Testing $H_0 : \rho_k = 0$

For an MA(q) process, the theoretical ACF has a cutoff to zero after lag q, i.e., $\rho_k = 0$ for $k > q$.

In the sample ACF, we expect an approximate cutoff to zero after lag q.

Therefore, to identify the order of an MA process, we need to be able to test the null hypothesis that $\rho_k = 0$.

If $\rho_j = 0$ for all $j \geq k$, then an approximate standard error for r_k is given by

$$s(r_k) = \left(1 + 2\sum_{j=1}^{k-1} r_j^2\right)^{1/2} n^{-1/2}$$

and the ratio $r_k/s(r_k)$ has approximately a $N(0,1)$ distribution (if n is large enough).

An approximate level α test of $H_0 : \rho_k = 0$ rejects H_0 when

$$\left| \frac{r_k}{s(r_k)} \right| > z_{\alpha/2} \quad \text{or equivalently} \quad |r_k| > z_{\alpha/2}\, s(r_k)$$

For $\alpha = .05$, use $z_{\alpha/2} = 1.96$.

The ACF plot produced by SAS PROC ARIMA has a band marked at "two standard errors" (at $\pm 2s(r_k)$). This can be used for tests at level $\alpha \approx .05$.

For an MA(q) process, after lag q we expect all (or nearly all) of the values r_k to lie within this band, most of them being well inside the band.

Testing $H_0 : \phi_{kk} = 0$

For an AR(p) process, the theoretical PACF has a cutoff to zero after lag p, i.e., $\phi_{kk} = 0$ for $k > p$.

In the sample PACF, we expect an approximate cutoff to zero after lag p.

Therefore, to identify the order of an AR process, we need to be able to test the null hypothesis that $\phi_{kk} = 0$.

When $\phi_{kk} = 0$, an approximate standard error for $\hat{\phi}_{kk}$ is given by

$$s(\hat{\phi}_{kk}) = n^{-1/2}$$

and the ratio $\hat{\phi}_{kk}/s(\hat{\phi}_{kk})$ has approximately a $N(0,1)$ distribution (if n is large enough).

An approximate level α test of $H_0 : \phi_{kk} = 0$ rejects H_0 when

$$\left| \frac{\hat{\phi}_{kk}}{s(\hat{\phi}_{kk})} \right| > z_{\alpha/2} \quad \text{or equivalently} \quad |\hat{\phi}_{kk}| > z_{\alpha/2}\, n^{-1/2}$$

The PACF plot produced by PROC ARIMA has a band marked at "two standard errors" (at $\pm 2n^{-1/2}$). This can be used for tests at level $\alpha \approx .05$.

For an AR(p) process, after lag p we expect all (or nearly all) of the values $\hat{\phi}_{kk}$ to lie within this band, most of them being well inside the band.

Note: the band for the PACF has constant width. The band for the ACF has increasing width.

Autocorrelation Check for White Noise

This item in the PROC ARIMA output displays the Ljung-Box test computed at lags which are multiples of 6.

A white noise process (random shocks) has $\rho_k = 0$ for all $k \neq 0$.

A test of $H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$ uses the statistic

$$Q(m) = n(n+2)\sum_{k=1}^{m} \frac{r_k^2}{n-k}$$

where n is the length of the time series.

If a series $z_1, z_2, \ldots, z_n$ consists of random shocks (or more generally, of independent identically distributed random variables with enough moments) and n is large enough, then $Q(m)$ has approximately a χ^2_m distribution.

If any of $\rho_1, \rho_2, \ldots, \rho_m$ are non-zero, $Q(m)$ will be "larger than χ^2_m".

We reject H_0 at level α if $Q(m) > \chi^2_m(\alpha)$, the upper α point of the chi-squared distribution with m degrees of freedom. SAS displays the p-values, so we reject if p-value $< \alpha$.

PROC ARIMA displays $Q(6)$, $Q(12)$, $Q(18)$, $Q(24)$.

