Lent Term 2001 Richard Weber

Time Series — Examples Sheet

This is the examples sheet for the M. Phil. course in Time Series. A copy can be found at: `http://www.statslab.cam.ac.uk/~rrw1/timeseries/`

Throughout, unless otherwise stated, the sequence $\{\epsilon_t\}$ *is white noise, variance* σ^2.

1. Find the Yule-Walker equations for the AR(2) process

$$X_t = \tfrac{1}{3}X_{t-1} + \tfrac{2}{9}X_{t-2} + \epsilon_t\,.$$

Hence show that it has autocorrelation function

$$\rho_k = \tfrac{16}{21}\left(\tfrac{2}{3}\right)^{|k|} + \tfrac{5}{21}\left(-\tfrac{1}{3}\right)^{|k|}, \quad k \in \mathbb{Z}\,.$$

[The Yule-Walker equations are

$$\rho_k - \tfrac{1}{3}\rho_{k-1} - \tfrac{2}{9}\rho_{k-2} = 0\,, \quad k \geq 2.$$

On trying $\rho_k = A\lambda^k$, we require $\lambda^2 - \frac{1}{3}\lambda - \frac{2}{9} = 0$. This has roots $\frac{2}{3}$ and $-\frac{1}{3}$, so

$$\rho_k = A\left(\tfrac{2}{3}\right)^{|k|} + B\left(-\tfrac{1}{3}\right)^{|k|}\,,$$

where $\rho_0 = A + B = 1$. We also require $\rho_1 = \frac{1}{3} + \frac{2}{9}\rho_1$. Hence $\rho_1 = \frac{3}{7}$, and thus we require $\frac{2}{3}A - \frac{1}{3}B = \frac{3}{7}$. These give $A = \frac{16}{21}$, $B = \frac{5}{21}$.]

2. Let $X_t = A\cos(\Omega t+U)$, where A is an arbitrary constant, Ω and U are independent random variables, Ω has distribution function F over $[0,\pi]$, and U is uniform over $[0,2\pi]$. Find the autocorrelation function and spectral density function of $\{X_t\}$. Hence show that, for any positive definite set of covariances $\{\gamma_k\}$, there exists a process with autocovariances $\{\gamma_k\}$ such that every realization is a sine wave.

[Use the following definition: $\{\gamma_k\}$ are positive definite if there exists a nondecreasing function F such that $\gamma_k = \int_{-\pi}^{\pi} e^{ik\omega} dF(\omega)$.]

[

$$\mathbb{E}[X_t \mid \Omega] = \tfrac{1}{2\pi}\int_0^{2\pi} A\cos(\Omega t+u)\,du = \tfrac{1}{2\pi}\,A\sin(\Omega t+u)\big|_0^{2\pi} = 0$$

$$\begin{aligned}
\mathbb{E}[X_{t+s}X_t] &= \tfrac{1}{2\pi}\int_0^{\pi}\int_0^{2\pi} A\cos(\Omega(t+s)+u)A\cos(\Omega t+u)\,du\,dF(\Omega)\\
&= \tfrac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} A^2[\cos(\Omega(2t+s)+2u)+\cos(\Omega s)]\,du\,dF(\Omega)\\
&= \tfrac{1}{2}\int_0^{\pi} A^2\cos(\Omega s)dF(\Omega)\\
&= \tfrac{1}{4}\int_0^{\pi} A^2\left[e^{i\Omega s}+e^{-i\Omega s}\right]dF(\Omega)\\
&= \tfrac{1}{4}A^2\int_{-\pi}^{\pi} e^{i\Omega s}d\bar{F}(\Omega)
\end{aligned}$$

where we define over the range $[-\pi,\pi]$ the nondecreasing function $\bar{F}$, by $\bar{F}(-\Omega) = F(\pi)-F(\Omega)$ and $\bar{F}(\Omega) = F(\Omega)+F(\pi)-2F(0)$, $\Omega\in[0,\pi]$.]

3. Find the spectral density function of the AR(2) process

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t\,.$$

What conditions on (ϕ_1, ϕ_2) are required for this process to be an indeterministic second order stationary? Sketch in the (ϕ_1, ϕ_2) plane the stationary region.

