Exponential Family of Densities:-

*** def:**

Suppose that X has a p.d.f. $f(x;\theta)$, $\theta \in \Omega$ where θ is a single parameter. A family of density $f(x;\theta)$ is called a member of one parameter exponential family if it can be written in the following form:-

$$f(x;\theta) = \exp\{T(x)\, g(\theta) + D(\theta) + S(x)\}$$

where $T(x)$, $g(\theta)$, $D(\theta)$ and $S(x)$ are suitable functions.

*** Example:**

Suppose X has a normal distribution with unknown mean μ and known variance σ_0^2. Then

$$f(x;\mu) = \frac{1}{\sqrt{2\pi}\,\sigma_0} e^{-\frac{1}{2\sigma_0^2}(x-\mu)^2}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_0} e^{-\frac{1}{2\sigma_0^2}(x^2-2x\mu+\mu^2)}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_0} e^{-\frac{x^2}{2\sigma_0^2}} \cdot e^{\frac{x\mu}{\sigma_0^2}} \cdot e^{-\frac{\mu^2}{2\sigma_0^2}}$$

$$\frac{1}{\sqrt{2\pi}\,\sigma_0} = e^{\ln\frac{1}{\sqrt{2\pi}\,\sigma_0}} = e^{\ln 1 - \ln\sqrt{2\pi}\,\sigma_0} = e^{-\ln\sqrt{2\pi}\,\sigma_0}$$

$$= \exp\left\{\frac{x\mu}{\sigma_0^2} - \frac{\mu^2}{2\sigma_0^2} - \left(\frac{x^2}{2\sigma_0^2} + \ln\sqrt{2\pi}\,\sigma_0\right)\right\}$$

Therefore, this distribution is a one-parameter exponential family with:

$T(x) = x$, $g(\theta) = \frac{\mu}{\sigma_0^2}$, $D(\theta) = \frac{-\mu^2}{2\sigma_0^2}$ and $S(x) = -\left(\frac{x^2}{2\sigma_0^2} + \ln\sqrt{2\pi}\,\sigma_0\right)$

***Example:**

Suppose X has a binomial distribution with parameters n and θ, $0<\theta<1$. Then

$$P(x;\theta) = C^n_x \, \theta^x (1-\theta)^{n-x} \qquad x=0,1,2,\dots,n$$

$$= C^n_x \left(\frac{\theta}{1-\theta}\right)^x (1-\theta)^n$$

$$= \exp\left\{x \ln\left(\frac{\theta}{1-\theta}\right) + n\ln(1-\theta) + \ln C^n_x\right\}$$

Therefore, this distribution is a one parameter exponential family with

$T(x) = x$, $g(\theta) = \ln \dfrac{\theta}{1-\theta}$, $D(\theta) = n\ln(1-\theta)$ and

$S(x) = \ln C^n_x$

Remark:

Under the random sample $x_1, x_2, \dots, x_n$ if the p.d.f. $f(x;\theta)$ is a member of exponential family, then:-

$$P(x_1, x_2, \dots, x_n;\theta) = \exp\left\{g(\theta)\sum_{i=1}^{n} T(x_i) + D(\theta) + S(x)\right\}$$

Applying the Factorization theorem, we can prove that $\sum_{i=1}^{n} T(x_i)$ is a sufficient statistic.

Remark:-

We might note that one parameter exponential family can be generalized to the k-parameter exponential family.

*** Def:**

Suppose X has a p.d.f. $f(x;\theta)$, $\theta \in \Omega$, where θ may be a vector. A family of densities $f(x;\theta_1,\theta_2,\dots,\theta_k)$ is called a member of k-parameter exponential family if it can be written in the following form:-

$$f(x;\theta_1,\theta_2,\dots,\theta_k) = \exp\left\{\sum_{i=1}^{k} T_i(x)\, g_i(\theta) + D(\theta) + S(x)\right\}$$

where $T_i(x)$, $g_i(\theta)$, $D(\theta)$ and $S(x)$ are suitable function, and k depends on the number of the parameters.

*** Example:**

Let X has a normal distribution with mean μ and variance σ^2, where both are unknown. To show that the normal distribution is a member of exponential family.

The p.d.f is given by:-

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad -\infty < x < \infty$$

which it can be written as:-

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

$$= \exp\left\{-\ln\sqrt{2\pi}\,\sigma - \frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}\right\}$$

where $T_1(x) = x$, $T_2(x) = x^2$, $g_1(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$, $g_2(\mu, \sigma^2) = \frac{-1}{2\sigma^2}$

$D(\theta) = \frac{-\mu^2}{2\sigma^2} - \ln \sqrt{2\pi}\,\sigma$ and $S(x) = 1$

Hence the normal distribution is a member of exponential family.

Remark:

Under the random sample $x_1, x_2, \dots, x_n$, the p.d.f. $f(x;\theta)$. is a member of k-parameter exponential family, then:-

$$f(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) = \exp\left\{ \sum_{i=1}^{k} g_i(\theta) \sum_{j=1}^{n} T_i(x_j) + D(\theta) + S(x) \right\}$$

and so by Factorization theorem $\sum_{j=1}^{n} T_1(x_j), \dots,$ $\sum_{j=1}^{n} T_k(x_j)$ are jointly sufficient statistics for the parameters $\theta_1, \theta_2, \dots, \theta_k$.

* Example:

Let $f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x > 0$, what is the sufficient statistic for the parameter θ?.

solu:

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0$$

$$\ln f(x;\theta) = -\ln\theta - \frac{x}{\theta}$$

$$\Rightarrow f(x;\theta) = \exp\left\{-\ln\theta - \frac{x}{\theta}\right\}$$

وبأخذ الـ e للطرفين

$\therefore f(x;\theta)$ is a member of one-parameter exponential family .

