Exponential Families

A family $(f(x;\theta))_{\theta\in\Omega}$ is called a regular exponential family if $\Omega = \{\gamma < \theta < \delta\}$ and

$$f(x;\theta) = \begin{cases} e^{p(\theta)K(x) + S(x) + q(\theta)} & x \in S \\ 0 & x \notin S. \end{cases}$$

where S is the support of X.

Assumptions

1. S does not depend on θ
2. If X is cont $K'(x) \not\equiv 0$ & $S(x)$ is cont
3. If X is discrete, $K(x)$ is a nontrivial function of x.

Ex $f(x;\theta) \sim N(0,\theta)$

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

For $X \sim (f(x;\theta))$ from an exp family then $X_1, \ldots, X_n$ has joint pmf/pdf given by

$$e^{p(\theta)\sum_{i=1}^n K(x_i) + \sum_{i=1}^n S(x_i) + n q(\theta)}$$

$$= e^{p(\theta)\sum_{i=1}^n K(x_i) + nq(\theta)} \, e^{\sum S(x_i)}$$

From the factorization result before it follows that $T = \sum K(X_i)$ is a sufficient statistics.

Theorem ① $f_T(y_1;\theta) = R(y_1)\, e^{p(\theta)y_1 + nq(\theta)}$

② $E[Y_1] = -n\dfrac{q'(\theta)}{p'(\theta)}$ ③ $Var(Y_1) = \dfrac{n}{p'(\theta)^3}\left(p''(\theta)q'(\theta) - q''(\theta)p'(\theta)\right)$

where neither the support of Y_1 nor $R(y_1)$ depend on θ.

Theorem If $f(x;\theta)$ is an exponential family and $x_1, \dots, x_n$ is a sample, then $T = \sum k(x_i)$ is a complete and sufficient statistic.

Pf Sufficiency follows from the factorization. Completness from the fact that

$$E[u(T)] = \int u(t) R(t) e^{p(\theta)t + n q(\theta)} dt = 0$$

and this means that

$$\int u(t) R(t) e^{p(\theta)t} dt = 0 \quad \forall \ \theta$$

In particular since $p(\theta)$ is non-trivial and continuous, we have for a range of s on the real line

$$\int u(t) R(t) e^{st} dt = 0$$ which means that the Laplace transform of $u \cdot R$ is 0. Thus $u \cdot R \equiv 0$ and because $R \not\equiv 0$, $u \equiv 0$.

Conditions for the existence of a sufficient statistic

If T is a sufficient statistic for a family $f(x;\theta)$, then

$$\prod f(x_i;\theta) = g(T;\theta) \cdot k(x_1, \dots, x_n)$$

Taking ln and differentiating w.r.t θ we obtain

$$\sum_{i=1}^{m} \frac{\partial \ln f(x_i;\theta)}{\partial \theta} = \frac{\partial g}{\partial \theta}(T;\theta) = \underline{G(T;\theta)}$$

Now differentiating w.r.t x_i we get

$$\frac{\partial^2 \ln f(x_i;\theta)}{\partial \theta \partial x_i} = \frac{\partial G}{\partial T}(T,\theta) \cdot \frac{\partial T}{\partial x_i}$$

For a fixed value of $\theta = \theta_0$ we also get that

$$\left.\frac{\partial^2 \ln f(x;\theta)}{\partial \theta \partial x_i}\right|_{\theta=\theta_0} = \frac{\partial G}{\partial T}(T,\theta_0)\frac{\partial T}{\partial x_i}$$

Thus the ratio is

$$\frac{\dfrac{\partial^2 \ln f(x;\theta)}{\partial \theta \partial x_i}}{\left.\dfrac{\partial^2 \ln f(x;\theta)}{\partial \theta \partial x_i}\right|_{\theta=\theta_0}} = \frac{\dfrac{\partial G}{\partial T}(T,\theta)}{\dfrac{\partial G}{\partial T}(T,\theta_0)} = \lambda(\theta)$$

since the right hand side does not depend on i, we get that all of the left hand sides are equal and independent of i, thus

$$\frac{\partial G}{\partial T}(T,\theta) = \lambda(\theta) \cdot g(T)$$

In particular, integrating w.r.t T we also get:

$$\sum_{i=1}^{m} \frac{\partial \ln f(x_i;\theta)}{\partial \theta} = \lambda(\theta)\tilde{g}(T).$$