[We have

$$f_X(\omega)\left|1 - \phi_1 e^{i\omega} - \phi_2 e^{2i\omega}\right|^2 = \sigma^2/\pi$$

Hence

$$f_X(\omega) = \frac{\sigma^2}{\pi\left[1 + \phi_1^2 + \phi_2^2 + 2(-\phi_1 + \phi_1\phi_2)\cos(\omega) - 2\phi_2\cos(2\omega)\right]}$$

The Yule-Walker equations have solution of the form $\rho_k = A\lambda_1^k + B\lambda_2^k$ where λ_1, λ_2 are roots of

$$g(\lambda) = \lambda^2 - \phi_1\lambda - \phi_2 = 0.$$

The roots are $\lambda = \left[\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}\right]/2$. To be indeterministic second order stationary these roots must have modulus less that 1. If $\phi_1^2 + 4\phi_2 > 0$ then the roots are real and lie in the range $[-1, 1]$ if and only if $g(-1) > 0$ and $g(1) > 0$, i.e., $\phi_1 + \phi_2 < 1$, $\phi_1 - \phi_2 > -1$. If $\phi_1^2 + 4\phi_2 < 0$ then the roots are complex and their product must be less than 1, i.e., $\phi_2 > -1$. The union of these two regions, corresponding to possible (ϕ_1, ϕ_2) for real and imaginary roots, is simply the triangular region

$$\phi_1 + \phi_2 < 1, \quad \phi_1 - \phi_2 > -1, \quad \phi_2 \geq -1\,.$$

]

4. For a stationary process define the covariance generating function

$$g(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k, \quad |z| < 1 .$$

Suppose $\{X_t\}$ satisfies $X = C(B)\epsilon$, that is, it has the Wold representation

$$X_t = \sum_{r=0}^{\infty} c_r \epsilon_{t-r} ,$$

where $\{c_r\}$ are constants satisfying $\sum_0^\infty c_r^2 < \infty$ and $C(z) = \sum_{r=0}^\infty c_r z^r$. Show that

$$g(z) = C(z)C(z^{-1})\sigma^2 .$$

Explain how this can be used to derive autocovariances for the ARMA(p, q) model. Hence show that for ARMA$(1, 1)$, $\rho_2^2 = \rho_1\rho_3$. How might this fact be useful?

[We have

$$\gamma_k = \mathbb{E}X_t X_{t+k} = \mathbb{E}\left[\sum_{r=0}^{\infty} c_r \epsilon_{t-r} \sum_{s=0}^{\infty} c_s \epsilon_{t+k-s}\right]$$
$$= \sigma^2 \sum_{r=0}^{\infty} c_r c_{k+r}$$

Now

$$C(z)C(z^{-1}) = \sum_{r=0}^{\infty} c_r z^r \sum_{s=0}^{\infty} c_s z^{-s}$$

The coefficients of z^k and z^{-k} are clearly

$$c_k c_0 + c_{k+1}c_1 + c_{k+2}c_3 + \cdots$$

from which the result follows.

For the ARMA(p, q) model $\phi(B)X = \theta(B)\epsilon$ or

$$X = \frac{\theta(B)}{\phi(B)}\epsilon$$

where ϕ and θ are polynomials of degrees p and q in z. Hence

$$C(z) = \frac{\theta(z)}{\phi(z)}$$

and γ_k can be found as the coefficient of z^k in the power series expansion of $\sigma^2\theta(z)\theta(1/z)/\phi(z)\phi(1/z)$. For ARMA$(1,1)$ this is

$$\sigma^2(1+\theta z)(1+\theta z^{-1})(1+\phi z+\phi^2 z^2+\cdots)(1+\phi z^{-1}+\phi^2 z^{-2}+\cdots)$$

from which we have

$$\begin{aligned}\gamma_1 &= \Bigl(\theta(1+\phi^2+\phi^4+\cdots)\\ &\quad+(\phi+\phi^3+\phi^5+\cdots)(1+\theta^2)+\theta(\phi^2+\phi^4+\phi^6+\cdots)\Bigr)\sigma^2\\ &= \frac{\theta+\phi(1+\theta^2)+\phi^2\theta}{1-\phi^2}\sigma^2\end{aligned}$$

and similarly

$$\gamma_2 = \phi\frac{\theta+\phi(1+\theta^2)+\phi^2\theta}{1-\phi^2}\sigma^2 \qquad \gamma_3 = \phi^2\frac{\theta+\phi(1+\theta^2)+\phi^2\theta}{1-\phi^2}\sigma^2$$

Hence $\rho_2^2=\rho_1\rho_3$. This might be used as a diagnostic to test the appropriateness of an ARMA$(1,1)$ model, by reference to the correlogram, where we would expect to see $r_2^2=r_1r_3$.]

