LECTURE 9 - BAYESIAN INFERENCE FOR NORMAL DISTRIBUTION

DATA $x_1, \dots, x_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\boxed{\text{CASE 1}}$ μ UNKNOWN, σ^2 KNOWN

- PRIOR $\mu \sim N(\phi, \tau^2)$
- LIKELIHOOD $f(\underline{x} \mid \mu) \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{\Sigma(x_i - \mu)^2}{2\sigma^2} \right\}$
- POSTERIOR $\pi(\mu \mid x) \propto p(\mu) f(\underline{x} \mid \mu)$

$$\vdots$$

$$\mu \mid x \sim N\left(\frac{\tau^2 \Sigma x_i + \sigma^2 \phi}{n\tau^2 + \sigma^2} ; \frac{\sigma^2 \tau^2}{n\tau^2 + \sigma^2} \right)$$

NOTE

- if $\tau^2 \to 0$ (INFORMATIVE PRIOR) $\Rightarrow E_\pi(\mu) = \phi$, $VAR_\pi(\mu) \to 0$
- if $\tau^2 \to \infty$ (UNINFORMATIVE PRIOR) $\Rightarrow E_\pi(\mu) = \bar{x}$; $VAR_\pi(\mu) = \frac{\sigma^2}{n}$

CASE 2 — μ KNOWN, σ^2 UNKNOWN

- PRIOR $\sigma^2 \sim \Gamma^{-1}(\alpha, \beta)$; $p(\sigma^2) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{ -\dfrac{\beta}{\sigma^2} \right\}$
- LIKELIHOOD THE SAME AS CASE 1
- POSTERIOR $\pi(\sigma^2 | \underline{x}) \propto f(\underline{x} | \sigma^2)\, p(\sigma^2)$

$$\propto (\sigma^2)^{-n/2} \exp\left\{ -\frac{\sum (x_i - \mu)^2}{2\sigma^2} \right\} (\sigma^2)^{-\alpha-1} \exp\left\{ -\frac{\beta}{\sigma^2} \right\}$$

$$= (\sigma^2)^{-\left(\frac{n}{2}+\alpha+1\right)} \exp\left\{ -\frac{1}{\sigma^2} \left[\frac{\sum (x_i - \mu)^2}{2} + \beta \right] \right\}$$

$$\sigma^2 | x \sim \Gamma^{-1}\left(\frac{n}{2} + \alpha \, ; \, \frac{\sum (x_i - \mu)^2}{2} + \beta \right)$$

$\boxed{\text{CASE 3}}$ μ UNKNOWN, σ^2 UNKNOWN

- BAYES THEOREM (MULTI PARAMETER CASE)

$\underline{\theta} = \{\theta_1, \dots, \theta_n\}$; DATA $\underline{x}$

$$\Rightarrow \pi(\underline{\theta} \mid \underline{x}) = \frac{f(\underline{x} \mid \underline{\theta})\, p(\underline{\theta})}{f(\underline{x})}$$

NOTE

- $f(\underline{x}) = \int_{\underline{\theta}} f(\underline{x} \mid \underline{\theta})\, p(\underline{\theta})\, d\underline{\theta}$ PROBLEMATIC !
- $p(\underline{\theta}) = p(\theta_1, \dots, \theta_n) \rightarrow p(\underline{\theta}) = \prod_{i=1}^{n} p(\theta_i)$
- $p(\underline{\theta} \mid \phi) = \prod_{i=1}^{n} p(\theta_i \mid \phi)$ WHERE $p(\phi) \sim f(\phi)$
- $\pi(\underline{\theta} \mid \underline{x}) \propto f(\underline{x} \mid \underline{\theta})\, p(\underline{\theta})$ THEN $\pi(\theta_1 \dots \theta_n \mid \underline{x})$ HOW TO SUMMARISE THIS POSTERIOR ?

⇒ WE WOULD LIKE TO LOOK AT MARGINAL POSTERIORS

$$\pi(\theta_1 \mid \underline{x}) = \int \pi(\underline{\theta} \mid \underline{x})\, d\theta_2 \sim d\theta_n$$

OFTEN TOO COMPLEX TO COMPUTE!

$\boxed{\text{CASE 3}}$ μ UNKNOWN, σ^2 UNKNOWN

- PRIOR $p(\mu, \sigma^2) = p(\mu) \cdot p(\sigma^2)$ $\begin{cases} \mu \sim N(\phi, \tau^2) \\ \sigma^2 \sim \Gamma^{-1}(\alpha, \beta) \end{cases}$

- POSTERIOR $\pi(\mu, \sigma^2 \mid \underline{x}) \propto f(\underline{x} \mid \mu, \sigma^2)\, p(\mu)\, p(\sigma^2)$

$$= \left[\prod^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}\right] (2\pi\tau^2)^{-1/2} \exp\left\{-\frac{(\mu - \phi)^2}{2\tau^2}\right\} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left\{\frac{-\beta}{\sigma^2}\right\} \cdot$$

$$\propto (\sigma^2)^{-\left(\frac{n}{2}+\alpha+1\right)} \exp\left\{-\left[\frac{\sum(x_i-\mu)^2}{2\sigma^2} + \frac{(\mu-\phi)^2}{2\tau^2} + \frac{\beta}{\sigma^2}\right]\right\}$$

→ MARGINALS $\begin{cases} \cdot\ \pi(\mu \mid \underline{x}) \propto \int_0^\infty \pi(\mu, \sigma^2 \mid \underline{x})\, d\sigma^2 \\ \cdot\ \pi(\sigma^2 \mid \underline{x}) \propto \int_{-\infty}^{+\infty} \pi(\mu, \sigma^2 \mid \underline{x})\, d\mu \end{cases}$

LET'S TRY SOMETHING ELSE!

$\Rightarrow$ <u>CONDITIONAL</u> DISTRIBUTIONS

$$\pi(\mu \mid \underline{x}, \sigma^2) \propto \exp\left\{-\frac{1}{2}\left[\frac{\sum(x_i-\mu)^2}{\sigma^2} + \frac{(\mu-\phi)^2}{\tau^2}\right]\right\} \sim$$

IT WILL TAKE US BACK TO CASE 1! $N(\cdot, \cdot)$

$$\pi(\sigma^2 \mid \underline{x}, \mu) \propto \sim$$

BACK TO CASE 2! $\Gamma^{-1}(\cdot, \cdot)$

$\rightarrow$ WHAT IF WE SET UP AN ITERATIVE PROCESS WHICH SIMULATES FROM THESE CONDITIONAL DISTRIBUTIONS IN TURN?