A note on updating beliefs when the number of observations increases in each period.

Christopher J. Ellis*

*Department of Economics, University of Oregon, Eugene, OR 97403

May 2001.

Abstract

This paper derives the updating rules for beliefs when each successive period has one more new observation than the previous one. There are many applications, for example learning the quality of the ships produced by a shipyard when a new vessel is launched in each period, or the merits of a plumber who repairs one plumbing system a day.

JEL Classification: C11.

I thank George Evans and Steve Haynes for their helpful comments.

1 Introduction.

In many circumstances in economics we model how an agent learns about some characteristic of an object by making repeated observations and then Bayes updating their prior distribution on the basis of this new information. Often the arrival of information will take a very particular form. One new observation is made in the first period, two new observations are made in the second, three in the third and so on. Examples would include the sale of a sequence of large durable goods such as ships from a shipyard, the privatization of a sequence of state run industries, or a sequence of services such as repairs performed by a plumber. In each case one observation is made on each unit of the object per period and the number of objects increases per period. The question we address is, if information arrives in the manner described above, how should an observer update their beliefs about some characteristic of the object if the signals that are received are noisy.

Obviously many variants of this type of problem have been addressed in the literature, but, surprisingly, we believe that the solution to this particular problem is not available elsewhere.

1.1 The Set-Up.

We assume that an economic agent wishes to learn about some characteristic of an object. The true characteristic of the object is $\bar{\theta}_t$, but the agent observes the noisy signal $\theta_t = \bar{\theta}_t + \varepsilon_t$ where $\varepsilon_t \sim N(0;1)\ \forall t$. In each period t the agent makes k simultaneous observations $\theta_t^k = \bar{\theta}_t + \varepsilon_t^k$, where the ε_t^k are also assumed $i{:}i{:}d$ across the k^0s. We also assume that the agents initial prior is distributed $\bar{\theta}_0 \sim N(\mu_0; \sigma_0^2)$, so from the perspective of the agent each θ_t^i is composed of the sum of normally distributed independent random variables and is thus also normally distributed. There are a total of I objects that will appear at the rate of 1 per period. This might simply reflect that there are a finite number of units of an object to be sold, or, it might be that each unit has a finite life and I represents the steady state stock of units.

The updating of expectations occurs as follows; at the beginning of any period, say t; the agent holds a prior distribution over the characteristic of interest $\bar{\theta} \sim N(\mu_t; \sigma_t^2)$. Given that beliefs are conditioned on all observations across all previously observed units of the object this implies that in periods $t \leq I$ when computing this distribution the agent has $t-1$ observations on the first unit sold, $t-2$ on the second, $t-3$ on the third and so on, giving at the end of period t a total number of observations of $n(t) = \frac{1}{2}t(t+1)$. In periods $t > I$ the total number of observations are given by $n(t) = \frac{1}{2}I(I+1) + (t-I)I$ where the term $(t-I)I$ reflects the idea that there is an upper bound on the number of units that may exist at one time. Now let the sale of the unit $k = t$ take place and the observations θ_t^k, $k = 1;:::;t$ of all the t units be revealed. The agent then calculates the posterior p.d.f. conditional on these observations. These beliefs are then used as the prior distribution of $\bar{\theta}_{t+1}$. Since both $\bar{\theta}_t$ and ε_t^k are normally distributed we may use the theory of linear projections to develop updating rules for the mean and variance of the agents beliefs.

1.2 Some Preliminaries.

Following Hamilton [1] pp 100-102 we know that if Y_1 is an $(n_1 \times 1)$ vector of normally distributed random variables with mean μ_1, and Y_2 is an $(n_2 \times 1)$ vector of normally distributed random variables with mean μ_2, where the variance covariance matrix is given by

$$\begin{bmatrix} E(Y_1-\mu_1)(Y_1-\mu_1)' & E(Y_1-\mu_1)(Y_2-\mu_2)' \\ E(Y_2-\mu_2)(Y_1-\mu_1)' & E(Y_2-\mu_2)(Y_2-\mu_2)' \end{bmatrix} = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$$

then

$$Y_2 | Y_1 \sim N([\mu_2 + \Omega_{21}\Omega_{11}^{-1}(y_1-\mu_1)];[\Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}])$$