Integrating w.r.t θ gives

$$\sum_{i=1}^{m} \ln f(x_i;\theta) = \alpha(\theta)\tilde{g}(T) + R(x_1,..,x_m) + P(\theta)$$

Fixing now some of the values $x_2, x_3, -, x_m$, we get that

$$\ln f(x_1;\theta) = p(\theta)K(x_1) + S(x_1) + q(\theta).$$

Ex Take $X_1, \ldots, X_n \sim N(\theta, \sigma^2)$ where σ^2 is known. Then

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

$$= e^{\frac{\theta}{\sigma^2}x - \frac{x^2}{2\sigma^2} - \log\sqrt{2\pi\sigma^2} - \frac{\theta^2}{2\sigma^2}}$$

Thus $p(\theta) = \frac{\theta}{\sigma^2}$, $K(x) = x$

$$S(x) = -\frac{x^2}{2\sigma^2}, \quad q(\theta) = -\frac{\theta^2}{2\sigma^2} - \log\sqrt{2\pi\sigma^2}.$$

Therefore $\sum X_i$ is a sufficient, complete statistic and in addition

$$E\left[\sum X_i\right] = -n\frac{q'(\theta)}{p'(\theta)} = n\theta$$

which implies that $\overline{X}$ is the MVUE.

Ex Let X be a Poisson (θ), then
Supp $(S) = \{0, 1, 2, \ldots\}$ which does not depend on θ. In addition

$$f(x;\theta) = e^{-\theta}\frac{\theta^x}{x!} = e^{(\log\theta)x + \log\left(\frac{1}{x!}\right) + (-\theta)}$$

Hence this is a regular exponential family $K(x) = x$

Then $T = \sum X_i$ is a sufficient and complete statistic.

Moreover $E[T] = n\theta$ so $\bar{T} = \frac{T}{n}$ is actually also the MVUE.

$p(\theta) = \log\theta$, $q(\theta) = -\theta$. Thus

$$E[Y_1] = -n\frac{q'(\theta)}{p'(\theta)} = -n\frac{-1}{1/\theta} = n\theta$$

$$Var(T) = n \cdot \frac{1}{p'(\theta)^3}\left(p''(\theta)\cdot q'(\theta) - q''(\theta)p'(\theta)\right)$$

$$= n \cdot \frac{1}{\left(\frac{1}{\theta}\right)^3}\left(-\frac{1}{\theta^2}\cdot(-1)\right) = n\theta.$$

Appendix:

If $Y_1 = \sum_{i=1}^{n} K(X_i)$ is the statistic for an exponential family, the proof that Y_1 has density

$$f_{Y_1}(y_1;\theta) = H(y_1)\, e^{p(\theta) y_1 + n q(\theta)}$$

goes through a change of variable:

$$(x_1, x_2, \dots, x_n) \longrightarrow (y_1, x_2, \dots, x_n)$$

and this is enough to compute the joint pdf of $(y_1, x_2, \dots, x_n)$. Then from this integrating w.r.t $x_2, \dots, x_n$ we get the density of Y_1.

For the computation of $E[Y_1]$, we notice that

$$\int H(y_1)\, e^{p(\theta) y_1 + n q(\theta)}\, dy_1 = 1$$

so differentiating w.r.t. θ we get

$$\int \left(p'(\theta) y_1 + n q'(\theta)\right) e^{p(\theta) y_1 + n q(\theta)}\, dy_1 = 0$$

so $E[Y_1] = -\dfrac{n q'(\theta)}{p'(\theta)}$,

To get Var we need to differentiate this one more time.