5. Consider the ARMA$(2, 1)$ process defined as

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t + \theta_1 \epsilon_{t-1} \,.$$

Show that the coefficients of the Wold representation satisfy the difference equation

$$c_k = \phi_1 c_{k-1} + \phi_2 c_{k-2}, \quad k \geq 2 \,,$$

and hence that

$$c_k = A z_1^{-k} + B z_2^{-k} \,,$$

where z_1 and z_2 are zeros of $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$, and A and B are constants. Explain how in principle one could find A and B.

[The recurrence is produced by substituting $X_t = \sum_{r=0}^{\infty} c_r \epsilon_{t-r}$ into the defining equation, and similarly for X_{t-1} and X_{t-2}, multiplying by ϵ_{t-k}, $k \geq 2$, and taking expected value.

The general solution to such a second order linear recurrence relation is of the form given and we find A and B by noting that

$$X_t = \phi_1 \left(\phi_1 X_{t-2} + \phi_2 X_{t-3} + \epsilon_{t-1} + \theta_1 \epsilon_{t-2} \right) + \phi_2 X_{t-2} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

so that $c_0 = 1$ and $c_1 = \theta_1 + \phi_1$. Hence $A + B = 1$ and $A z_1^{-1} + B z_2^{-1} = \theta_1 + \phi_1$. These can be solved for A and B.]

6. Suppose

$$Y_t = X_t + \epsilon_t, \qquad X_t = \alpha X_{t-1} + \eta_t\,,$$

where $\{\epsilon_t\}$ and $\{\eta_t\}$ are independent white noise sequences with common variance σ^2. Show that the spectral density function of $\{Y_t\}$ is

$$f_Y(\omega) = \frac{\sigma^2}{\pi}\left\{\frac{2 - 2\alpha\cos\omega + \alpha^2}{1 - 2\alpha\cos\omega + \alpha^2}\right\}.$$

For what values of p, d, q is the autocovariance function of $\{Y_t\}$ identical to that of an ARIMA(p, d, q) process?

[

$$\begin{aligned} f_Y(\omega) = f_X(\omega) + f_\epsilon(\omega) &= \frac{1}{|1 - \alpha e^{i\omega}|^2} f_\eta(\omega) + f_\epsilon(\omega) \\ &= \frac{\sigma^2}{\pi}\left\{\frac{1}{1 - 2\alpha\cos\omega + \alpha^2} + 1\right\} = \frac{\sigma^2}{\pi}\left\{\frac{2 - 2\alpha\cos\omega + \alpha^2}{1 - 2\alpha\cos\omega + \alpha^2}\right\}. \end{aligned}$$

We recognise this as the spectral density of an ARMA$(1,1)$ model. E.g., $Z_t - \alpha Z_{t-1} = \xi_t - \theta\xi_{t-1}$, choosing θ and σ_ξ^2 such that

$$(1 - 2\theta\cos\omega + \theta^2)\sigma_\xi^2 = (\sigma^2/\pi)(2 - 2\alpha\cos\omega + \alpha^2)$$

I.e., choosing θ such that $(1+\theta^2)/\theta = (2+\alpha^2)/\alpha$.]

7. Suppose $X_1, \ldots, X_T$ are values of a time series. Prove that

$$\left\{ \hat{\gamma}_0 + 2\sum_{k=1}^{T-1} \hat{\gamma}_k \right\} = 0\,,$$

where $\hat{\gamma}_k$ is the usual estimator of the kth order autocovariance,

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^{T} (X_t - \bar{X})(X_{t-k} - \bar{X})\,.$$

Hint: Consider $0 = \sum_{t=1}^{T}(X_t - \bar{X})$.

