Chapter 7. Bayesian model:

① review of conditional distribution:

If (X, Y) have joint distribution $f(x, y)$, and X has marginal distribution $f_X(x)$. Then the conditional distribution of Y given $X = x$ is

$$f(y|x) = \frac{f(x,y)}{f_X(x)}.$$

<u>Eg1.</u> If the joint pdf of (X, Y) is

$$f(x,y) = \begin{cases} 4xy. & 0 \le x, y \le 1. \\ 0. & o/w. \end{cases}$$

Find $f(y|x)$ for any $0 \le x \le 1$.

Sol: First find the marginal pdf of X. For $0 \le x \le 1$.

$$f_X(x) = \int_0^1 f(x,y)\,dy = \int_0^1 4xy\,dy = 2x.$$

so. $$f(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} 2y, & 0 \le y \le 1. \\ 0. & o/w \end{cases}$$

Eg 2. If (X, Y) have joint pdf

$$f(x,y) = \begin{cases} 8xy, & \text{if } 0 \le y \le x \le 1. \\ 0, & \text{o/w.} \end{cases}$$

Find $f(y|x)$ for any $0 \le x \le 1$.

Sol: $$f_X(x) = \begin{cases} 4x^3, & 0 \le x \le 1. \\ 0, & \text{o/w} \end{cases}$$

So. $$f(y|x) = \begin{cases} 2x^{-2}y, & \text{if } 0 \le y \le x. \\ 0, & \text{o/w.} \end{cases}$$

Ex 1. If (X, Y) have joint pdf

$$f(x,y) = \begin{cases} x+y, & \text{if } 0 \le x, y \le 1. \\ 0, & \text{o/w} \end{cases}$$

Find $f(y|x)$. for any $0 \le x \le 1$.

Sol: $$f_X(x) = \begin{cases} \frac{1}{2}(1+x), & \text{if } 0 \le x \le 1. \\ 0, & \text{o/w} \end{cases}$$

So $$f(y|x) = \begin{cases} \frac{2(x+y)}{1+x}, & \text{if } 0 \le y \le 1. \\ 0, & \text{o/w.} \end{cases}$$

On the other hand, if we know the conditional distribution of Y given $X=x$, i.e., $f(y|x)$, and also the marginal distribution of $f_X(x)$, then the joint distribution of (X, Y) is

$$f(x,y) = f(y|x)\, f_X(x).$$

Eg 3. If $f(y|x)$ is given below.

$$f(y|x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}. \quad \text{and} \quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

then the joint pdf of (X, Y) is

$$f(x,y) = f(y|x)\, f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$= \frac{1}{2\pi} e^{-\frac{x^2+(y-x)^2}{2}}.$$

Ex 2. If $f(y|x)$ is given by

$$f(y|x) = \begin{cases} \alpha e^{-(y-x)\alpha}, & y \geq x \\ 0. & \text{o/w.} \end{cases}$$

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x \geq 0. \\ 0, & x < 0. \end{cases} \quad \text{Find } f(x,y).$$

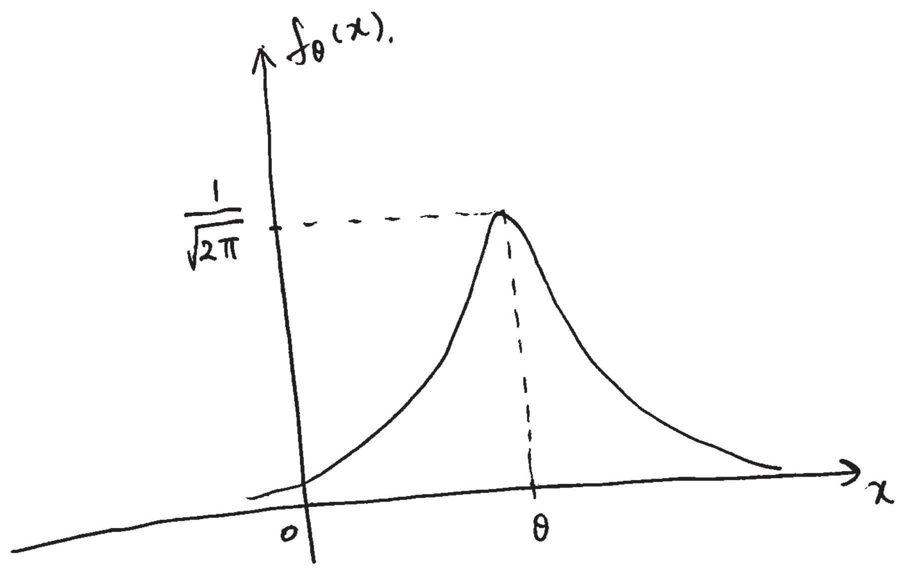
Sol:

$$f(x,y)=\begin{cases}\alpha^2 e^{-\alpha y}, & \text{if } \quad y \geq x \geq 0. \\ 0, & \text{o/w.}\end{cases}$$

② Bayesian model: Eg5. Suppose $X \sim N(\theta, 1)$. That is, the pdf of X is

$$f_\theta(x)=\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(x-\theta)^2}{2}}, \qquad x \in \mathbb{R}.$$

The plot of $f_\theta(x)$ is as below:



We don't know the value of θ. In previous sections and chapters, we assume θ is deterministic but unknown, we proposed methods to estimate θ, e.g. use $\hat{\theta}_{ML}$, the MLE of θ. We all agree with this method.