To find the sufficient statistics for θ , we find

$$f(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

$$= \left(\frac{1}{\theta}\right)^n e^{-\frac{\sum x_i}{\theta}}$$

$$= \exp\left\{-n \ln\theta - \frac{\sum x_i}{\theta}\right\}$$

بالمقارنة مع صيغة العائلة الاسية ذات المعلمة الواحدة نلاحظ ان

$$\sum_{i=1}^{n} T(x_i) = \sum_{i=1}^{n} x_i$$

$\therefore \sum_{i=1}^{n} x_i$ is the sufficient statistics for θ.

* Example :

Let $x_1, x_2, \dots, x_n \sim$ poisson(θ) . What is the sufficient statistics for θ ?

Soln.

$$f(x;\theta) = \frac{e^{-\theta}\,\theta^{x}}{x!} \qquad x = 0, 1, 2, \dots$$

$$= \exp\{x \ln\theta - \theta - \ln x!\}$$

$\therefore f(x;\theta)$ is a member of one-parameter exponential family with

$T(x) = x$, $g(\theta) = \ln\theta$, $D(\theta) = -\theta$, $S(x) = \ln x!$

To find the sufficient statistic for θ we find the joint p.d.f. for x as follows:

$$f(x_1, x_2, \dots, x_n ; \theta) = \prod_{i=1}^{n} f(x_i ; \theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\, \theta^{x_i}}{x_i!}$$

$$= \prod_{i=1}^{n} e^{-\theta} \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} = e^{-n\theta} \frac{\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$$= \exp\left\{ \sum x_i \ln\theta - n\theta - \ln \prod_{i=1}^{n} x_i! \right\}$$

$\sum_{i=1}^{n} T(x_i) = \sum_{i=1}^{n} x_i$, $g(\theta) = \ln\theta$, $D(\theta) = -n\theta$, $S(x) = -\ln \prod_{i=1}^{n} x_i!$

$\therefore \sum_{i=1}^{n} x_i$ is a sufficient statistics for θ.

Example:

Let $x_1, x_2, \dots, x_n$ be a r.s. from a population with Beta distribution. What are the jointly sufficient statistics for the parameters of Beta distribution.

Solu:

$$f(x ; \alpha, \beta) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} & 0 < x < 1 \\ 0 & o.w. \end{cases}$$

$$\ln f(x ; \alpha, \beta) = \underbrace{\ln 1}_{=0} - \ln B(\alpha, \beta) + (\alpha-1)\ln x + (\beta-1)\ln(1-x)$$

وبأخذ الـ e للطرفين

$$f(x ; \alpha, \beta) = \exp\{-\ln B(\alpha, \beta) + (\alpha-1)\ln x + (\beta-1)\ln(1-x)\}$$

$\therefore f(x ; \alpha, \beta)$ is a member of exponential family since:

$T_1(x) = \ln x$, $T_2(x) = \ln(1-x)$, $g_1(\theta) = \alpha - 1$, $g_2(\theta) = \beta - 1$
$D(\theta) = -\ln B(\alpha,\beta)$ and $S(x) = \ln(1)$

$$f(x_1, x_2, \dots, x_n ; \alpha, \beta) = \prod_{i=1}^{n} f(x_i ; \alpha, \beta)$$

$$= \left(\frac{1}{B(\alpha,\beta)}\right)^n \prod_{i=1}^{n} x_i^{\alpha-1} (1-x_i)^{\beta-1}$$

$$= \left(\frac{1}{B(\alpha,\beta)}\right)^n \prod_{i=1}^{n} x_i^{\alpha-1} \prod_{i=1}^{n} (1-x_i)^{\beta-1}$$

$$\ln f(x_1, \dots, x_n ; \alpha, \beta) = -n \ln B(\alpha,\beta) + \ln(x_1^{\alpha-1} \cdot x_2^{\alpha-1} \cdots x_n^{\alpha-1})$$
$$+ \ln[(1-x_1)^{\beta-1} \cdot (1-x_2)^{\beta-1} \cdots (1-x_n)^{\beta-1}]$$

$$= -n \ln B(\alpha,\beta) + [(\alpha-1)\ln x_1 + (\alpha-1)\ln x_2 + \cdots$$
$$+ (\alpha-1)\ln x_n] + [(\beta-1)\ln(1-x_1) + (\beta-1)\ln(1-x_2) + \cdots +$$
$$(\beta-1)\ln(1-x_n)]$$

$$= -n \ln B(\alpha,\beta) + (\alpha-1)[\ln x_1 + \cdots + \ln x_n] + (\beta-1)$$
$$[\ln(1-x_1) + \cdots + \ln(1-x_n)]$$
$$= -n \ln B(\alpha,\beta) + (\alpha-1)\sum_{i=1}^{n} \ln x_i + (\beta-1)\sum_{i=1}^{n} \ln(1-x_i)$$

وبأخذ الـ e للطرفين

$$f(x_1, x_2, \dots, x_n ; \alpha, \beta) = \exp\left\{-n \ln B(\alpha,\beta) + (\alpha-1)\sum_{i=1}^{n} \ln x_i + (\beta-1)\sum_{i=1}^{n} \ln(1-x_i)\right\}$$

$\therefore \sum_{i=1}^{n} \ln x_i$ and $\sum_{i=1}^{n} \ln(1-x_i)$ are jointly sufficient statistics for α and β.