Hence deduce that not all ordinates of the correlogram can have the same sign.

Suppose $f(\cdot)$ is the spectral density and $I(\cdot)$ the periodogram. Suppose f is continuous and $f(0) \neq 0$. Does $\mathbb{E}I(2\pi/T) \to f(0)$ as $T \to \infty$?

[The results follow directly from

$$\frac{1}{T}\left[\sum_{t=1}^{T}(X_t - \bar{X})\right]^2 = 0\,.$$

Note that formally,

$$I(0) = \hat{\gamma}_0 + 2\sum_{k=1}^{T-1} \hat{\gamma}_k = 0\,.$$

so it might appear that $\mathbb{E}I(2\pi/T) \to I(0) \neq f(0)$ as $T \to \infty$. However, this would be mistaken. It is a theorem that as $T \to \infty$, $I(\omega_j) \sim f(\omega_j)\chi_2^2/2$. So for large T, $\mathbb{E}I(2\pi/T) \approx f(0)$.]

8. Suppose $I(\cdot)$ is the periodogram of $\epsilon_1, \ldots, \epsilon_T$, where these are i.i.d. $N(0,1)$ and $T = 2m+1$. Let ω_j, ω_k be two distinct Fourier frequencies, Show that $I(\omega_j)$ and $I(\omega_k)$ are independent random variables. What are their distributions?

If it is suspected that $\{\epsilon_t\}$ departs from white noise because of the presence of a single harmonic component at some unknown frequency ω a natural test statistic is the maximum periodogram ordinate

$$T = \max_{j=1,\ldots,m} I(\omega_j) \,.$$

Show that under the hypothesis that $\{\epsilon_t\}$ is white noise

$$P(T > t) = 1 - \left\{1 - \exp\left(-\pi t/\sigma^2\right)\right\}^m \,.$$

[The independence of $I(\omega_j)$ and $I(\omega_k)$ was proved in lectures. Their distributions are $(\sigma^2/2\pi)\chi_2^2$, which is equivalent to the exponential distribution with mean σ^2/π. Hence the probability that the maximum is less than t is the probability that all are, i.e.,

$$P(T < t) = \left\{1 - \exp\left(-\pi t/\sigma^2\right)\right\}^m \,.$$

]

9. Complete this sketch of the fast Fourier transform. From data $X_0, \dots, X_T$, with $T = 2^M - 1$, we want to compute the 2^{M-1} ordinates of the periodogram

$$I(\omega_j) = \frac{1}{\pi T} \left| \sum_{t=0}^{T} X_t e^{it2\pi j/2^M} \right|^2, \quad j = 1, \dots, 2^{M-1}.$$

This requires order T multiplications for each j and so order T^2 multiplications in all. However,

$$\begin{aligned}
\sum_{t=0,1,\dots,2^M-1} X_t e^{it2\pi j/2^M} &= \sum_{t=0,2,\dots,2^M-2} X_t e^{it2\pi j/2^M} + \sum_{t=1,3,\dots,2^M-1} X_t e^{it2\pi j/2^M} \\
&= \sum_{t=0,1,\dots,2^{M-1}-1} X_{2t} e^{i2t2\pi j/2^M} + \sum_{t=0,1,\dots,2^{M-1}-1} X_{2t+1} e^{i(2t+1)2\pi j/2^M} \\
&= \sum_{t=0,1,\dots,2^{M-1}-1} X_{2t} e^{it2\pi j/2^{M-1}} + e^{i2\pi j/2^M} \sum_{t=0,1,\dots,2^{M-1}-1} X_{2t+1} e^{it2\pi j/2^{M-1}}.
\end{aligned}$$

Note that the value of either sum on the right hand side at $j = k$ is the complex conjugate of its value at $j = (2^{M-1} - k)$; so these sums need only be computed for $j = 1, \dots, 2^{M-2}$. Thus we have two sums, each of which is similar to the sum on the left hand side, but for a problem half as large. Suppose the computational effort required to work out each right hand side sum (for all 2^{M-2} values of j) is $\Theta(M-1)$. The sum on the left hand side is obtained (for all 2^{M-1} values of j) by combining the right hand sums, with further computational effort of order 2^{M-1}. Explain

$$\Theta(M) = a2^{M-1} + 2\Theta(M-1).$$

Hence deduce that $I(\cdot)$ can be computed (by the FFT) in time $T \log_2 T$.