The Bayesian statisticians don't agree with this. they believe that, since θ is unknown, θ can be anywhere on the real line so it should be treated as a random variable. people should even place a distribution on θ, i.e. the socalled prior distribution.

In this chapter, we introduce this somewhat new perspective.

Let us go back to our example, Bayesians believe we should ~~treat~~ treat the location θ as a random variable, and the assume θ follows the following pdf $\theta \sim N(0,1)$. i.e. the pdf of θ is

$$f(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}.$$

Using this idea, the ~~joint~~ $f_\theta(x)$ should be considered as the conditional distribution of X given θ. i.e,

$$f(x \mid \theta) = f_\theta(x).$$

the joint pdf of (X, θ) becomes.

$$f(x, \theta) = f(x \mid \theta)\, f(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}.$$

$$= \frac{1}{2\pi} e^{-\frac{\theta^2 + (x-\theta)^2}{2}}.$$

Now we are able to find the conditional distribution of θ given $X=x$.

The marginal pdf of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x,\theta)\, d\theta = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{\theta^2+(x-\theta)^2}{2}}\, d\theta$$

$$= \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}.$$

$$\text{so. } f(\theta \mid x) = \frac{f(x,\theta)}{f_X(x)} = \frac{\frac{1}{2\pi} e^{-\frac{\theta^2+(x-\theta)^2}{2}}}{\frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}} = \frac{1}{\sqrt{\pi}} e^{-(\theta-\frac{1}{2}x)^2}.$$

so $\theta \mid X=x \sim N(\frac{1}{2}x, \frac{1}{2})$.

Change to the following motivating example:

Eg4.

If ~~the~~ $X \sim \text{Exponential}(\theta)$. that is, $f_\theta(x)$ is

$$f_\theta(x) = \begin{cases} \theta e^{-x\theta}, & x \geq 0. \\ 0, & \text{o/w}. \end{cases}$$

Assume θ follows $\theta \sim \text{Exponential}(1)$.

$$f(\theta) = \begin{cases} e^{-\theta}, & \theta \geq 0 \\ 0, & \text{o/w}. \end{cases}$$

Then the joint pdf of (X, θ) is

$$f(x,\theta) = f(x|\theta)\, f(\theta)$$

$$= \begin{cases} \theta e^{-(x+1)\theta}, & \theta, x \geq 0. \\ 0, & o/w. \end{cases}$$

So the marginal pdf of X is

$$f_X(x) = \int_0^\infty f(x,\theta)\,d\theta = \int_0^\infty \theta e^{-(x+1)\theta}\,d\theta$$

$$= -\frac{1}{x+1} e^{-(x+1)\theta}\,\theta \Big|_0^\infty + \int_0^\infty \frac{1}{x+1} e^{-(x+1)\theta}\,d\theta$$

$$= -\frac{1}{(x+1)^2} e^{-(x+1)\theta}\Big|_0^\infty = \frac{1}{(x+1)^2}.$$

So. $$f(\theta|x) = \begin{cases} (x+1)^2\,\theta e^{-(x+1)\theta}. & \theta \geq 0. \\ 0. & o/w. \end{cases}$$

In summary. Suppose we know the distribution of X, i.e., $f_\theta(x)$.

step 1. Treat $f_\theta(x)$ as the conditional distribution of X given θ, i.e.

$$f(x|\theta) = f_\theta(x).$$

step 2. place a prior distribution $f(\theta)$ on θ.

step 3. Find joint distribution of (X, θ) by

$$f(x,\theta) = f(x|\theta)\, f(\theta) = f_\theta(x)\, f(\theta).$$

step 4. Find the ~~conditional~~ posterior distribution of θ given x by.

$$f(\theta|x) = \frac{f(x,\theta)}{f_X(x)} = \frac{f(x,\theta)}{\int f(x,\theta)\,d\theta}.$$, or simply, by

$$f(\theta|x) \propto f(x,\theta) = f_\theta(x)\, f(\theta).$$

also we can use the following example:

Eg 6. Suppose $X|\theta \sim unif(0,\theta)$. θ has ~~pdf~~ prior

$$f(\theta) = \begin{cases} 2\theta, & 0 \le \theta \le 1. \\ 0. & o/w \end{cases}$$

Find the posterior of θ given x.

then $f(x|\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta. \\ 0. & o/w. \end{cases}$ so $f(x|\theta) = \frac{1}{\theta} \mathbb{1}_{0 \le x \le \theta}$.

so $f(x,\theta) = f(x|\theta)\, f(\theta) = \begin{cases} 2\,\mathbb{1}_{0 \le x \le \theta}, & 0 \le x \le \theta \le 1, \\ 0. & o/w. \end{cases}$

so $f_X(x) = \int_x^1 f(x,\theta)\,d\theta = \int_x^1 2\,d\theta = 2(1-x).$

so $$f(\theta \mid x) = \begin{cases} 2(1-x), & x \le \theta \le 1. \\ 0. & o/w. \end{cases}$$