Exercise sheet 4

Exercise 1. Consider the first-order polynomial trend model with state $\boldsymbol{\theta}_k$ and transition noise covariance matrix U_k at time step k and assume that the posterior distribution of $\boldsymbol{\theta}_{k-1}$ given $\boldsymbol{y}_{0:k-1} = y_{0:k-1}$ is normal with mean $\hat{m}_{k-1}$ and variance $\hat{P}_{k-1}$. We want to use an intervention at time step k to obtain a predicted distribution for $\boldsymbol{\theta}_k$ at time step k of the form

$$\mathrm{N}(\cdot; \hat{m}_{k-1}, P_k + \tilde{U}_k)$$

with $P_k = \hat{P}_{k-1} + U_k$ and with $\tilde{U}_k$ some extra variance.

i) Define the distribution of the noise $\boldsymbol{u}_k$ at time step k that encompasses all the uncertainty in the prediction.

ii) Show that an alternative approach is to modify the state equation at time step k to

$$\boldsymbol{\theta}_k = \breve{F}_k \boldsymbol{\theta}_{k-1} + c_k + \breve{\boldsymbol{u}}_k, \qquad \breve{\boldsymbol{u}}_k \sim \mathrm{N}(\cdot\,; 0, \breve{U}_k).$$

with

$$\breve{F}_k = \sqrt{1 + \frac{\tilde{U}_k}{P_k}}, \qquad c_k = (1 - \breve{F}_k)\hat{m}_{k-1} \qquad \text{and} \qquad \breve{U}_k = \Big(1 + \frac{\tilde{U}_k}{P_k}\Big) U_k.$$

Can you think of a situation where simply adding an independent error as in (i) is not going to work, but we can still use the change of model considered in (ii)? Briefly discuss in the context of a first-order polynomial trend model.

Exercise 2. You want to learn the probability for a (possibly unfair) die to show each of its $K = 6$ faces, the probability of a given face $i \in \{1, \dots, K\}$ being defined as a random variable $\boldsymbol{\theta}_i$. The observation process $\boldsymbol{y}_n$ for the n-th observation, $n \geq 1$, is to roll the die and observe the outcome $y_n \in \{1, \dots, K\}$, which is conditionally independent of all other informations given the parameter $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K)$. In this case, the likelihood is a categorical distribution

$$p_{\boldsymbol{y}_n | \boldsymbol{\theta}_{1:K}}(y_n \,|\, \theta_{1:K}) = \theta_{y_n}$$

and a conjugate prior is the Dirichlet distribution

$$p_{\boldsymbol{\theta}_{1:K}}(\theta_{1:K}) = \frac{\Gamma\big(\sum_{i=1}^K \alpha_i\big)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K \theta_k^{\alpha_k - 1},$$

with parameters $\alpha_1, \dots, \alpha_K$ defined as positive real numbers. The Dirichlet distribution is defined on the set

$$\mathrm{S}_{K-1} = \Big\{\theta_1, \dots, \theta_K \in (0,1) : \sum_{i=1}^K \theta_i = 1\Big\}$$

of all probability mass functions on $\{1, \dots, K\}$ as required.

i) Express the posterior distribution $p_{\boldsymbol{\theta}_{1:K} | \boldsymbol{y}_{1:n}}(\cdot \,|\, y_{1:n})$ as a function of the parameters $\alpha_1, \dots, \alpha_K$ and of the total number of times m_i the face $i \in \{1, \dots, K\}$ has been observed.

ii) Express the marginal likelihood $p_{\boldsymbol{y}_{1:n}}(y_{1:n})$ as a function of the parameters $\alpha_1, \dots, \alpha_K$ and of $m_1, \dots, m_K$.

iii) You have been given some prior information which you have taken into account by specifying some given values $\alpha_1, \dots, \alpha_K$ of the parameters. Given that the case $\alpha_1 = \cdots = \alpha_K = 1$ corresponds to a uniform prior, how would you assess the evidence in favour of/against the validity of the prior information you have received?