Exponential Families.

- Statistics: data $\longrightarrow$ knowledge
- exponential families: classes of distributions.
- $h: \mathbb{R}^n \to \mathbb{R}$ nonnegative function

$T_1 \cdots T_s$ — measurable functions: $\mathbb{R}^n \to \mathbb{R}$.

$$A(\eta) = \log \int e^{\sum_{i=1}^{s} \eta_i T_i(x)} h(x)\, d\mu(x).$$

μ — measure on $\mathbb{R}^n$

$\eta = (\eta_1 \cdots \eta_s) \in \mathbb{R}^s$.

$$p_\eta(x) = e^{\sum_{i=1}^{s} \eta_i T_i(x) - A(\eta)} \cdot h(x) \qquad x \in \mathbb{R}^n.$$

$\int p_\eta(x)\, d\mu = 1$. $\longrightarrow$ a family of probability densities indexed by η.

$\Xi = \{\eta : A(\eta) < \infty\}$. —— natural parameter space

$\{p_\eta : \eta \in \Xi\}$ —— s-parameter exponential family in canonical form.

ex. exponential density: $p_\eta(x) = -\eta\, e^{+\eta x}$ $\quad \eta < 0 \quad x > 0$

$p_\eta(x) = e^{+\eta x + \log(-\eta)}$ $\qquad h = 1_{(0, +\infty)} \quad s = 1$.

$T_1(x) = x$. $\qquad A(\eta) = \log \int e^{+\eta x} dx = \begin{cases} \log(-\frac{1}{\eta}) & \eta < 0 \\ \infty & \eta \geq 0 \end{cases}$

- other parameterization.

$$p_\theta(x) = e^{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)} \cdot h(x) \qquad \theta \in \Omega. \quad x \in \mathbb{R}^n.$$

$$B(\theta) = A(\eta(\theta)).$$

$\{p_\theta : \theta \in \Omega\}$ —— s-parameter exponential families.

ex: $N(\mu, \sigma^2)$:

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

$\theta = (\mu, \sigma^2)$. Two-parameter exponential family.

$T_1(x) = x$. $T_2(x) = x^2$. $\eta_1(\theta) = \frac{\mu}{\sigma^2}$. $\eta_2(\theta) = -\frac{1}{2\sigma^2}$. $B(\theta) = \frac{\mu^2}{2\sigma^2} + \log\sigma$.

$h(x) = \frac{1}{\sqrt{2\pi}}$.

ex. $X_1 \cdots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

$$p(x_1 \cdots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 - n\left(\frac{\mu^2}{2\sigma^2} + \log\sigma\right)}$$

$T_1(x) = \sum_{i=1}^{n} x_i$ $\quad T_2(x) = \sum_{i=1}^{n} x_i^2$ $\quad \eta_1(\theta) = \frac{\mu}{\sigma^2}$ $\quad \eta_2(\theta) = -\frac{1}{2\sigma^2}$

$B(\theta) = n\left[\frac{\mu^2}{2\sigma^2} + \log\sigma\right]$ $\quad h(x) = \frac{1}{(2\pi)^{n/2}}$.

$X_1 \cdots X_n \overset{iid}{\sim} \exp\left[\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right] h(x)$

joint $\exp\left[\sum_{i=1}^{s} \eta_i(\theta)\left(\sum_{j=1}^{n} T_i(x_j)\right) - n B(\theta)\right] \prod_{j=1}^{n} h(x_j)$

s-para. exp. family. $\tilde{T}_i(x) = \sum_{j=1}^{n} T_i(x_j)$ $\tilde{B}(\theta) = n B(\theta)$

$\tilde{h}(x) = \prod_{i=1}^{n} h(x_i)$.

- Relate moments and cumulants for the statistic $T_1 \cdots T_s$ to derivatives of A.

Thm. if $\int |f(x)| e^{\sum \eta_i T_i(x)} h(x)\, d\mu(x) < \infty$

$\Rightarrow$ $g(\eta) = \int f(x) e^{\sum \eta_i T_i(x)} h(x)\, d\mu(x)$ is continuous and has continuous partial derivatives of all orders of η.

these derivatives can be computed by differentiation under the integral sign.

$$g(\eta) = e^{A(\eta)} = \int e^{\sum_i \eta_i T_i(x)} h(x)\, d\mu(x).$$

$$e^{A(\eta)} \frac{\partial A(\eta)}{\partial \eta_j} = \int \frac{\partial}{\partial \eta_j} e^{\sum_i \eta_i T_i(x)} h(x)\, d\mu(x) = \int T_j(x)\, e^{\sum \eta_i T_i(x)} h(x) d\mu(x)$$

$$\frac{\partial A(\eta)}{\partial \eta_j} = \int T_j(x)\, P_\eta(x)\, d\mu(x). \Rightarrow \quad E_\eta T_j(X) = \frac{\partial A(\eta)}{\partial \eta_j}.$$

$T = (T_1 \cdots T_s)' \in \mathbb{R}^s$. a random vector.