In our problem $Y_1 = (\theta_1, ..., \theta_{n(t)})$, $Y_2 = (\bar{\theta}_1, ..., \bar{\theta}_{n(t)})$, Ω_{11} is an $(n(t) \times n(t))$ covariance matrix of the observations $\theta_1, ..., \theta_{n(t)}$, Ω_{12} is an $(1 \times n(t))$ covariance matrix between the observations $\theta_1, ..., \theta_{n(t)}$ and the variable $\bar{\theta}_t$, Ω_{21} is an $(n(t) \times 1)$ covariance matrix between the variable $\bar{\theta}_t$ and the observations $\theta_1, ..., \theta_{n(t)}$, Ω_{22} is the variance of $\bar{\theta}_t$ denoted σ_t^2.

To ease the calculations to follow we first compute Ω_{11} it's inverse Ω_{11}^{-1}, Ω_{12}, Ω_{21}, and Ω_{22}.

1. Ω_{11} and Ω_{11}^{-1}, where Ω_{11} is the variance covariance matrix of observations $(\theta_1, ..., \theta_{n(t)})$:

   The covariance is defined as

$$\begin{aligned} Cov(\theta_i, \theta_j) &= E(\theta_i - \mu)(\theta_j - \mu) \\ &= E(\bar{\theta}_i - \mu + \varepsilon_i)(\bar{\theta}_j - \mu + \varepsilon_j) \\ &= \begin{cases} \sigma_t^2 \text{ if } i \neq j \\ \sigma_t^2 + 1 \text{ if } i = j \text{ since the } \varepsilon_i\text{'s have unit variance.} \end{cases} \end{aligned}$$

thus

$$\Omega_{11} = \begin{bmatrix} \sigma_t^2 + 1, & \cdots & \sigma_t^2 \\ \vdots & \ddots & \vdots \\ \sigma_t^2, & \cdots & \sigma_t^2 + 1 \end{bmatrix}$$

to invert this matrix note first that it may be rewritten

$$\Omega_{11} = \sigma_t^2 H + I$$

where I is an $(n(t) \times n(t))$ identity matrix, and H is an $(n(t) \times n(t))$ matrix with each element unity.

To find Ω_{11}^{-1} apply the method of undetermined coefficients, guess

$$\Omega_{11}^{-1} = aH + bI$$

if correct this implies

$$\begin{aligned}(aH + bI)(\sigma_t^2 H + I) &= I \\ a\sigma_t^2 H^2 + b\sigma_t^2 H + aH + bI &= I\end{aligned}$$

using $H^2 = n(t)H$ it follows that the above holds if $b = 1$ and $(a\sigma_t^2 n(t) + b\sigma_t^2 + a)H = 0$, hence $a = -\frac{\sigma_t^2}{n(t)\sigma_t^2+1}$ and $b = 1$. Inserting these values back into the original guess gives the inverse

$$\begin{aligned}\Sigma_{11}^{-1} &= -\frac{\sigma_t^2}{n(t)\sigma_t^2 + 1}H + I \\ &= \begin{bmatrix} 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & , -\frac{\sigma_t^2}{n(t)\sigma_t^2+1} \\ \vdots & \vdots & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ -\frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & , 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1} \end{bmatrix}\end{aligned}$$

which is an $(n(t) \times n(t))$ matrix.

2 Σ_{12} is an $(1 \times n(t))$ covariance matrix between the observations $\Pi_1, \ldots, \Pi_{n(t)}$ and the variable $\overline{\Pi}_t$, the covariance is defined by

$$\begin{aligned}\mathrm{Cov}(\Pi_i, \overline{\Pi}_t) &= E(\Pi_i - \mu)(\overline{\Pi}_t - \mu) \\ &= E(\overline{\Pi}_t + \varepsilon_i - \mu)(\overline{\Pi}_t - \mu) \\ &= \sigma_t^2\end{aligned}$$

3 Σ_{21} is an $(n(t) \times 1)$ covariance matrix between the variable $\overline{\Pi}_t$ and the observations $\Pi_1, \ldots, \Pi_{n(t)}$, this is clearly just the transpose of Σ_{12}.

