to \( \sigma^2 \), while the rest are all zero, yielding \( \gamma_X(t + 1, t) = \theta \sigma^2 \)

- when \( h = -1 \), the 2nd cov is equal to \( \sigma^2 \), while the rest are all zero, yielding \( \gamma_X(t - 1, t) = \theta \sigma^2 \)

- when \( |h| \geq 2 \), all four cov’s are zero, yielding \( \gamma_X(t + h, t) = 0 \) for \( h \neq 0, 1 \) or \(-1\)
Example – First-Order Moving Average Process: III

- since $\gamma_X(t+h, t)$ is independent of $t$ for all $h$, the process $\{X_t\}$ is stationary with ACVF

\[
\gamma_X(h) = \begin{cases} 
\sigma^2(1 + \theta^2), & h = 0, \\
\sigma^2\theta, & h = \pm1, \\
0, & \text{otherwise},
\end{cases}
\]

and, when $\sigma^2 > 0$, ACF

\[
\rho_X(h) = \begin{cases} 
1, & h = 0, \\
\theta/(1 + \theta^2), & h = \pm1, \\
0, & \text{otherwise}
\end{cases}
\]

- $\{X_t\}$ is called a first-order moving average or MA(1) process
ACF for MA(1) Process with $\theta = 1$
ACF for MA(1) Process with $\theta = 1/2$
ACF for MA(1) Process with $\theta = 0$
ACF for MA(1) Process with $\theta = -1/2$
ACF for MA(1) Process with $\theta = -1$
Example – First-Order Moving Average Process: IV

- as examples, generate realizations of MA(1) process with $\theta = 1$ and $\theta = -1$ using WN(0,1) processes with Gaussian, uniform, double exponential and discrete distributions
$\theta = 1$ MA(1) $x_t$ from Gaussian WN(0,1)
$\theta = -1 \text{ MA}(1) \ x_t$ from Gaussian WN(0,1)
\[ \theta = 1 \text{ MA}(1) x_t \text{ from Uniform WN}(0,1) \]
$\theta = -1$ MA(1) $x_t$ from Uniform WN(0,1)
$\theta = 1$ MA(1) $x_t$ from Double Exponential WN(0,1)
$\theta = -1$ MA(1) $x_t$ from Double Exponential WN(0,1)
\[ \theta = 1 \text{ MA}(1) x_t \text{ from Discrete WN}(0,1) \]
$$\theta = -1 \text{ MA}(1)$$

$$x_t$$ from Discrete WN(0,1)
Example – First-Order Autoregressive Process: I

- assume that \( \{X_t\} \) is a stationary process satisfying

\[
X_t = \mu + \phi(X_{t-1} - \mu) + Z_t, \quad t \in \mathbb{Z},
\]

where \( \mu \) and \( \phi \) are real-valued constants with \( |\phi| < 1 \), and \( \{Z_t\} \sim \text{WN}(0, \sigma^2) \) with \( \text{cov} \{X_s, Z_t\} = 0 \) for all \( s < t \)

- since

\[
E\{X_t\} = \mu + \phi(E\{X_{t-1}\} - \mu) + E\{Z_t\}
\]

implies

\[
E\{X_t\} - \mu = \phi(E\{X_t\} - \mu)
\]

we can conclude that \( \mu = E\{X_t\} \) if \( \phi \neq 0 \); if in fact \( \phi = 0 \), then \( \mu = E\{X_t\} \) follows immediately from

\[
X_t = \mu + Z_t
\]
Example – First-Order Autoregressive Process: II

- to find the ACVF, subtract $\mu$ from both sides of 
  \[ X_t = \mu + \phi(X_{t-1} - \mu) + Z_t, \]

  multiply each side by $X_{t-h} - \mu$ for $h > 0$ and take expectations

  \[ E\{(X_t - \mu)(X_{t-h} - \mu)\} = \phi E\{(X_{t-1} - \mu)(X_{t-h} - \mu)\} + E\{(X_{t-h} - \mu)Z_t\} \]

