Lecture 3: Exponential Families and Sufficiency

Probability: Given $X_n \overset{iid}{\sim} f_\theta$ some parameter

Ex: If $\theta = 5$ calculate $P(X_n = \cdots)$

Statistics: Given $X_n \overset{iid}{\sim} f_\theta$

Ex: I don't know θ, how can I estimate it?

Generally we'll work w/ parameterized fams of dists:

e.g.
- $N(\mu, \sigma^2)$ when $\mu \in \mathbb{R}, \sigma^2 > 0$
- $Exp(\lambda)$ where $\lambda > 0$
- $U(0, \theta)$ where $\theta > 0$

Exponential Families

Assume we have a fam of dists parameterized by $\theta \in \Theta \subset \mathbb{R}$

so that $X_n \overset{iid}{\sim} f_\theta$

where

Joint dist

fns of θ not x

$$f_\theta(\underset{\sim}{x}) = h(\underset{\sim}{x})\,c(\theta)\exp\left(T(\underset{\sim}{x})\,w(\theta)\right)$$

fns of x not θ

then we say the X_n's are from an exponential family.

Ex, Poisson, Exp, Normal, Gamma, Beta, ...

Ex, $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$ for $\lambda > 0$

θ Θ

$$\boxed{f(x) = \frac{1}{x!}\lambda^x e^{-\lambda}\,\mathbb{1}(x \in \mathbb{N}_0)}$$

$$f(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n) = \prod_{n=1}^{N} \frac{1}{x_n!}\lambda^{x_n} e^{-\lambda}\,\mathbb{1}(x_n \in \mathbb{N}_0)$$

$$= \prod_n\left(\frac{1}{x_n!}\right)\prod_n\left(\lambda^{x_n}\right)\prod_n e^{-\lambda}\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)$$

$$\boxed{\begin{aligned} a &= e^{\log a} \\ \log(a^b) &= b\log a \end{aligned}}$$

$$= \prod_n\left(\frac{1}{x_n!}\right)\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)\,\lambda^{\sum_n x_n} e^{-N\lambda}$$

$$e^{\log(\lambda^{\sum x_n})} = e^{(\sum_n x_n)\log\lambda}$$

$$= \underbrace{\prod_n\left(\frac{1}{x_n!}\right)\prod_n \mathbb{1}(x_n \in \mathbb{N}_0)}_{h(\underset{\sim}{x})}\,\underbrace{e^{-N\lambda}}_{c(\lambda)}\exp\Big(\underbrace{\left(\textstyle\sum_n x_n\right)}_{T(\underset{\sim}{x})}\underbrace{\log\lambda}_{w(\lambda)}\Big)$$

So $\text{Pois}(\lambda)$ is an exp. fam.

Short-cut: can just check the marginal.

If

$$f_\theta(x) = h_0(x)\, c_0(\theta) \exp(T_0(x)\, \omega_0(\theta))$$

marginal

then

$$f_\theta(\underset{\sim}{x}) = \prod_n f_\theta(x_n) = \prod_n h_0(x_n) c_0(\theta) \exp(T_0(x_n)\, \omega_0(\theta))$$

$$= \prod_n h_0(x_n) \prod_n c_0(\theta) \exp\left(\sum_n T_0(x_n)\, \omega_0(\theta)\right)$$

$$\exp\left(\omega_0(\theta) \sum_n T_0(x_n)\right)$$

$$f_\theta(\underset{\sim}{x}) = \underbrace{\prod_n h_0(x_n)}_{h(\underset{\sim}{x})} \underbrace{\prod_n c_0(\theta)}_{c(\theta)} \exp\left(\underbrace{\left(\sum_n T_0(x_n)\right)}_{T(\underset{\sim}{x})} \underbrace{\omega_0(\theta)}_{\omega(\theta)}\right)$$

punchline:
$$h(\underset{\sim}{x}) = \prod_n h_0(x_n)$$
$$c(\theta) = \prod_n c_0(\theta)$$
$$T(\underset{\sim}{x}) = \sum_n T_0(x_n)$$
$$\omega(\theta) = \omega_0(\theta)$$

Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$

$$f(x_n) = \frac{1}{x_n!} \lambda^{x_n} e^{-\lambda}\, \mathbb{1}(x_n \in \mathbb{N}_0)$$

$$= \underbrace{\frac{1}{X_n!} \mathbb{1}(X_n \in \mathbb{N}_0)}_{h_0(X_n)} \underbrace{e^{-\lambda}}_{c_0(\lambda)} \exp(\underbrace{X_n}_{T_0(X_n)} \underbrace{\log \lambda}_{w(\lambda)})$$

$$T(\underline{X}) = \sum_n T_0(X_n) = \sum_n X_n .$$

Ex. $X_n \overset{iid}{\sim} U(0, \theta)$, $\theta > 0$.

Exp fam?

$$f_\theta(x_n) = \frac{1}{\theta} \text{ for } 0 < x_n < \theta$$

$$= \frac{1}{\theta} \underbrace{\mathbb{1}(0 < X_n < \theta)}_{\text{no way to write as } (\text{fn of } X) \cdot (\text{fn } \theta)}$$

No, not an exp. fam.

