

# HW 7 Shuang Gao

## Question 1

we can put a prior on the parameters  $\omega = (\omega_1, \dots, \omega_L)$

- A convenient choice is the Dirichlet distribution  $\omega \sim Di_L(\alpha)$  conjugate prior
- The posterior probability for  $\omega$   $\omega \sim Di_L(\alpha + N\hat{\omega})$

As  $\alpha \rightarrow 1$ ,  $\omega \sim Di_L(N\hat{\omega} + 1)$

$\hat{\omega}$  sample

Exercise: verify that the multinomial distribution and the Dirichlet distribution have approximately the mean and variances

HW 1

posterior: $p(\omega   N\hat{\omega}) \propto p(N\hat{\omega}   \omega) \pi(\omega)$ $= \omega_1^{N\hat{\omega}_1 + \alpha - 1} \cdots \omega_L^{N\hat{\omega}_L + \alpha - 1}$	posterior: $p(\omega   N\hat{\omega}) \propto p(N\hat{\omega}   \omega) \pi(\omega)$ $= \omega_1^{N\hat{\omega}_1 + \alpha - 1} \cdots \omega_L^{N\hat{\omega}_L + \alpha - 1}$
---	---

$\hat{w}_j$  the observed proportion in category j

$w_j$  the probability that a sample point falls in category j

① prior  $\omega \sim Di_L(\alpha)$

data  $N\hat{\omega} \sim \text{Multinomial}(N, \omega)$

$\omega | N\hat{\omega} \propto P(N\hat{\omega} | \omega) \pi(\omega)$

$$= \frac{N!}{(N\hat{\omega}_1)! \cdots (N\hat{\omega}_L)!} \omega_1^{N\hat{\omega}_1} \cdots \omega_L^{N\hat{\omega}_L} \frac{\prod_{i=1}^L \Gamma(\alpha_i)}{\prod_{i=1}^L \Gamma(\alpha_i)} \prod_{i=1}^L \omega_i^{\alpha_i - 1}$$

$$\propto \omega_1^{N\hat{\omega}_1 + \alpha - 1} \cdots \omega_L^{N\hat{\omega}_L + \alpha - 1}$$

thus  $\omega | N\hat{\omega} \sim \text{Dirichlet}_L(N\hat{\omega} + \alpha)$

let  $\alpha \rightarrow 0$  (flat prior),  $\omega | N\hat{\omega} \sim \text{Dirichlet}_L(N\hat{\omega})$

② bootstrap dist

$N\hat{\omega}^* \sim \text{Multi}(N, \hat{\omega})$

Comparing ① & ②,

①  $\omega | N\hat{\omega} \sim \text{Dir}_L(N\hat{\omega})$

$$\text{mean } E(\omega_i) = \frac{N\hat{\omega}_i}{\sum_{i=1}^L N\hat{\omega}_i} = \hat{w}_i$$

$$\begin{aligned} \text{cov}(\omega_i, \omega_j) &= \frac{-N\hat{\omega}_i N\hat{\omega}_j}{\left(\sum_{i=1}^L N\hat{\omega}_i\right)^2 \left(\sum_{i=1}^L N\hat{\omega}_i + 1\right)} \\ &= \frac{-N^2 \hat{w}_i \hat{w}_j}{N^2 (N+1)} \\ &= -\frac{\hat{w}_i \hat{w}_j}{N+1} \end{aligned}$$

②  $N\hat{\omega}^* \sim \text{Multi}(N, \hat{\omega})$

$$\begin{aligned} \text{mean } E(N\hat{\omega}_i^*) &= N\hat{\omega}_i \\ \Rightarrow E(\hat{w}_i^*) &= \hat{w}_i \end{aligned}$$

$$\begin{aligned} \text{cov}(N\hat{\omega}_i^*, N\hat{\omega}_j^*) &= -N\hat{\omega}_i \hat{\omega}_j = N^2 \text{cov}(\hat{w}_i^*, \hat{w}_j^*) \\ \Rightarrow \text{cov}(\hat{w}_i^*, \hat{w}_j^*) &= -\frac{\hat{w}_i \hat{w}_j}{N} \end{aligned}$$

Conclusion: 2 distributions have same mean and approximately same cov.

Question 2:

$\beta_1, \beta_4, \beta_5$  and  $\beta_6$

From simple algebra, it can be shown that

$$(*) \quad f(X) = (\beta_1 + \beta_4 X) + \beta_5(X - \xi_1)_+ + \beta_6(X - \xi_2)_+ \quad (\text{Hw.})$$

• So  $f(X)$  can be expressed by 4 basis functions:

•  $\tilde{h}_1(X) = 1, \tilde{h}_2(X) = X, \tilde{h}_3(X) = (X - \xi_1)_+$  and  $\tilde{h}_4(X) = (X - \xi_2)_+$

$$\begin{array}{c} (\tilde{h}_1(X))_+ \\ \uparrow \\ \text{Graph of } (X - \xi_1)_+ \end{array} \quad \begin{array}{c} (\tilde{h}_2(X))_+ \\ \uparrow \\ \text{Graph of } (X - \xi_2)_+ \end{array} = \begin{cases} X - \xi_1 & \text{if } X \geq \xi_1 \\ 0 & \text{if } X < \xi_1 \end{cases}$$

original form

$$f(X) = (\beta_1 + \beta_4 X) I(X < \xi_1) + (\beta_2 + \beta_5 X) I(\xi_1 \leq X < \xi_2) + (\beta_3 + \beta_6 X) I(X \geq \xi_2)$$

with 2 boundary condition

$$\beta_2 = \beta_1 - (\beta_5 - \beta_4) \xi_1$$

$$\beta_3 = \beta_1 - (\beta_5 - \beta_4) \xi_1 - (\beta_6 - \beta_5) \xi_2$$

$$f(X) = (\beta_1 + \beta_4 X) I(X < \xi_1) + (\beta_2 + \beta_5 X) I(\xi_1 \leq X < \xi_2) + (\beta_3 + \beta_6 X) I(X \geq \xi_2)$$

$$= (\beta_1 + \beta_4 X) I(X < \xi_1) + [\beta_1 - (\beta_5 - \beta_4) \xi_1 + \beta_5 X] I(\xi_1 \leq X < \xi_2)$$

$$+ [\beta_1 - (\beta_5 - \beta_4) \xi_1 - (\beta_6 - \beta_5) \xi_2 + \beta_6 X] I(X \geq \xi_2)$$

$$\textcircled{1} \quad X < \xi_1, \quad f(X) = \beta_1 + \beta_4 X$$

$$\textcircled{2} \quad \xi_1 \leq X < \xi_2, \quad f(X) = \beta_1 - (\beta_5 - \beta_4) \xi_1 + \beta_5 X \\ = \beta_1 + \beta_4 \xi_1 + \beta_5 (X - \xi_1)$$

$$\textcircled{3} \quad X \geq \xi_2, \quad f(X) = \beta_1 - (\beta_5 - \beta_4) \xi_1 - (\beta_6 - \beta_5) \xi_2 + \beta_6 X \\ = \beta_1 + \beta_4 \xi_1 + \beta_5 (\xi_2 - \xi_1) + \beta_6 (X - \xi_2)$$

From above 3 conditions, we can see 4 basis shown in the final

form are  $h_1(X) = 1 \quad h_2(X) = X$

$$h_3(X) = (X - \xi_1)_+$$

$$h_4(X) = (X - \xi_2)_+$$