For a random shock sequence, we expect all of these to be non-significant (but of course, for each of them there is a probability α of rejecting H_0 by chance).

Properties of AR(1) Processes

AR(1): $z_t = C + \phi_1 z_{t-1} + a_t$ with a_t independent $N(0, \sigma_a^2)$.

• An AR(1) process is stationary if $|\phi_1| < 1$, and non-stationary if $|\phi_1| \geq 1$.

For a stationary AR(1) process:

- $\mu_z = \dfrac{C}{1 - \phi_1}$
- $\sigma_z^2 = \dfrac{\sigma_a^2}{1 - \phi_1^2}$
- $z_t \sim N(\mu_z, \sigma_z^2)$.

  (This is the marginal or unconditional distribution of z_t when given no information about past values.)
- $\rho_k = (\phi_1)^k$, for $k = 0, 1, 2, \ldots$
  (exponential decay to 0 as $k \to \infty$)

Some derivations

Let X, Y, Z be random variables (with finite means and variances), and a, b, c, d be constants.

Rules for Expected Values:

- $E(X+Y) = EX + EY$,
  $E(X+Y+Z) = EX + EY + EZ$, etc.
- $E(bX) = b(EX)$.
- $E(c+X) = c + EX$.
- In combination: $E(bX+c) = b(EX) + c$,
  $E(aX+bY+cZ+d) = a\,EX + b\,EY + c\,EZ + d$, etc.

Derivation 1:

$$
\begin{aligned}
& z_t = C + \phi_1 z_{t-1} + a_t \\
\Rightarrow\ & E(z_t) = E(C + \phi_1 z_{t-1} + a_t) \qquad \text{(taking } E \text{ on both sides)} \\
\Rightarrow\ & E(z_t) = C + \phi_1 E(z_{t-1}) + E(a_t) \qquad \text{(applying the rules)} \\
\Rightarrow\ & \mu_z = C + \phi_1 \mu_z + 0 \qquad \text{(now solve for } \mu_z\text{)} \\
\Rightarrow\ & (1 - \phi_1)\mu_z = C \\
\Rightarrow\ & \mu_z = \frac{C}{1 - \phi_1}
\end{aligned}
$$

In the above we used:

- C and ϕ_1 are constants; z_t, z_{t-1}, a_t are random variables.
- $Ez_t = Ez_{t-1} = \mu_z$ since the process is stationary.
- $Ea_t = 0$.

Some Rules for Variances:

- $\text{Var}(X + b) = \text{Var}(X)$.
- $\text{Var}(cX) = c^2\text{Var}(X)$.
- In combination: $\text{Var}(cX + b) = c^2\text{Var}(X)$.
- If X and Y are independent (or even just uncorrelated), then $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
  Similarly, if X, Y, Z are independent, then $\text{Var}(X + Y + Z) = \text{Var}(X) + \text{Var}(Y) + \text{Var}(Z)$, etc.
- In combination: If X, Y are independent, then $\text{Var}(bX + cY + d) = b^2\,\text{Var}(X) + c^2\,\text{Var}(Y)$, etc.

Derivation 2:

$$
\begin{aligned}
& z_t = C + \phi_1 z_{t-1} + a_t \\
\Rightarrow\ & \text{Var}(z_t) = \text{Var}(C + \phi_1 z_{t-1} + a_t) \qquad \text{(taking Var on both sides)} \\
\Rightarrow\ & \text{Var}(z_t) = \phi_1^2 \text{Var}(z_{t-1}) + \text{Var}(a_t) \qquad \text{(applying the rules)} \\
\Rightarrow\ & \sigma_z^2 = \phi_1^2 \sigma_z^2 + \sigma_a^2 \qquad \text{(now solve for } \sigma_z^2\text{)} \\
\Rightarrow\ & (1 - \phi_1^2)\sigma_z^2 = \sigma_a^2 \\
\Rightarrow\ & \sigma_z^2 = \frac{\sigma_a^2}{1 - \phi_1^2}
\end{aligned}
$$

In the above we used:

- $\text{Var}(z_t) = \text{Var}(z_{t-1}) = \sigma_z^2$ since the process is stationary .
- $\text{Var}(a_t) = \sigma_a^2$.
- z_{t-1} and a_t are independent. (Why?)