[The derivation of the recurrence for $\Theta(M)$ should be obvious. We have $\Theta(1) = 1$, and hence $\Theta(M) = aM2^M = O(T \log_2 T)$.]

10. Suppose we have the ARMA$(1,1)$ process

$$X_t = \phi X_{t-1} + \epsilon_t + \theta\epsilon_{t-1}\,,$$

with $|\phi| < 1$, $|\theta| < 1$, $\phi + \theta \neq 0$, observed up to time T, and we want to calculate k-step ahead forecasts $\hat{X}_{T,k}$, $k \geq 1$.

Derive a recursive formula to calculate $\hat{X}_{T,k}$ for $k = 1$ and $k = 2$.

[

$$\begin{aligned}
\hat{X}_{T,1} &= \phi X_T + \hat{\epsilon}_{T+1} + \theta\hat{\epsilon}_T = \phi X_T + \theta(X_T - \hat{X}_{T-1,1}) \\
\hat{X}_{T,2} &= \phi\hat{X}_{T,1} + \hat{\epsilon}_{T+2} + \theta\hat{\epsilon}_{T+1} = \phi\hat{X}_{T,1}
\end{aligned}$$

]

11. Consider the stationary scalar-valued process $\{X_t\}$ generated by the moving average, $X_t = \epsilon_t - \theta\epsilon_{t-1}$.

Determine the linear least-square predictor of X_t, in terms of $X_{t-1}, X_{t-2}, \ldots$.

[We can directly apply our results to give

$$\begin{aligned}\hat{X}_{t-1,1} &= -\theta\hat{\epsilon}_{t-1} \\ &= -\theta[X_{t-1} - \hat{X}_{t-2,1}] \\ &= -\theta X_{t-1} + \theta\hat{X}_{t-2,1} \\ &= -\theta X_{t-1} + \theta[-\theta X_{t-2} + \theta\hat{X}_{t-3,1}] \\ &= -\theta X_{t-1} - \theta^2 X_{t-2} - \theta^3 X_{t-3} - \cdots\end{aligned}$$

Alternatively, take the linear predictor as $\hat{X}_{t-1,1} = \sum_{r=1}^{\infty} a_r X_{t-r}$ and seek to minimize $\mathbb{E}[X_t - \hat{X}_{t-1,1}]^2$. We have

$$\begin{aligned}\mathbb{E}[X_t - \hat{X}_{t-1,1}]^2 &= \mathbb{E}\left[\epsilon_t - \theta\epsilon_{t-1} - \sum_{r=1}^{\infty} a_r(\epsilon_{t-r} - \theta\epsilon_{t-r-1})\right]^2 \\ &= \sigma^2\left[1 + (\theta + a_1)^2 + (\theta a_1 - a_2)^2 + (\theta a_2 - a_3)^2 + \cdots\right]\end{aligned}$$

Note that all terms, but the first, are minimized to 0 by taking $a_r = -\theta^r$.]

12. Consider the ARIMA$(0,2,2)$ model

$$(I-B)^2X = (I - 0.81B + 0.38B^2)\epsilon$$

where $\{\epsilon_t\}$ is white noise with variance 1.

(a) With data up to time T, calculate the k-step ahead optimal forecast of $\hat{X}_{T,k}$ for all $k \geq 1$. By giving a general formula relating $\hat{X}_{T,k}$, $k \geq 3$, to $\hat{X}_{T,1}$ and $\hat{X}_{T,2}$, determine the curve on which all these forecasts lie.

[The model is

$$X_t = 2X_{t-1} - X_{t-2} + \epsilon_t - 0.81\epsilon_{t-1} + 0.38\epsilon_{t-2}\,.$$

Hence

$$\begin{aligned}\hat{X}_{T,1} &= 2X_T - X_{T-1} + \hat{\epsilon}_{T+1} - 0.81\hat{\epsilon}_T + 0.38\hat{\epsilon}_{T-1}\\ &= 2X_T - X_{T-1} - 0.81[X_T - \hat{X}_{T-1,1}] + 0.38[X_{T-1} - \hat{X}_{T-2,1}]\end{aligned}$$

and similarly

$$\begin{aligned}\hat{X}_{T,2} &= 2\hat{X}_{T,1} - X_T + 0.38[X_T - \hat{X}_{T-1,1}]\\ \hat{X}_{T,k} &= 2\hat{X}_{T,k-1} - \hat{X}_{T,k-2}, \quad k \geq 3\,.\end{aligned}$$

This implies that the forecasts lie on a straight line.]