4 Σ_{22} is the variance of $\overline{\Pi}_t$ denoted σ_t^2.

1.3 The Updating Rule for the Mean (μ).

Consider first the mean, we have

$$\mu_{t+1} = \mu_t + \Sigma_{21}\Sigma_{11}^{-1}(\Pi_{n(t)} - \mu_t)$$

where $\Pi_{n(t)} - \mu_t$ is an $(n(t) \times 1)$ vector of data expressed as deviations from the current mean. Hence

$$\mu_{t+1} = \mu_t + \left[\sigma_t^2, \ldots\ldots, \sigma_t^2\right] \begin{bmatrix} 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & , -\frac{\sigma_t^2}{n(t)\sigma_t^2+1} \\ \vdots & \vdots & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ -\frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & , 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1} \end{bmatrix} \begin{bmatrix} \Pi_1 - \mu_t \\ \vdots \\ \vdots \\ \vdots \\ \Pi_{n(t)} - \mu_t \end{bmatrix}$$

hence

$$\mu_{t+1} = \mu_t + \left[\sigma_t^2, \ldots, \sigma_t^2\right] \begin{bmatrix} 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & ; -\frac{\sigma_t^2}{n(t)\sigma_t^2+1} \\ \vdots & \ddots & \vdots \\ -\frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & ; 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1} \end{bmatrix} \begin{bmatrix} \theta_1 + \varepsilon_1^1 - \mu_t \\ \theta_2 + \varepsilon_2^1 - \mu_t \\ \theta_2 + \varepsilon_2^2 - \mu_t \\ \vdots \\ \theta_t + \varepsilon_t^t - \mu_t \end{bmatrix}$$

computing the matrix multiplication yields

$$\begin{aligned} \mu_{t+1} = {} & \mu_t + \sigma_t^2 \left(1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}\right)\left(\theta_1 + \varepsilon_1^1 - \mu_t\right) - \cdots - \left(\frac{(\sigma_t^2)^2}{n(t)\sigma_t^2+1}\right)\left(\theta_t + \varepsilon_t^t - \mu_t\right) \\ & - \left(\frac{(\sigma_t^2)^2}{n(t)\sigma_t^2+1}\right)\left(\theta_1 + \varepsilon_1^1 - \mu_t\right) + \left(\frac{(\sigma_t^2)^2}{n(t)\sigma_t^2+1}\right)\left(\theta_2 + \varepsilon_1^2 - \mu_t\right) - \cdots \\ & - \left(\frac{(\sigma_t^2)^2}{n(t)\sigma_t^2+1}\right)\left(\theta_t + \varepsilon_t^t - \mu_t\right) - \cdots + \sigma_t^2 \left(1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}\right)\left(\theta_t + \varepsilon_t^t - \mu_t\right) \end{aligned}$$

gathering terms this yields the required expression

$$\mu_{t+1} = \left(\frac{1}{n(t)\sigma_t^2+1}\right)\mu_t + \left(\frac{\sigma_t^2}{n(t)\sigma_t^2+1}\right)\left(\sum_{j=1} j\left(\theta_j\right) + \sum_{k=1} \varepsilon_j^k\right) \tag{1}$$

1.4 The Updating Rule for the Variance (σ_t^2).

Next consider the variance, we have

$$\sigma_{t+1}^2 = \sigma_t^2 - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

which may be written

$$\sigma_{t+1}^2 = \sigma_t^2 - \left[\sigma_t^2, \ldots, \sigma_t^2\right] \begin{bmatrix} 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & ; -\frac{\sigma_t^2}{n(t)\sigma_t^2+1} \\ \vdots & \ddots & \vdots \\ -\frac{\sigma_t^2}{n(t)\sigma_t^2+1}, & \ldots & ; 1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1} \end{bmatrix} \begin{bmatrix} \sigma_t^2 \\ \vdots \\ \sigma_t^2 \end{bmatrix}$$

carrying out the matrix multiplication yields

$$\sigma_{t+1}^2 = \sigma_t^2 - n(t)(\sigma_t^2)^2 \left(1 - \frac{\sigma_t^2}{n(t)\sigma_t^2+1}\right) - n(t)(n(t)-1)\left(\frac{(\sigma_t^2)^3}{n(t)\sigma_t^2+1}\right)$$

simple algebra now reduces this expression to the updating rule for the variance.