General Facts: For any stationary ARMA(p, q) process:

- For any time t, the random variable z_t will be independent of all the "future" random shocks $\{a_{t+1}, a_{t+2}, a_{t+3} \dots\}$.
- The process can be re-written as an MA(∞) process:

$$z_t = \mu_z + a_t + \sum_{i=1}^{\infty} \psi_i a_{t-i}$$

where $\psi_i \to 0$ as $i \to \infty$ (approaching zero at an eventually exponential rate).

For simplicity, we will show this only for a stationary AR(1) process with $C = 0$.

We know that:

(A): $z_t = a_t + \phi_1 z_{t-1}$,
(B): $z_{t-1} = a_{t-1} + \phi_1 z_{t-2}$,
(C): $z_{t-2} = a_{t-2} + \phi_1 z_{t-3}$, etc.

Substituting (B) into (A) gives:

$$\begin{aligned} z_t &= a_t + \phi_1(a_{t-1} + \phi_1 z_{t-2}) \\ &= a_t + \phi_1 a_{t-1} + \phi_1^2 z_{t-2} \, . \end{aligned}$$

Substituting (C) into the above gives:

$$\begin{aligned} z_t &= a_t + \phi_1 a_{t-1} + \phi_1^2(a_{t-2} + \phi_1 z_{t-3}) \\ &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \phi_1^3 z_{t-3} \end{aligned}$$

and the pattern is clear.

Repeated substitution shows for any $k > 0$ that

$$\begin{aligned} z_t &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \cdots + \phi_1^k a_{t-k} + \phi_1^{k+1} z_{t-k-1} \\ &= a_t + \sum_{i=1}^{k} \phi_1^i a_{t-i} + \phi_1^{k+1} z_{t-k-1} \,. \end{aligned}$$

Since $|\phi_1| < 1$, the term $\phi_1^{k+1} z_{t-k-1}$ goes to zero as $k \to \infty$ and the sum converges to give

$$z_t = a_t + \sum_{i=1}^{\infty} \phi_1^i a_{t-i}$$

which has the desired form with $\psi_i = \phi_1^i$.

The same process of repeated substitution works for any ARMA process.

Since the random shocks are all independent of each other, and z_t is expressed entirely as a combination of the "present" shock a_t and "past" shocks $a_{t-1}, a_{t-2}, \ldots$, it will be independent of the "future" shocks $a_{t+1}, a_{t+2}, \ldots$ QED

The Mean of a General Stationary ARMA(p, q) Process

Suppose

$$z_t = C + \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \qquad (*)$$

is stationary. Taking expected values on both sides gives

$$\mu_z = C + \phi_1 \mu_z + \cdots + \phi_p \mu_z + 0 - \theta_1 0 - \cdots - \theta_q 0 \qquad (**)$$

and solving for μ_z produces

$$\mu_z = \frac{C}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}\,.$$

Mean Centering

Suppose $\{z_t\}$ is a stationary ARMA(p, q) process.

For all t, define $\tilde{z}_t = z_t - \mu_z$ so that $E\tilde{z}_t = 0$. We call $\tilde{z}_t$ the "mean centered process".

Subtracting $(**)$ from $(*)$ on the previous page, we find

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \cdots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \,.$$

(The constant C has disappeared.)

Another Rule for Expected Values: Products of Independent RV's

- If X and Y are independent (with finite means), then

$$E(XY) = (EX)(EY)\,.$$

Special Cases: Suppose $\{z_t\}$ is a stationary ARMA process and $\{a_t\}$ is the sequence of random shocks used to generate $\{z_t\}$.

- $E(a_s a_t) = 0$ if $s \neq t$.
- $E(z_s a_t) = 0$ if $s < t$.

