Chapter 3: Facts about Common Statistical Models *

3 Facts about common statistical models

3.1 Bayes Models

- Probability Model on Data We have distributions $\mathcal{P} \equiv \{P_\theta : \theta \in \Theta\}$ for X on $(\mathcal{X}, \mathcal{B})$, where $\mathcal{P} \ll \mu$ ($\sigma-$ finite measure) and R -N derivatives

$$\frac{dP_\theta}{d\mu}(x) = f_\theta(x)$$

- Prior on Parameter We now add an assumption of a distribution G on $(\Theta, \mathcal{C})$ with $G \ll \nu(\sigma$ -finite measure) and R-N derivative

$$\frac{dG}{d\nu}(\theta) = g(\theta)$$

- Joint Distribution for (X, θ) : Here we consider $f_\theta(x)$ as a function of both x and θ (i.e., measurable in (x, θ)). If $f_\theta(x)$ is $\mathcal{B} \times \mathcal{C}$ -measurable, then there exists a joint probability distribution for (X, θ) on $(\mathcal{X} \times \Theta, \mathcal{B} \times \mathcal{C})$ defined, for $A \in \mathcal{B} \times \mathcal{C}$, by

$$\pi^{X,\theta}(A) \equiv P((X, \theta) \in A) = \int_A f_\theta(x) d(\mu \times G)(x, \theta) = \int f_\theta(x) g(\theta) d(\mu \times \nu)(x, \theta)$$

where

$$\frac{d\pi^{X,\theta}}{d(\mu \times G)} \equiv f_\theta(x), \quad \frac{d\pi^{X,\theta}}{d(\mu \times \nu)} \equiv f_\theta(x) g(\theta)$$

- Marginal Distributions
  - for X $(B \in \mathcal{B})$

$$\pi^X(B) \equiv P(X \in B) = \pi^{X,\theta}(B \times \Theta) = \int_{B \times \Theta} f_\theta(x) d(\mu \times G)(x, \theta) \overset{Fubini}{=} \int_B \left[\int_\Theta f_\theta(x) dG(\theta) \right] d\mu(x)$$

$$= \int_B \left[\int_\Theta f_\theta(x) g(\theta) d\nu \right] d\mu(x)$$

$$0 \leq \frac{d\pi^X(x)}{d\mu} = \int_\Theta f_\theta(x) dG(\theta) = \int_\Theta f_\theta(x) g(\theta)$$

  - for θ $(C \in \mathcal{C})$

$$\pi^\theta(C) \equiv P(\theta \in C) = \pi^{X,\theta}(\mathcal{X} \times C) = \int_{\mathcal{X} \times C} f_\theta(x) d(\mu \times G)(x, \theta) = \int_C \left[\int_{\mathcal{X}} f_\theta(x) d\mu(x) \right] dG(\theta) = G(C)$$

  Marginal distribution of θ is prior distribution G.

- Conditional distributions
  - for $X \mid \theta$

$$\pi^{X|\theta}(B \mid \theta) \equiv P_{X|\theta}(X \in B \mid \theta) = \int_B f_\theta(x) d\mu(x) = P_\theta(B), \quad B \in \mathcal{B}$$

$$\frac{d\pi^{X|\theta}(x)}{d\mu} = \frac{dP_\theta(x)}{d\mu} = f_\theta$$

*STA643: Advanced Theory of Statistical Inference. Instructed by Dr. Daniel Nordman. Arranged by Zhiling Gu

- for $\theta \mid X$

$$\pi^{\theta|X}(C \mid x) \equiv P_{\theta|X}(\theta \in C \mid X = x) = \int_C \left[\frac{f_\theta(x)}{\int_\Theta f_\theta(x) dG(\theta)}\right] dG(\theta) = \int_C \frac{f_\theta(x) g(\theta)}{\int_\Theta f_\theta(x) g(\theta) d\nu(\theta)} d\nu(\theta)$$

$$\frac{d\pi^{\theta|X}(\theta)}{dG} = \frac{f_\theta(x) g(\theta)}{\int_\Theta f_\theta(x) dG(\theta)}, \quad \frac{d\pi^{\theta|X}(\theta)}{d\nu} = \frac{f_\theta(x) g(\theta)}{\int_\Theta f_\theta(x) g(\theta) d\nu(\theta)}, G \ll \nu$$

Note priors does not necessarily have a density, i.e. G is not necessarily dominated by some ν. But you can always write the density of posterior with respect to G.