A probability density function or probability mass function is called an exponential family if it can be expressed as

\[ f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right). \]

Note: \(h(x), t_1(x), \ldots, t_k(x)\) do not depend on \(\theta\) and \(c(\theta)\) does not depend on \(x\).

Example: Consider \(X \sim b(n, p)\) with \(n\) fixed. Show that \(p(x) = \binom{n}{x}p^x(1-p)^{n-x}\) can be expressed in the exponential family form.

\[
\begin{align*}
p(x) &= \binom{n}{x}p^x(1-p)^{n-x} \\
&= \binom{n}{x}\left(\frac{p}{1-p}\right)^x (1-p)^n \\
&= \binom{n}{x}(1-p)^n e^{x\log\left(\frac{p}{1-p}\right)} \\
&= \binom{n}{x}(1-p)^n e^{x\log\left(\frac{p}{1-p}\right)}
\end{align*}
\]

Therefore this pmf is an exponential family with \(h(x) = \binom{n}{x}, c(p) = (1-p)^n, t_1(x) = x, w_1(p) = \log\frac{p}{1-p}\).

Theorem: Suppose a random variable \(X\) has a pdf or pmf that can be expressed in the form of exponential family. Then,

\[
\begin{align*}
(a) \quad &E\left(\sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j}t_i(x)\right) = -\frac{\partial}{\partial \theta_j}\log c(\theta). \\
\text{and} \quad &\text{var}\left(\sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j}t_i(x)\right) = -\frac{\partial^2}{\partial \theta_j^2}\log c(\theta) - E\left(\sum_{i=1}^{k} \frac{\partial^2 w_i(\theta)}{\partial \theta_j^2}t_i(x)\right).
\end{align*}
\]

Note: Here log is the natural logarithm.
Proof of (a):

\[
\int_{x} f(x|\theta) dx = 1 \\
\int_{x} h(x) c(\theta) \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx = 1
\]

Differentiate both sides w.r.t. \( \theta_j \):

\[
\int_{x} h(x) \frac{\partial c(\theta)}{\partial \theta_j} \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx \\
+ \int_{x} h(x) c(\theta) \sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(x) \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx = 0
\]

Multiply the first integral by \( \frac{c(\theta)}{c(\theta)} \) and note that \( \frac{\partial \log c(\theta)}{\partial \theta_j} = \frac{\partial c(\theta)}{\partial \theta_j} \frac{1}{c(\theta)} \).

\[
\int_{x} h(x) \frac{\partial c(\theta)}{\partial \theta_j} \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) c(\theta) \frac{c(\theta)}{c(\theta)} dx \\
+ \int_{x} h(x) c(\theta) \sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(x) \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx = 0
\]

After rearranging we get

\[
\int_{x} h(x) c(\theta) \sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(x) \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx = \\
- \frac{\partial \log c(\theta)}{\partial \theta_j} \int_{x} h(x) c(\theta) \exp \left( \sum_{i=1}^{k} w_i(\theta)t_i(x) \right) dx
\]

Or

\[
E \left( \sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(x) \right) = - \frac{\partial}{\partial \theta_j} \log c(\theta).
\]

To prove statement (b) of the theorem differentiate a second time and rearrange.