- Today: Notes 7
- Wed/Fri: Notes 8, Lab
- Mon 10/23: review
wed: quiz 02

- Final: Dec 12 9-12
- Quiz 3 moved from wed ← Mon
- HW 5 due wed
- HW 6 completion based

- OH today 11:20 - 12:20
wed 2:30-4
- Faculty lecture Thurs @ 4:30
Scheuer Room

07: EXPONENTIAL FAMILIES

Rice 8.8.1

Prof Amanda Luby

---

1 Exponential Families

Many "nice" distributions that we've spent a lot of time with are members of the **exponential family**.

> **Exponential Family**
>
> One-parameter members of the exponential family have density functions of the form:
>
> $$f(x;\theta) = \exp[\eta(\theta)T(x) - A(\theta) + B(x)]$$
> $$= h(x)\exp[\eta(\theta)T(x) - A(\theta)]$$
> $$= h(x)g(\theta)\exp[\eta(\theta)T(x)]$$
>
> * support of f_x does not depend on θ
>
> $\eta(\theta)$ = "natural parameter"
>
> $T(x)$ = sufficient statistic

Writing densities in this form requires a bit of work.

Example: Poisson distribution

* Good strategy is to write pdf as $\exp(\ln(f_x))$

$$P(Y=y) = \frac{\lambda^y}{y!}e^{-\lambda} \quad y \geq 0$$

$$= \exp\left(\ln\left(\frac{\lambda^y}{y!}e^{-\lambda}\right)\right)$$

$$= \exp[y\ln\lambda - \ln y! - \lambda\ln e]$$

$$= \exp[\underbrace{\ln\lambda}_{\eta(\theta)} \cdot \underbrace{y}_{T(x)} - \underbrace{\lambda}_{A(\lambda)} + \underbrace{(-\ln y!)}_{B(x)}]$$

$\Rightarrow$ Poisson distributions are members of the exponential family

Recall: $T(x)$ is sufficient for θ if $L(\theta) = g(T(x_i), \theta) \cdot b(x_i)$

$$f(x;\theta) = \exp[\eta(\theta)T(x) - A(\theta) + B(x)]$$

BUT, the payoff is that once we write the pdf in *exponential form*, we can immediately identify a sufficient statistic:

$$L(\theta) = \prod_{i=1}^{n} f(x;\theta) = \prod \exp(\eta(\theta)T(x) - A(\theta) + B(x))$$

$$= \underbrace{\exp\left(\eta(\theta)\sum T(X_i) - nA(\theta)\right)}_{g(\sum T(X_i),\, \theta)} \underbrace{\exp\left(\sum B(X_i)\right)}_{b(X_i)}$$

$n=1$: $T(x)$ is sufficient

$n=n$: $\sum T(x)$ is sufficient

$\Rightarrow$ By factorization theorem, $\sum T(X_i)$ is sufficient for θ !!!

We also have the following results:

Sufficiency and the MLE

If $T(x)$ is sufficient for θ, the MLE is a function of $T(x)$

To maximize $L(\theta)$ WRT θ, need to maximize $g(\sum T(X_i), \theta)$ WRT θ. g depends on the data only through $\sum T(X_i)$, so MLE must be a function of $\sum T(X_i)$.

Mean and Variance of Exponential Families $f(x;\theta) = \exp[\eta(\theta)T(x) - A(\theta) + B(x)]$

If Y is a member of exponential family,

$$E(Y) = \frac{\partial}{\partial \eta} A(\eta)$$

$$V(Y) = \frac{\partial^2}{\partial \eta^2} A(\eta)$$

Conjugate priors If $f_X(x|\theta)$ is a member of exponential family, there is a conjugate prior: $f_\theta(\theta)$ where $f_{\theta|X}(\theta|x)$ has same family as f_θ.

original: $f_x(x;\theta)$
export: $f_x(x;\eta)$

$$f(x;\theta) = \exp[\eta(\theta)T(x) - A(\theta) + B(x)]$$

Example: Poisson distribution

$$f_x = \exp\{\underbrace{\ln\lambda}_{\eta(\lambda)} \cdot \underbrace{y}_{T(y)} - \underbrace{\lambda}_{A(\lambda)} + \underbrace{(-\ln y!)}_{B(y)}\}$$

① $\sum T(x_i) = \sum y_i$ is sufficient for λ

② $A(\lambda) = \lambda$, $\eta = \ln\lambda$ } $A(\eta) = e^{\ln\lambda} = e^{\eta}$

$$E[Y] = \frac{\partial}{\partial\eta}A(\eta) = \frac{\partial}{\partial\eta}e^{\eta} = e^{\eta} = e^{\ln\lambda} = \lambda$$

$$V[Y] = \frac{\partial^2}{\partial\eta^2}A(\eta) = \frac{\partial^2}{\partial\eta^2}e^{\eta} = e^{\eta} = e^{\ln\lambda} = \lambda$$

k-parameter exponential family

$$f(x;\vec{\theta}) = \exp\left[\sum_{i=1}^{k}\eta_i(\theta)T_i(x) - A(\theta) + B(x)\right]$$

Example: $X_1, ..., X_n \sim N(\mu, \sigma^2)$

$$= h(x)g(\theta)\underbrace{\exp[\eta(\theta)T(x)]}_{\exp[\sum_{i=1}^{k}\eta_i(\theta)T_i(x)]}$$

$$f_y = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(y-\mu)^2/2\sigma^2}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)e^{-\frac{y^2}{2\sigma^2} + \frac{2y\mu}{2\sigma^2} - \frac{\mu^2}{2\sigma^2}}$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)e^{-\mu^2/2\sigma^2}}_{g(\theta)} \cdot e^{\underbrace{-\frac{1}{2\sigma^2}\cdot y^2}_{\eta_1(\theta)\cdot T_1(x)} + \underbrace{\frac{\mu}{\sigma^2}\cdot y}_{\eta_2(\theta)\cdot T_2(x)}}$$

$\eta_1 = -1/2\sigma^2$ $\quad$ $\eta_2 = \frac{\mu}{\sigma^2}$

$T_1 = y^2$ $\quad$ $T_2 = y$

$\Rightarrow N(\mu,\sigma^2)$ is a member of the 2-parameter E.F.

$\rightarrow$ sufficient stats for $n=1$

$T_1 = y \quad T_2 = y^2$

$\Rightarrow$ S.S. for sample size n

$T_1 = \sum y_i \quad T_2 = \sum y_i^2$

1.1 More distributions