- mgf $\quad M_T(u) = E e^{u_1 T_1 + \cdots + u_s T_s} = E e^{u^* T} \quad u \in \mathbb{R}^s$.

- cumulant generating function : $K_T(u) = \log M_T(u)$

Lem 2.7. If $M_X(u)$ and $M_Y(u)$ are finite and agree for u in some set with a nonempty interior, then X and Y have the same distribution. $P_X = P_Y$.

- moments: $\quad a_{r_1 \cdots r_s} = E[T_1^{r_1} \times \cdots \times T_s^{r_s}]$

Thm. $\quad a_{r_1 \cdots r_s} = \frac{\partial^{r_1}}{\partial u_1^{r_1}} \cdots \frac{\partial^{r_s}}{\partial u^{r_s}} M_T(u) \Big|_{u=0}$.

- $K_{r_1 \cdots r_s} = \frac{\partial^{r_1}}{\partial u_1^{r_1}} \cdots \frac{\partial^{r_s}}{\partial u_s^{r_s}} K_T(u) \Big|_{u=0}$: cumulants.

$\underline{s=1}$. $K_T' = \frac{M_T'}{M_T}$ $K_T'' = \frac{M_T M_T'' - M_T'^2}{M_T^2}$

$k_1 = ET$. $k_2 = ET^2 - (ET)^2 = Var(T)$

Lem. $X \perp Y \Rightarrow E(XY) = EX\, EY$.

$\not\Leftarrow$.

If $T = Y_1 + \cdots + Y_n$. $Y_1 \cdots Y_n$ independent.

$M_T(u) = E e^{u'T} = M_{Y_1}(u) \times \cdots \times M_{Y_n}(u)$

$K_T(u) = K_{Y_1}(u) + \cdots + K_{Y_n}(u)$.

- $X \sim e^{\eta T(x) - A(\eta)} h(x) \Rightarrow T$ has mgf

$$E_\eta e^{u \cdot T(x)} = \int e^{u T(x)} e^{\eta T(x) - A(\eta)} h(x)\, d\mu(x)$$

$$= e^{A(u+\eta) - A(\eta)} \int e^{(u+\eta) T(x) - A(u+\eta)} h(x)\, d\mu(x)$$

$$= e^{A(u+\eta) - A(\eta)}$$

$K_T(u) = A(u+\eta) - A(\eta)$.

$$K_{r_1 \cdots r_s} = \frac{\partial^{r_1}}{\partial \eta_1^{r_1}} \cdots \frac{\partial^{r_s}}{\partial \eta_s^{r_s}} A(\eta)$$

ex. $X \sim \text{Poisson}(\lambda)$ $\quad P(X=x) = \dfrac{\lambda^x e^{-\lambda}}{x!} = \dfrac{1}{x!} e^{x\log\lambda - \lambda}$ $\quad x = 0, 1, \cdots$

not in canonical form. The canonical para. is $\eta = \log\lambda$.

$$P(X=x) = \frac{1}{x!} e^{\eta x - e^{\eta}} \qquad x = 0, 1, \cdots \qquad A(\eta) = e^{\eta}$$

all of the cumulants of $T = X$ are $e^{\eta} = \lambda$.

ex. $N(\mu, \sigma^2)$. $\quad \sigma^2$ is known.

$$p_\mu(x) = e^{\frac{\mu x}{\sigma^2} - \frac{\mu^2}{2\sigma^2}} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

$T(x) = x$. $\quad \eta = \dfrac{\mu}{\sigma^2}$. $\quad A(\eta) = \dfrac{\sigma^2 \eta^2}{2}$.

$\kappa_1 = A'(\eta) = \sigma^2\eta = \mu$. $\quad \kappa_2 = A''(\eta) = \sigma^2$. $\quad$ higher-order cumulants $\kappa_3, \kappa_4, \cdots$ are zero

- Calculate moments from cumulants. when $s = 1$.

$$M = e^K \qquad M' = K'e^K \qquad M'' = (K'' + K'^2)e^K$$

$$M''' = (K''' + 3K'K'' + K'^3)e^K \qquad M'''' = (K'''' + 3K''^2 + 4K'K''' + 6K'^2K'' + K'^4)e^K$$

At zero:

$$ET = \kappa_1 \qquad ET^2 = \kappa_2 + \kappa_1^2$$

$$ET^3 = \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 \qquad ET^4 = \kappa_3 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + 6\kappa_1^2\kappa_2 + \kappa_1^4$$

ex $X \sim \text{Pois}(\lambda)$ $\quad EX = \lambda$ $\quad EX^2 = \lambda + \lambda^2$. $\quad EX^3 = \lambda + 3\lambda^2 + \lambda^3$

$EX^4 = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4$.

$X \sim N(\mu, \sigma^2)$. $\quad EX^3 = 3\mu\sigma^2 + \mu^3$. $\quad EX^4 = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4$.