Proofs:

- $s \neq t \;\Rightarrow\; a_s, a_t$ indep. $\;\Rightarrow\; E(a_s a_t) = (Ea_s)(Ea_t) = 0 \cdot 0 = 0.$
- $s < t \;\Rightarrow\; z_s, a_t$ indep. $\;\Rightarrow\; E(z_s a_t) = (Ez_s)(Ea_t) = \mu_z \cdot 0 = 0.$

Deriving the Theoretical ACF of Stationary AR(1)

Suppose $k > 0$.

$$
\begin{aligned}
\mathrm{Cov}(z_t, z_{t-k}) =& E(z_t - \mu_z)(z_{t-k} - \mu_z) \\
=& E\tilde{z}_t\tilde{z}_{t-k} \\
=& E[(\phi_1\tilde{z}_{t-1} + a_t)\tilde{z}_{t-k}] \\
=& E(\phi_1\tilde{z}_{t-1}\tilde{z}_{t-k} + a_t\tilde{z}_{t-k}) \\
=& \phi_1 E\tilde{z}_{t-1}\tilde{z}_{t-k} + Ea_t\tilde{z}_{t-k} \\
=& \phi_1 \mathrm{Cov}(z_{t-1}, z_{t-k}) + 0 \\
& \text{(since } a_t \text{ is a future shock relative to } z_{t-k}\text{)}
\end{aligned}
$$

Therefore (dividing by σ_z^2)

$$
\frac{\mathrm{Cov}(z_t, z_{t-k})}{\sigma_z^2} = \phi_1 \frac{\mathrm{Cov}(z_{t-1}, z_{t-k})}{\sigma_z^2} \quad \Rightarrow \quad \rho_k = \phi_1 \rho_{k-1}
$$

Setting $k = 1$ gives $\rho_1 = \phi_1 \rho_0 = \phi_1$ since $\rho_0 = 1$.

Setting $k = 2$ gives $\rho_2 = \phi_1 \rho_1 = \phi_1^2$.

Setting $k = 3$ gives $\rho_3 = \phi_1 \rho_2 = \phi_1^3$, etc.

Thus $\rho_k = \phi_1^k$ for all $k = 0, 1, 2, \ldots$

A similar approach works to derive recursions for the ACF of any ARMA process. (But it is messier.)

Autocovariance

For any stationary process $\{z_t\}$ define the **autocovariances**:

$$\gamma_k = \text{Cov}(z_t, z_{t-k}) = \text{Cov}(z_t, z_{t+k}) \quad \text{for } k = 0, \pm 1, \pm 2, \ldots$$

Note:

- $\gamma_0 = \text{Var}(z_t) = \sigma_z^2$.
- $\gamma_k = \gamma_{-k}$ for all k.
- $\rho_k = \dfrac{\gamma_k}{\gamma_0}$ for all k.

Mean, Variance, and Autocovariances of an MA(q) Process

For convenience, define $\psi_0 = 1$, and $\psi_i = -\theta_i$ for $i = 1, \ldots, q$ so that

$$\begin{aligned} z_t &= C + \quad a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \\ &= C + \psi_0 a_t + \psi_1 a_{t-1} + \cdots + \psi_q a_{t-q} \end{aligned}$$

(1) $\mu_z = C$ (We already knew this one.)

(2) $\sigma_z^2 = \gamma_0 = \sigma_a^2 \sum_{i=0}^{q} \psi_i^2$

(3) $\gamma_k = \sigma_a^2 \sum_{i=0}^{q-k} \psi_i \psi_{i+k}$ for $k = 0, 1, \ldots, q$

(4) $\gamma_k = 0$ for $k > q$

Fact (4) gives the cutoff to zero after lag q in the ACF.

