Single models

Unbalanced coin

We start our explanation of how to proceed with the model of a coin.

Let $y_t \in \{1, 2\}$ be a binary variable (describing an experiment with a coin: $y_t = 1$ denotes head and $y_t = 2$ denotes tail) and we want to estimate the probability of $y_t = 1$ (head).

The estimation is evident even without any theory. We perform experiments and count number of realized heads and tails. (Number of heads + tails gives number of experiments performed.) The estimated probability will be the number of heads divided the number of experiments (this follows directly from the statistical definition of probability). With no prior information, we start counting from zeros.

If we denote n_1 is number of heads and n_2 that of tails then the probability of head P_h is

$$P_h = \frac{n_1}{n_1 + n_2}$$

and at the beginning it is $n_1 = n_2 = 0$

Now, if we believe that the head is much more probable we ca formulate this knowledge like this: we performed preliminary 10 experiments and obtained 9 times head and only 1 tail. So, we can set $n_1 = 9$ and $n_2 = 1$ at the beginning. If our belief is not correct then we can say that one or two results of out experiment does not make much change and only say 10 new results can compete with our imbedded knowledge. If our belief is even stronger, we can start with $n_1 = 90$ and $n_2 = 10$ which is $10\,[9,\ 1]$.

As we can see, we can formulate our prior knowledge as $k\,[n_1, n_2]$ where $[n_1, n_2]$ defines our knowledge about our the probability and k expresses the strength of our belief.

This idea can be a base for our approach to setting the prior knowledge to all other models.

Models with summation statistics

Parameters of many models are estimated as sample averages. Then the statistics are sum S_t and counter κ_t with the update

$$S_t = S_{t-1} + y_t$$
$$\kappa_t = \kappa_{t-1} + 1$$

with the point estimate

$$\hat{p}_t = \frac{S_t}{\kappa_t}$$

Now, if we have some knowledge about the estimate of the parameter e.g. $\hat{p} = \hat{p}_0$. Then we can select the strength by the value of κ_0 and then the statistics S_0 is given as

$$S_0 = \kappa_0 \hat{p}_0.$$

Remark

For other cases, e.g. if $\hat{p}_t = \frac{\kappa_t}{S_t}$ we can proceed analogically.

Estimation of normal model

Here the situation is more complicated, however, the basic procedure is still the same. The statistics are

$$V_t = V_{t-1} + \Psi_t \Psi_t^{'}$$
$$\kappa_t = \kappa_t + 1$$

and the point estimate is

$$\hat{\theta} = V_p^{-1} V_{yp}, \text{ with } V = \begin{bmatrix} V_y & V_{yp}^{'} \\ V_{yp} & V_p \end{bmatrix}$$

If V_p is a unit matrix, then $\hat{\theta} = V_{yp}$. If we want to introduce $\hat{\theta} = \hat{\theta}_0$ we can set

$$V_0 = \begin{bmatrix} 1 & \hat{\theta}_0^{'} \\ \hat{\theta}_0 & I \end{bmatrix}$$

where I is a unit matrix.

Now, if we do k steps with the same data, we have

$$V_0 = k \Psi_t \Psi_t^{'}$$
$$\kappa_0 = k$$

That is, the final setting with the strength k is

$$V_0 = k \begin{bmatrix} 1 & \hat{\theta}_0' \\ \hat{\theta}_0 & I \end{bmatrix}$$

$$\kappa_0 = k$$

Example

For regression model $y_t = .4y_t + 1.2u_t + e_t$ and the strength $k = 10$ we set

$$V_0 = 10 \begin{bmatrix} 1 & 0.4 & 1.2 \\ 0.4 & 1 & 0 \\ 1,2 & 0 & 1 \end{bmatrix}$$

$$\kappa_0 = 10$$

Then $V_p = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$, $V_{yp} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}$ and $\hat{\theta} = \frac{1}{10} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 12 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 1.2 \end{bmatrix}$ which is just the prior estimate we wanted. The strength shows itself in the covariance matrix of the estimate $C\left(\hat{\theta}_t\right) = \frac{\hat{r}_t}{\kappa_t}$ which is 10 times smaller.