2.3 SUFFICIENCY, EXPONENTIAL FAMILIES AND CONJUGACY.

DEFINITION (SUFFICIENCY).

A STATISTIC $t(X)$ (POSSIBLY MULTIVARIATE) IS SAID TO BE SUFFICIENT FOR $X$ FOR LEARNING ABOUT $\theta$ IF WE CAN WRITE:

$$f(x|\theta) = g(t, \theta) L(x) \quad (*)$$

WHERE $g(t, \theta)$ DEPENDS UPON $t(X)$ AND $\theta$ AND $L(x)$ DOES NOT DEPEND UPON $\theta$ BUT MAY DEPEND UPON $x$.

EQUIVALENT STATEMENTS TO (*) ARE:

i) $f(x|t, \theta)$ DOES NOT DEPEND UPON $\theta$.

ii) $f(\theta|x, t)$ DOES NOT DEPEND UPON $X$. [i.e. $\theta \perp x | t$]

[$f(\theta|x, t) = f(\theta|t) \Rightarrow (\theta \perp x | t)$].

IDEA IS THAT GIVEN $t(X)$ NOTHING FURTHER CAN BE LEARNED ABOUT $\theta$ FROM OBSERVING $X$.

DEFINITION ($k$-PARAMETER EXPONENTIAL FAMILY).

A PROBABILITY DENSITY $f(x|\theta)$ FOR $\theta = (\theta_1, ..., \theta_k)$ IS SAID TO BELONG TO THE $k$-PARAMETER EXPONENTIAL FAMILY IF IT IS OF THE FORM

$$f(x|\theta) = \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) u_j(x) + g(\theta) + L(x) \right\}$$
WHERE $\phi(\theta) = (\phi_1(\theta), \ldots, \phi_k(\theta))$ [k functions of $\theta$]$$

$u(x) = (u_1(x), \ldots, u_k(x))$ [k functions of $x$]$$

THE FAMILY IS REGULAR IF THE SAMPLE SPACE OF $x$ DOES NOT DEPEND UPON $\theta$, OTHERWISE IT IS NON-REGULAR.

EXAMPLE (BERNOULLI DISTRIBUTION).$$

$f(x|\theta) = \theta^x (1-\theta)^{1-x} \quad x \in \{0, 1\}$

$= \exp \left\{ x \log \theta + (1-x) \log (1-\theta) \right\}$

$= \exp \left\{ \left( \frac{\log \theta}{1-\theta} \right) x + \log (1-\theta) \right\}$

A MEMBER OF THE REGULAR 1-PARAMETER EXPONENTIAL FAMILY WITH $$

\phi_1(\theta) = \frac{\log \theta}{1-\theta}, \quad u_1(x) = x, \quad g(\theta) = \log (1-\theta), L(x) = 0$$

THEN $\sum_{i=1}^{n} u_i(x_i) = \left[ n, \frac{\sum_{i=1}^{n} u_1(x_i)}{n}, \ldots, \frac{\sum_{i=1}^{n} u_k(x_i)}{n} \right]$ [EXCHANGEABLE LIKELIHOOD]
PROOF.

\[ \prod_{i=1}^{\infty} \mathcal{E} \mathcal{F}_{\text{FC}}(x_i | g_1, h_1, \phi, \theta) = \prod_{i=1}^{\infty} \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) w_j(x_i) + g(\theta) + l(x_i) \right\} \]

\[ = \exp \left\{ \sum_{i=1}^{\infty} \left( \sum_{j=1}^{k} \phi_j(\theta) w_j(x_i) + g(\theta) + l(x_i) \right) \right\} \]

\[ = \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) \left( \sum_{i=1}^{\infty} w_j(x_i) \right) + n g(\theta) + \sum_{i=1}^{\infty} l(x_i) \right\} \]

\[ = \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) \left( \sum_{i=1}^{\infty} w_j(x_i) \right) + n g(\theta) \right\} \exp \left\{ \sum_{i=1}^{\infty} l(x_i) \right\} \]

\[ = \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) \left( \sum_{i=1}^{\infty} w_j(x_i) \right) + n g(\theta) \right\} \]

\[ = \exp \left\{ \sum_{j=1}^{k} \phi_j(\theta) \left( \sum_{i=1}^{\infty} w_j(x_i) \right) \right\} \exp \left\{ \sum_{i=1}^{\infty} l(x_i) \right\} \]

FUNCTION OF \( t, \theta \) ONLY.

\[ t_n \] IS SUFFICIENT FOR \( x = (x_1, \ldots, x_n) \) FOR LEARNING ABOUT \( \theta \).

I SOMETIMES Omit EXPlicit Reference of \( n \) To focus on "Interesting" Functions of \( x_1, \ldots, x_n \).

Example (Bernoulli Distribution (continued))

\[ f(x_i | \theta) = \exp \left\{ \left( \log \frac{\theta}{1-\theta} \right) x_i + \log (1-\theta) \right\} \]

If \( x_1, \ldots, x_n \) is an exchangeable sequence with \( x_i \sim \text{Bernoulli}(\theta) \)

Then:

\[ t_n = \left[ n, \sum_{i=1}^{n} x_i, (x_1) \right] = \left[ n, \sum_{i=1}^{n} x_i \right] \]

is a sufficient statistic.
2.3.1 EXPONENTIAL FAMILIES AND CONJUGATE PRIORS.

IF \( f(x; \Theta) \) IS A MEMBER OF A (REGULAR) \( k \)-PARAMETER EXPONENTIAL FAMILY, IT IS EASY TO SEE THAT A CONJUGATE PRIOR CAN BE FOUND IN THE \((k+1)\)-PARAMETER EXPONENTIAL FAMILY OVER \( \Theta \).

[Note: \( x = (x_1, \ldots, x_n) \) possible]

\[
\begin{align*}
  f(x; \Theta) &= \exp \left\{ \sum_{j=1}^{k} \phi_j(\Theta) y_j(x) + g(\Theta) + h(x) \right\} \\
  \text{REGARD AS A FUNCTION OF } \Theta,
\end{align*}
\]

\[
= \exp \left\{ \sum_{j=1}^{k} y_j(x) \phi_j(\Theta) + g(\Theta) + h(x) \right\}
\]

[EXCHANGEABLE: \( \exp \left\{ \sum_{j=1}^{k} \left( \frac{1}{n} \sum_{i=1}^{n} y_j(x_i) \right) \phi_j(\Theta) + n g(\Theta) + \sum_{i=1}^{n} h(x_i) \right\} \)]

TAKE \( f(\Theta) = \exp \left\{ \sum_{j=1}^{k} a_j \phi_j(\Theta) + d g(\Theta) + c(a,d) \right\} \)

WHERE \( c(a,d) = -\log \int_{\Theta} \exp \left\{ \sum_{j=1}^{k} a_j \phi_j(\Theta) + d g(\Theta) \right\} \, d\Theta \)

IS THE NORMALISING CONSTANT AND \( a = (a_1, \ldots, a_k) \) SO,

\[
\begin{align*}
  f(x; \Theta) &= \exp \left\{ \sum_{j=1}^{k} (a_j + y_j(x)) \phi_j(\Theta) + (d+1) g(\Theta) + c(\tilde{a}, \tilde{d}) \right\} \\
  \text{WHERE } (\tilde{a}, \tilde{d}) \text{ IS THE NORMALISING CONSTANT},
\end{align*}
\]

\[
\tilde{a}_j = a_j + y_j(x), \quad \tilde{d} = d+1
\]

Thus, \( f(x; \Theta) \) AND \( f(\Theta) \) ARE FROM THE SAME FAMILY GIVING (CONJUGATE.