  to get

  \[ \text{cov}\ \{X_t, X_{t-h}\} = \phi \text{cov}\ \{X_{t-1}, X_{t-h}\} + \text{cov}\ \{X_{t-h}, Z_t\}; \]

  however, since $\text{cov}\ \{X_{t-h}, Z_t\} = 0$ by assumption, we have

  \[ \text{cov}\ \{X_t, X_{t-h}\} = \phi \text{cov}\ \{X_{t-1}, X_{t-h}\}, \]

  i.e.,

  \[ \gamma_X(h) = \phi \gamma_X(h - 1) \]
Example – First-Order Autoregressive Process: III

- repetitive use of $\gamma_X(h) = \phi \gamma_X(h - 1)$ yields
  
  $$\gamma_X(h) = \phi \gamma_X(h - 1) = \phi^2 \gamma_X(h - 2) = \cdots = \phi^h \gamma_X(0)$$

- noting that
  
  $$\gamma_X(-h) = \text{cov} \{X_{t-h}, X_t\} = \text{cov} \{X_t, X_{t-h}\} = \text{cov} \{X_{t+h}, X_t\} = \gamma_X(h)$$

  (note: true for all ACVFs!), have, when $\gamma_X(0) > 0$,

  $$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \phi^{|h|}, \quad h \in \mathbb{Z}$$
Example – First-Order Autoregressive Process: IV

• in addition, we have

\[ \gamma_X(0) = \text{cov} \{X_t, X_t\} \]
\[ = \text{cov} \{\mu + \phi(X_{t-1} - \mu) + Z_t, \mu + \phi(X_{t-1} - \mu) + Z_t\} \]
\[ = \phi^2 \text{cov} \{X_{t-1}, X_{t-1}\} + \phi \text{cov} \{X_{t-1}, Z_t\} + \phi \text{cov} \{Z_t, X_{t-1}\} + \text{cov} \{Z_t, Z_t\} \]
\[ = \phi^2 \gamma_X(0) + \sigma^2, \]

from which we can conclude

\[ \gamma_X(0) = \frac{\sigma^2}{1 - \phi^2} \]

• process \( \{X_t\} \) is called a first-order autoregressive process or AR(1) process (when \( \phi > 0 \), called ‘red noise’ in the geophysical literature)
ACF for AR(1) Process with $\phi = 0.9$
ACF for AR(1) Process with $\phi = 0.5$
ACF for AR(1) Process with $\phi = 0$
ACF for AR(1) Process with $\phi = -0.5$
ACF for AR(1) Process with $\phi = -0.9$
Example – First-Order Autoregressive Process: V

• as examples, let’s generate some realizations of AR(1) processes \( \{X_t\} \) with \( \mu = 0 \) and with \( \{Z_t\} \) taken to be Gaussian WN(0,1)

• to do so, note that \( X_1 \) is Gaussian with zero mean and variance \( 1/(1 - \phi^2) \), which we can generate from a standard normal RV \( Z_1 \) using \( X_1 = Z_1/\sqrt{(1 - \phi^2)} \)

• deviates \( X_2, X_3, \ldots \) can be generated using the defining equation:

\[
X_t = \phi X_{t-1} + Z_t, \quad t = 2, 3, \ldots
\]

• note: \( x_t \) for \( \phi = 0.9 \) is reminiscent of a random walk (why?)

