Exponential Family Facts
Stat 543 Spring 2016

Many standard families of distributions can be treated with a single set of analyses by recognizing them to be of a common "exponential family" form. The following is a list of facts about such families. (See, for example, Schervish Sections 2.2.1 and 2.2.2, or scattered results in Shao or the old books by Lehmann (*TSH* and *TPE*) for details.)

Definition 1 *Suppose that $\mathcal{P} = \{P_\theta\}$ is defined by pdf's or pmf's $f_\theta(x)$ for X. $\mathcal{P}$ is called an* ***exponential family*** *if for some $h(x) \geq 0$,*

$$f_\theta(x) \doteq \frac{dP_\theta}{d\mu}(x) = \exp\left(a(\theta) + \sum_{i=1}^{k} \eta_i(\theta) T_i(x)\right) h(x) \ \ \forall\theta \quad .$$

Claim 2 *Let $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_k)$, $\mathcal{E} = \{\boldsymbol{\eta} \in \mathcal{R}^k | \int h(x) \exp\left(\sum \eta_i T_i(x)\right) d\mu(x) < \infty\}$, and consider also the family of distributions with pdf's or pmf's of the form*

$$f_{\boldsymbol{\eta}}(x) = C(\boldsymbol{\eta}) \exp\left(\sum_{i=1}^{k} \eta_i T_i(x)\right) h(x) \quad , \quad \boldsymbol{\eta} \in \mathcal{E} \quad .$$

Call this family $\mathcal{P}^*$. This set of distributions for X is at least as large as $\mathcal{P}$ (and this second parameterization is mathematically nicer than the first). $\mathcal{E}$ is called the natural parameter space for $\mathcal{P}^*$. It is a convex subset of $\mathcal{R}^k$. (TSH, page 57.) If $\mathcal{E}$ lies in a subspace of dimension less than k, then $f_{\boldsymbol{\eta}}(x)$ (and therefore $f_\theta(x)$) can be written in a form involving fewer than k statistics T_i. We will henceforth assume $\mathcal{E}$ to be fully k-dimensional. Note that depending upon the nature of the functions $\eta_i(\theta)$ and the parameter space Θ, $\mathcal{P}$ may be a proper subset of $\mathcal{P}^*$. That is, defining $\mathcal{E}_\Theta = \{(\eta_1(\theta), \eta_2(\theta), ..., \eta_k(\theta)) \in \mathcal{R}^k | \theta \in \Theta\}$, $\mathcal{E}_\Theta$ can be a proper subset of $\mathcal{E}$.

Claim 3 *The "support" of P_θ, defined as $\{x | f_\theta(x) > 0\}$, is clearly $\{x | h(x) > 0\}$, which is independent of θ.*

Claim 4 *From the Factorization Theorem, the statistic $\boldsymbol{T} = (T_1, T_2, ..., T_k)$ is sufficient for $\mathcal{P}$.*

Claim 5 *$\boldsymbol{T}$ has distributions derived from the distributions for X, say $\{P_\theta^{\boldsymbol{T}} | \theta \in \Theta\}$, which also form an exponential family.*

Claim 6 *If $\mathcal{E}_\Theta$ contains an open rectangle in $\mathcal{R}^k$, then $\boldsymbol{T}$ is complete for $\mathcal{P}$. (See pages 142-143 of TSH.)*

Claim 7 *If $\mathcal{E}_\Theta$ contains an open rectangle in $\mathcal{R}^k$ (and actually under the much weaker assumptions given on page 44 of TPE) $\boldsymbol{T}$ is minimal sufficient for $\mathcal{P}$.*

Claim 8 *If g is any real-valued function such that $E_{\boldsymbol{\eta}}|g(X)| < \infty$, then*

$$E_{\boldsymbol{\eta}} g(X) = \int g(x) f_{\boldsymbol{\eta}}(x) d\mu(x)$$

is continuous on $\mathcal{E}$ and has continuous partial derivatives of all orders on the interior of $\mathcal{E}$. These can be calculated as

$$\frac{\partial^{\alpha_1+\alpha_2+\cdots+\alpha_k}}{\partial\eta_1^{\alpha_1}\partial\eta_2^{\alpha_2}\cdots\partial\eta_k^{\alpha_k}} E_{\boldsymbol{\eta}} g(X) = \int g(x) \frac{\partial^{\alpha_1+\alpha_2+\cdots+\alpha_k}}{\partial\eta_1^{\alpha_1}\partial\eta_2^{\alpha_2}\cdots\partial\eta_k^{\alpha_k}} f_{\boldsymbol{\eta}}(x) d\mu(x) \quad .$$

(See page 59 of TSH.)

Claim 9 *If for* $\boldsymbol{u} = (u_1, u_2, ..., u_k)$*, both* $\boldsymbol{\eta}_0$ *and* $\boldsymbol{\eta}_0 + \boldsymbol{u}$ *belong to* $\mathcal{E}$

$$E_{\boldsymbol{\eta}_0} \exp\Big\{u_1 T_1(X) + u_2 T_2(X) + \cdots + u_k T_k(X)\Big\} = \frac{C(\boldsymbol{\eta}_0)}{C(\boldsymbol{\eta}_0 + \boldsymbol{u})} \ .$$

Further, if $\boldsymbol{\eta}_0$ *is in the interior of* $\mathcal{E}$*, then*

$$E_{\boldsymbol{\eta}_0}(T_1^{\alpha_1}(X) T_2^{\alpha_2}(X) \cdots T_k^{\alpha_k}(X)) = C(\boldsymbol{\eta}_0) \frac{\partial^{\alpha_1 + \alpha_2 + \cdots + \alpha_k}}{\partial \eta_1^{\alpha_1} \partial \eta_2^{\alpha_2} \cdots \partial \eta_k^{\alpha_k}} \left(\frac{1}{C(\boldsymbol{\eta})} \right) \bigg|_{\boldsymbol{\eta} = \boldsymbol{\eta}_0} \ .$$

In particular, $E_{\boldsymbol{\eta}_0} T_j(X) = \frac{\partial}{\partial \eta_j} \left(-\ln C(\boldsymbol{\eta})\right) \Big|_{\boldsymbol{\eta} = \boldsymbol{\eta}_0}$*,* $Var_{\boldsymbol{\eta}_0} T_j(X) = \frac{\partial^2}{\partial \eta_j^2} \left(-\ln C(\boldsymbol{\eta})\right) \Big|_{\boldsymbol{\eta} = \boldsymbol{\eta}_0}$*, and* $Cov_{\boldsymbol{\eta}_0}\left(T_j(X), T_l(X)\right) = \frac{\partial^2}{\partial \eta_j \partial \eta_l} \left(-\ln C(\boldsymbol{\eta})\right) \Big|_{\boldsymbol{\eta} = \boldsymbol{\eta}_0}$*.*

Claim 10 *If* $\mathbf{X} = (X_1, X_2, ..., X_n)$ *has iid components, with each* $X_i \sim P_\theta$*, then* $\mathbf{X}$ *generates a* k*-dimensional exponential family. The* k*-dimensional statistic* $\sum_{i=1}^n \boldsymbol{T}(X_i)$ *is sufficient for this family. Under the condition that* $\mathcal{E}_\Theta$ *contains an open rectangle, this statistic is also complete and minimal sufficient.*