1 Chapter 6 Learning as Inference

Problem 6.2 ***We build on the case study of inferring a Bernoulli probability in Section 6.2.***

(a) Prove equation (6.22) for the PME, starting from equation (6.14).

(b) Generalize equation (6.22) to the case of a beta prior with parameters α_0 and β_0.

(a)
To answer the question, I will continue to use the example of rain and dry in section 6.2 for the interest of convenience.
The PME for Bernoulli parameter r after observing t binary outcomes is given by:

$$\hat{r}_t = \frac{n_{rain} + 1}{t + 2} \tag{1}$$

Let $x_t \in \{0, 1\}$ be the observation on day t, with 1 indicating rain and 0 indicating dry. Then:

$$n_{rain,t} = n_{rain,t-1} + x_t \quad and \quad t = t - 1 + 1$$

Substituting:

$$\hat{r}_t = \frac{n_{rain,t-1} + x_t + 1}{(t-1) + 1 + 2} = \frac{n_{rain,t-1} + 1}{t + 1} + \frac{x_t - \hat{r}_{t-1}}{t + 2}$$

then we can get,

$$\hat{r}_t = \hat{r}_{t-1} + \frac{1}{t+2}(x_t - \hat{r}_{t-1}) \tag{2}$$

And this is the reinforcement learning update rule.

(b)
For a Beta prior $Beta(\alpha_0, \beta_0)$, the posterior mean after observing n_{rain} rain and n_{dry} dry days would be:

$$\hat{r}_t = \frac{\alpha_0 + n_{rain}}{\alpha_0 + \beta_0 + t} \tag{3}$$

Now we can define that $\hat{r}_{t-1} = \frac{\alpha_0 + n_{rain,t-1}}{\alpha_0 + \beta_0 + t - 1}$, and get that:

$$n_{rain,t} = n_{rain,t-1} + x_t \quad and \quad \hat{r}_t = \frac{\alpha_0 + n_{rain,t-1} + x_t}{\alpha_0 + \beta_0 + t}$$

In the generalized form:

$$\hat{r}_t = \hat{r}_{t-1} + \frac{1}{\alpha_0 + \beta_0 + t}(x_t - \hat{r}_{t-1})$$

This is the update equation for the PME under the $Beta(\alpha_0, \beta_0)$ prior.