**Estimation in time series**

**Assume:** \( \{X_t, t \in \mathbb{Z}\} \) is stationary with unknown mean \( \mu = E X_t \) \& autocovariance \( \gamma(k) = \text{cov}(X_t, X_{t+k}) \).

Assume \( \sup_{k=\infty} |\gamma(k)| < \infty \) \& define \( \sigma^2 = \sup_{k=\infty} \gamma(k) \).

**Data:** \( X_1, \ldots, X_n \)

Consider the sample mean \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \) as estimator of \( \mu \).

1) Show \( \bar{X} \) is unbiased for \( \mu \).
2) Show \( \text{var}(\bar{X}) \to 0 \) as \( n \to \infty \).
3) Show \( \text{var}(\sqrt{n} \bar{X}) \to \sigma^2 \) as \( n \to \infty \).

4) Just for part (4): assume \( X_t \) is MA(1), i.e. \( X_t = Z_t + \Theta Z_{t-1} \), where \( Z_t \sim \text{iid}(0, \sigma^2) \)

and show Central Limit Theorem (CLT)

i.e. that \( \sqrt{n} (\bar{X} - \mu) \approx N(0, \sigma^2) \) for large \( n \).

(HINT: reduce it to a CLT for the iid \( Z_t \) which is well known)

Fix \( k > 0 \). The goal is to estimate \( \gamma(k) \) now.

5) Define a new series \( Y_t = (X_t - \mu)(X_{t+k} - \mu) \)

Show that \( \bar{Y} = \frac{1}{n-k} \sum_{t=1}^{n-k} Y_t \) is a "good" estimator of \( \gamma(k) \), i.e. unbiased and consistent.

6) Denote \( \hat{\gamma}(k) = \bar{Y} = \frac{1}{n-k} \sum_{t=1}^{n-k} (X_t - \mu)(X_{t+k} - \mu) \)

Since \( \mu \) is unknown, it has to be estimated to make this practical. Let \( \bar{Y} = \frac{1}{n-k} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \mu) \).

Argue that \( \hat{\gamma}(k) \) is also a "good" estimator because \( |\hat{\gamma}(k) - \bar{\gamma}(k)| \) is small. How small is it?
Go back at \( \hat{y}(k) \) or \( \tilde{y}(k) \) and note that for large \( k \), they are unreliable. E.g., for \( k = n - 1 \) this is the average of just one (!) observation which can be anywhere. Since we know that \( y(k) \to 0 \) as \( k \to \infty \) [How do we know?]

We can define the final estimator

\[
\hat{y}(k) = \left( 1 - \frac{k}{n} \right) \bar{y}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})
\]

which is called the "sample autocovariance" & has the property that for large \( k \), \( \hat{y}(k) \approx 0 \).

7) Suppose for simplicity that \( \mu \) is known, so we can use the simpler equation \( \hat{y}(k) = \left( 1 - \frac{k}{n} \right) \bar{y}(k) \) for \( k > 0 \). Quantify the bias & the variance of \( \hat{y}(k) \) and show that—although unbiased—\( \tilde{y}(k) \) has bigger MSE (Mean Squared Error) than \( \hat{y}(k) \to \) so \( \hat{y}(k) \) is better!

8) Fit an AR(p) model to the data \( X_1, \ldots, X_n \) via the Yule-Walker equations (AKA the normal equations in projection/prediction) based on the estimated \( \hat{y}(k) \). Give all details including how to estimate the variance of the iid error driving the AR(p).

Generate a sample of size \( n = 200 \) from the AR(3) model:

\[
(1 - 0.5B)(1 + B + B^2) X_t = Z_t
\]

where \( Z_t \) is i.i.d \( \text{N}(0, \sigma^2) \), \( \sigma^2 = 2 \).

9) Estimate \( \hat{y}(k) \) and \( \hat{p}(k) = \hat{y}(k) / \hat{y}(0) \) and plot it; use function acf().

10) Fit an AR(p) model to your data. It is not known that \( p = 3 \) so try different \( p \)'s, i.e. \( p = 1, 2, 3, 4 \) or \( 5 \). What happens to your [estimate of \( c_{62} \)?]