• note: variance \( \gamma_X(0) = 1/(1 - \phi^2) \) of AR(1) process increases as \( |\phi| \) increases

• note: if \( \{Z_t\} \) is non-Gaussian, getting going is not so easy if we want marginal distributions to be the same
$\phi = 0.9 \text{ AR}(1) \ x_t \text{ from Gaussian WN}(0,1)$
\[ \phi = 0.5 \ AR(1) \ x_t \ from \ Gaussian \ WN(0,1) \]
$\phi = 0 \ AR(1) \ x_t$ from Gaussian WN(0,1)
\[ \phi = -0.5 \text{ AR}(1) x_t \text{ from Gaussian WN}(0,1) \]
$\phi = -0.9$ AR(1) $x_t$ from Gaussian WN(0,1)
Sample ACVF\text{s} and ACF\text{s}: I

• given a time series $x_1, x_2, \ldots, x_n$ that is presumed to be a realization of a portion $X_1, X_2, \ldots, X_n$ of a stationary process with mean $\mu$ and ACVF

$$\gamma_X(h) = \text{cov} \{X_{t+h}, X_t\} = E\{(X_{t+h} - \mu)(X_t - \mu)\},$$

we can estimate its ACVF for lags $h$ satisfying $-n < h < n$ using

$$\hat{\gamma}_X(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \text{ where } \bar{x} \equiv \frac{1}{n} \sum_{t=1}^{n} x_t,$$

where $\hat{\gamma}_X(\cdot)$ is called the sample autocovariance function

• note: we are dividing by $n$ rather than $n - |h|$ (?!)

BD–19; CC–46; SS–29 II–62
Sample ACVF\s and ACFs: II

- since the ACF is given by \( \rho_X(h) = \gamma_X(h)/\gamma_X(0) \), the corresponding sample autocorrelation function is given by

\[
\hat{\rho}_X(h) \equiv \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)} = \frac{\sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2}
\]

- sample correlation coefficient for \( u_t \) and \( v_t \), \( t = 1, \ldots, m \), is usually taken to be

\[
\frac{\sum_{t=1}^{m} (u_t - \bar{u})(v_t - \bar{v})}{\sqrt{[\sum_{t=1}^{m} (u_t - \bar{u})^2 \sum_{t=1}^{m} (v_t - \bar{v})^2]}}
\]

note that letting \( u_t = x_{t+|h|}, v_t = x_t \) and \( m = n - |h| \) in the above does not lead to \( \hat{\rho}_X(h) \), as a result of which it is possible to construct (non-pathological!) time series such that \( \hat{\rho}_X(h) \) does not reflect the strength of the linear relationship between \( \{x_1, x_2, \ldots, x_{n-|h|}\} \) and \( \{x_{|h|+1}, x_{|h|+2}, \ldots, x_n\} \)
Distribution of Sample ACF for IID($\mu, \sigma^2$) Noise

- if $\{X_t\}$ is IID($\mu, \sigma^2$) process & $n$ is large, can argue that RV
  \[
  \hat{\rho}_X(h) = \frac{\sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X})(X_t - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}
  \]
  is approximately $\mathcal{N}(0, 1/n)$, where $\mathcal{N}(\mu, \sigma^2)$ denotes RV with a normal (Gaussian) distribution, mean $\mu$ & variance $\sigma^2$

- for fixed $h'$, $\hat{\rho}_X(1), \ldots, \hat{\rho}_X(h')$ are IID $\mathcal{N}(0, 1/n)$ as $n \to \infty$

- implies about 95% of $\hat{\rho}_X(h)$'s should fall within $\pm 1.96/\sqrt{n}$

- note: above approximation breaks down for $h/n$ close to unity

- refined bounds are $\pm 1.96\left(\sqrt{n - |h|}\right)/n$ (Fuller, 1996, p. 336)

- let’s see how these pan out on realizations of Gaussian IID(0,1) $x_t$’s with $n = 200$ at lags $h = 1, \ldots, 40$ (note: 95% of 40 is 38)
Gaussian IID(0,1) $x_t$, 1st Realization
Gaussian IID(0,1) $x_t$, 2nd Realization
Gaussian IID(0,1) $x_t$, 3rd Realization
Sample ACF for 1st Realization
Sample ACF for 2nd Realization

ACF

h (lag)
Sample ACF for 3rd Realization
Reference