Proofs:

(4):

$$\begin{aligned} z_t =& C + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \\ z_{t-k} =& C + a_{t-k} - \theta_1 a_{t-k-1} - \cdots - \theta_q a_{t-k-q} \end{aligned}$$

So

$$\begin{aligned} &z_t \text{ depends on } \{a_t, a_{t-1}, \ldots, a_{t-q}\} \\ &z_{t-k} \text{ depends on } \{a_{t-k}, a_{t-k-1}, \ldots, a_{t-k-q}\} \end{aligned}$$

If $k > q$, there is no overlap in these two groups so that z_t and z_{t-k} are independent (since the two groups of random shocks are independent of each other). When z_t and z_{t-k} are independent, their covariance is zero. (Recall: Independent $\Rightarrow$ Uncorrelated)

(2): To simplify the notation, take $q = 2$ and $t = 3$. We illustrate the argument in this special case. (It works in general.)

$$\tilde{z}_3 = \psi_0 a_3 + \psi_1 a_2 + \psi_2 a_1$$

so that

$$\begin{aligned}
\mathrm{Var}(z_3) =& E\left[(z_3 - \mu_z)^2\right] = E\left[\,\tilde{z}_3^2\,\right] = E\left[(\psi_0 a_3 + \psi_1 a_2 + \psi_2 a_1)^2\right] \\
=& E\left[\psi_0^2 a_3^2 + \psi_1^2 a_2^2 + \psi_2^2 a_1^2\right. \\
& \qquad \left. + 2\psi_0\psi_1 a_3 a_2 + 2\psi_0\psi_2 a_3 a_1 + 2\psi_1\psi_2 a_2 a_1\right] \\
=& \psi_0^2 E a_3^2 + \psi_1^2 E a_2^2 + \psi_2^2 E a_1^2 \\
& \qquad + 2\psi_0\psi_1 E a_3 a_2 + 2\psi_0\psi_2 E a_3 a_1 + 2\psi_1\psi_2 E a_2 a_1 \\
=& \psi_0^2 \sigma_a^2 + \psi_1^2 \sigma_a^2 + \psi_2^2 \sigma_a^2 \\
& \qquad + 2\psi_0\psi_1\, 0 + 2\psi_0\psi_2\, 0 + 2\psi_1\psi_2\, 0 \\
=& \sigma_a^2(\psi_0^2 + \psi_1^2 + \psi_2^2)
\end{aligned}$$

(3): For simplicity, take $q = 2$ and particular values of t:

$$\tilde{z}_4 = \psi_0 a_4 + \psi_1 a_3 + \psi_2 a_2$$
$$\tilde{z}_3 = \psi_0 a_3 + \psi_1 a_2 + \psi_2 a_1$$

Then

$$\begin{aligned}
\gamma_1 &= \mathrm{Cov}(z_4, z_3) = E(z_4 - \mu_z)(z_3 - \mu_z) = E\tilde{z}_4\tilde{z}_3 \\
&= E(\psi_0 a_4 + \psi_1 a_3 + \psi_2 a_2)(\psi_0 a_3 + \psi_1 a_2 + \psi_2 a_1) \\
&= \psi_1\psi_0 E a_3^2 + \psi_2\psi_1 E a_2^2 + 0 + \cdots + 0 \\
&= \sigma_a^2(\psi_0\psi_1 + \psi_1\psi_2)
\end{aligned}$$

The argument for $\gamma_2 = \sigma_a^2\psi_0\psi_2$ is similar.

Why Does AR(p) PACF Cutoff to Zero After Lag p?

Here is an intuitive argument.

Consider the special case AR(2). We know

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t \qquad (*)$$

and that a_t is independent of $z_{t-1}, z_{t-2}, z_{t-3}, \ldots$ Thus $(*)$ is essentially a regression model with error term a_t. Since the error a_t is independent of z_{t-3}, adding z_{t-3} to the regression model cannot improve the fit. In other words, if we add z_{t-3} to the model, the regression coefficient ϕ_{33} must be zero.

Similarly, since both z_{t-3} and z_{t-4} are independent of a_t, adding them both to the model $(*)$ cannot improve the fit. In other words, if we add z_{t-3} and z_{t-4} to the model, both regression coefficients ϕ_{43} and ϕ_{44} must be zero.

Thus $\phi_{33} = \phi_{44} = 0$ and so on.