$$\sigma_{t+1}^2 = \frac{\sigma_t^2}{n(t)\sigma_t^2+1} \tag{2}$$

where $n(t) = \frac{1}{2}t(t+1)$ for periods $t \leq l$ and $n(t) = \frac{1}{2}l(l+1) + (t-l)l$ for periods $t > l$

1.5 Properties of the Solution.

The updating rules (1) and (2) constitute a system of first order non-linear difference equations the solution to which has the following properties

1.5.1 Properties of the Variance.

1. In the steady state $\sigma_t^2 = 0$.
2. The steady state is unique.
3. $\sigma_t^2 \to 0$ monotonically as t increases.

1.5.2 Properties of the Mean.

Lagging expression (1) backwards and substituting the result back into the original expression t times gives

$$\mu_{t+1} = \prod_{j=0}^{t} \theta_j \mu_0 + \theta_t \sigma_t^2 \left(\sum_{j=1}^{t} \left(j \left(\Pi_j\right) + \sum_{k=1}^{j} \varepsilon_j^k \right) \right) + \cdots$$
$$+ (\theta_t \times \cdots \times \theta_1)\, \sigma_1^2 \left(\sum_{j=1}^{1} \left(j \left(\Pi_j\right) + \sum_{k=1}^{j} \varepsilon_j^k \right) \right) \tag{3}$$

where $\theta_j \equiv \frac{1}{n(j)\sigma_j^2+1}$. Notice that (2) and the definition of $\theta_j \equiv \frac{1}{n(j)\sigma_j^2+1}$ imply that $\theta_j = \frac{\sigma_{j+1}^2}{\sigma_j^2}$ thus since $\theta_j\theta_{j+1} = \frac{\sigma_{j+1}^2}{\sigma_j^2}\frac{\sigma_{j+2}^2}{\sigma_{j+1}^2} = \frac{\sigma_{j+2}^2}{\sigma_j^2}$ similarly $\theta_j\theta_{j+1}\theta_{j+2} = \frac{\sigma_{j+3}^2}{\sigma_j^2}$ and so on, we may simplify (3) as

$$\mu_{t+1} = \frac{\sigma_{t+1}^2}{\sigma_0^2}\mu_0 + \sigma_{t+1}^2 \left(\sum_{j=1}^{t} \left(j \left(\Pi_j\right) + \sum_{k=1}^{j} \varepsilon_j^k \right) \right) + \cdots$$
$$+\sigma_{t+1}^2 \left(\sum_{j=1}^{1} \left(j \left(\Pi_j\right) + \sum_{k=1}^{j} \varepsilon_j^k \right) \right)$$
$$= \frac{\sigma_{t+1}^2}{\sigma_0^2}\mu_0 + \sigma_{t+1}^2 \sum_{i=1}^{t} \left(\sum_{j=1}^{i} \left(j \left(\Pi_j\right) + \sum_{k=1}^{j} \varepsilon_j^k \right) \right) \tag{4}$$

Hence this expression gives the agents expectation of the characteristic at $t+1$ in terms of the initial prior distribution $\Pi_0 \sim N(\mu_0; \sigma_0^2)$, the exogenous process that governs the evolution of σ_t^2, and the history of prior observations. From this expression we note that the effect of a change in Π_t on the expectation for the subsequent period $t+i$ may be written

$$\frac{d\mu_{t+i}}{d\Pi_t} = it(\sigma_{t+i}^2)$$

this simply says that changing the observation in period t will e¤ect that periods t observations, and that these will appear as data in the updating process i times from t to t + i.

References

[1] Hamilton, J. D., "Time Series Analysis," Princeton University Press, Princeton, New Jersey, 1994.