Bartlett's Formula

Suppose $(X_t) \sim \text{ARMA}\,(p, q)$. Then according to Bartlett's formula, under suitable conditions, as $n \to \infty$,

$$\sqrt{n}\left(\begin{bmatrix} r_1 \\ \vdots \\ r_k \end{bmatrix} - \begin{bmatrix} \rho(1) \\ \vdots \\ \rho(k) \end{bmatrix}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{W})$$

where

$$W_{ij} = \sum_{m=1}^{\infty} \left[\rho_X(m+i) + \rho_X(m-i) - 2\rho_X(i)\rho_X(m)\right]\left[\rho_X(m+j) + \rho_X(m-j) - 2\rho_X(j)\rho_X(m)\right]$$

1. Let X_t be an MA$\,(q)$ process with autocorrelation function $\rho_X(\cdot)$. **For $\boldsymbol{h > q}$**, use Bartlett's formula to find the approximate variance of the sample acf at lag h, r_h, in terms of $\rho_X(\cdot)$. Your answer should be a finite sum instead of an infinite series.

2. Consider the following MA (2) process:

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2}, \quad W_t \overset{iid}{\sim} N(0,1).$$

For each of the model parameter pairs (θ_1, θ_2) below,

(a) $(\theta_1, \theta_2) = (1, 1)$
(b) $(\theta_1, \theta_2) = (-0.9, 0.3)$
(c) $(\theta_1, \theta_2) = (-0.6, -0.3)$

(i) Simulate $n = 300$ observations of the process. Use 153 as your seed.

(ii) Plot the sample ACF with $\pm 1.96/\sqrt{n}$ bars: these are the default for the acf() function and represent the distribution of r_h for white noise, aka ARMA(0,0).

(iii) Now let's visualize the distribution of ARMA(0,2), usually simply called MA(2). Add to this plot the 95% intervals based on Bartlett's formula for lag $h > q$ in red that you calculated on problem 1.

(iv) Add to this plot the expectation and 95% interval for r_1 and r_2. You'll need Bartlett's formula again, but instead of simplifying, you can just let the computer do it. I would use a for loop over the "m" values; as m increases, the summand will eventually be zero.

For this problem, you'll need the ACF of MA(2). To "go the extra mile", you could also derive the ACF at lags 1 and 2 here. But if you know you can do this already, here it is.

$$\rho_{MA(2)}(h) = \begin{cases} 1 & h = 0 \\ \frac{\theta_1(1+\theta_2)}{(1+\theta_1^2+\theta_2^2)} & h = 1 \\ \frac{\theta_2}{(1+\theta_1^2+\theta_2^2)} & h = 2 \\ 0 & h \geq 3 \end{cases} \tag{1}$$

Best Prediction vs Best Linear Prediction

(Adapted from Shumway and Stoffer 3.14) Suppose we wish to find a prediction function $g(x)$ that minimizes

$$\text{MSE} = \mathbb{E}\left[(Y - g(X))^2\right],$$

where X and Y are jointly distributed random variables with density $f(x, y)$.

3. Show that the best prediction of Y from X is $g(x) = \mathbb{E}(Y|X = x)$. In other words, show that the MSE is minimized by the choice
$$g(x) = \mathbb{E}\left[Y \mid X = x\right]$$
Hint: "Add 0" to the MSE and expand.

4. Apply the above result to the model $Y = X^2 + 3X + Z - 1$, where X and Z are independent zero-mean normal variables with variance one. Find the MSE of the best prediction.

5. Note that Y and X both have mean zero. Find the best linear prediction of Y from X (in other words, find b that minimizes $MSE = E(Y - bX)^2)$. Find the MSE of the best linear prediction.