(b) Suppose now that $T = 95$. Calculate numerically the forecasts $\hat{X}_{95,k}$, $k = 1,2,3$ and their mean squared prediction errors when the last five observations are $X_{91} = 15.1$, $X_{92} = 15.8$, $X_{93} = 15.9$, $X_{94} = 15.2$, $X_{95} = 15.9$.

[You will need estimates for ϵ_{94} and ϵ_{95}. Start by assuming $\epsilon_{91} = \epsilon_{92} = 0$, then calculate $\hat{\epsilon}_{93} = \epsilon_{93} = X_{93} - \hat{X}_{92,1}$, and so on, until ϵ_{94} and ϵ_{95} are obtained.]

[Using the above formulae we obtain

t	X_t	$\hat{X}_{t,1}$	$\hat{X}_{t,2}$	$\hat{X}_{t,3}$	ϵ_t
91	15.1			0.000	0.000
92	15.8	16.500	17.200	17.900	0.000
93	15.9	16.486	16.844	17.202	−0.600
94	15.2	15.314	14.939	14.564	−1.286
95	15.9	15.636	15.596	15.555	0.586

Now

$$X_{T+k} = \sum_{r=0}^{\infty} c_r \epsilon_{T+k-r} \quad \text{and} \quad \hat{X}_{T,k} = \sum_{r=k}^{\infty} c_r \epsilon_{T+k-r}\,.$$

Thus

$$\mathbb{E}\left[X_{T+k}-\hat{X}_{T,k}\right]^2=\sigma_\epsilon^2\sum_{r=0}^{k-1}c_r^2\,.$$

where $\sigma_\epsilon^2=1$. Now

$$X_T=\epsilon_T+(2-0.81)\epsilon_{T-1}+(-2(0.81)-1+0.38)\epsilon_{T-2}+\cdots$$

Hence the mean square errors of $\hat{X}_{95,1}$, $\hat{X}_{95,2}$, $\hat{X}_{95,3}$ are respectively 1, 1.416, 5.018.]

13. Consider the state space model,

$$\begin{aligned} X_t &= S_t + v_t, \\ S_t &= S_{t-1} + w_t\,, \end{aligned}$$

where X_t and S_t are both scalars, X_t is observed, S_t is unobserved, and $\{v_t\}$, $\{w_t\}$ are Gaussian white noise sequences with variances V and W respectively. Write down the Kalman filtering equations for $\hat{S}_t$ and P_t.

Show that $P_t \equiv P$ (independently of t) if and only if $P^2 + PW = WV$, and show that in this case the Kalman filter for $\hat{S}_t$ is equivalent to exponential smoothing.

[This is the same as Section 8.4 of the notes. $F_t = 1$, $G_t = 1$, $V_t = V$, $W_t = W$. $R_t = P_{t-1} + W$. So if $(S_{t-1} \mid X_1, \ldots, X_{t-1}) \sim N\left(\hat{S}_{t-1}, P_{t-1}\right)$ then $(S_t \mid X_1, \ldots, X_t) \sim N\left(\hat{S}_t, P_t\right)$, where

$$\hat{S}_t = \hat{S}_{t-1} + R_t(V + R_t)^{-1}(X_t - \hat{S}_{t-1})$$

$$P_t = R_t - \frac{R_t^2}{V + R_t} = \frac{V R_t}{V + R_t} = \frac{V(P_{t-1} + W)}{V + P_{t-1} + W}\,.$$

P_t is constant if $P_t = P$, where P is the positive root of $P^2 + WP - WV = 0$.

In this case $\hat{S}_t$ behaves like $\hat{S}_t = (1 - \alpha)\sum_{r=0}^{\infty} \alpha^r X_{t-r}$, where $\alpha = V/(V + W + P)$. This is simple exponential smoothing.]