Short-cut 2: general fact, if the support of the dist depends on the parameter, its not an exp. family.

Defn: Sufficiency [$X_n \overset{iid}{\sim} f_\theta$, θ unknown]

We say a statistic $T = T(\underline{X})$ is sufficient for a parameter θ if

$f_{\underline{X} | T=t}(\underline{x})$ is "free" of θ.

free = formal for dist doesn't depend on θ.

<u>Ex.</u> If I have data $X_1, \ldots, X_N$ is $X_1 = T$ sufficient for θ?

No.

Ex. Is $T = \underset{\sim}{X}$ sufficient for θ?

<u>Yes.</u>

$$f(\underset{\sim}{x} \mid T = t) = \frac{f_{\underset{\sim}{X}, T}(\underset{\sim}{x}, t)}{f_T(t)} = \frac{f_{\underset{\sim}{X}, \underset{\sim}{X}}(\underset{\sim}{x}, \underset{\sim}{x})}{f_{\underset{\sim}{X}}(\underset{\sim}{x})}$$

$$= f_{\underset{\sim}{X}}(\underset{\sim}{x}) \Big/ f_{\underset{\sim}{X}}(\underset{\sim}{x}) = 1$$

no θ.

<u>Theorem: Factorization Theorem</u>

T is sufficient for θ <u>iff</u>

there is a fn $g(\theta, T)$ and $h(\underset{\sim}{x})$ so that

$$\text{joint} \rightarrow f_\theta(\underset{\sim}{x}) = \underbrace{g(\theta, T)}_{\substack{\text{fn of } \theta \\ \text{and data only through } T}} h(\underset{\sim}{x}) \leftarrow \text{no } \theta$$

Ex. Let $X_n \overset{iid}{\sim} \text{Bern}(\theta)$, $\theta \in [0,1]$

$$T = \frac{1}{N}\sum_{n=1}^{N} X_n = \bar{X}$$

$$\boxed{f(x) = \begin{cases} \theta & x=1 \\ 1-\theta & x=0 \end{cases} = \theta^{x}(1-\theta)^{1-x}\,\mathbb{1}(x=0,1)}$$

Is T sufficient for θ?

$$f_\theta(\underset{\sim}{x}) = \prod_n \theta^{x_n}(1-\theta)^{1-x_n}\,\mathbb{1}(x_n = 0 \text{ or } 1)$$

$$\boxed{\begin{array}{l} \bar{X} = \frac{1}{N}\sum_n X_n \\ \sum_n X_n = N\bar{X} \\ a^{b+c} = a^b a^c \end{array}}$$

$$= \theta^{\sum_n x_n}(1-\theta)^{\sum_n (1-x_n)}\prod_n \mathbb{1}(x_n = 0 \text{ or } 1)$$

$$= \theta^{\sum_n x_n}(1-\theta)^{N-\sum_n x_n}\prod_n \mathbb{1}(x_n = 0 \text{ or } 1)$$

$$= \theta^{N\bar{X}}(1-\theta)^{N-N\bar{X}}\prod_n \mathbb{1}(x_n = 0 \text{ or } 1)$$

$$= \theta^{N\bar{X}}(1-\theta)^{N}(1-\theta)^{-N\bar{X}}\prod_n \mathbb{1}\cdots$$

$$= \frac{\theta^{N\bar{X}}}{(1-\theta)^{N\bar{X}}}(1-\theta)^N\prod \mathbb{1}(\cdots)$$

$$= \underbrace{\left(\frac{\theta}{1-\theta}\right)^{N\bar{X}}(1-\theta)^N}_{g(T,\theta)}\;\underbrace{\prod_n \mathbb{1}(x_n = 0 \text{ or } 1)}_{h(\underset{\sim}{x})}$$

$$g(T,\theta) = \left(\frac{\theta}{1-\theta}\right)^{NT}(1-\theta)^N$$

So $\bar{X}$ is sufficient for θ.

Ex. Let $X_n \overset{iid}{\sim} U(0,\theta)$

$$f_\theta(x_n) = \frac{1}{\theta}\mathbb{1}(0<x_n<\theta)$$
$$= \frac{1}{\theta}\mathbb{1}(x_n>0)\mathbb{1}(x_n<\theta)$$

Can I find a SS?
one-dim'l

$$f_\theta(\underset{\sim}{x}) = \prod_n \frac{1}{\theta}\mathbb{1}(x_n>0)\mathbb{1}(x_n<\theta)$$

$$= \theta^{-N}\prod_n \mathbb{1}(x_n<\theta)\prod_n \mathbb{1}(x_n>0)$$

$$= \theta^{-N}\mathbb{1}(\text{all } x_n<\theta)\mathbb{1}(\text{all } x_n>0)$$

$$\prod \mathbb{1}(A_n) = \mathbb{1}(\text{all } A_n)$$

$$= \underbrace{\theta^{-N}\mathbb{1}(X_{(N)}<\theta)}_{g(T,\theta)}\underbrace{\mathbb{1}(X_{(1)}>0)}_{h(\underset{\sim}{x})}$$

$T = X_{(N)}$

So $X_{(N)}$ is sufficient for θ.