The Hammersley-Clifford theorem

Richard D. Gill
Mathematical Institute, University of Leiden, Netherlands
http://www.math.leidenuniv.nl/∼gill

October 4, 2011

Suppose X_v, $v \in \mathcal{V}$ is a finite collection of discrete random variables with strictly positive joint probability mass function p. Choose a fixed reference value x^* and define for all $A \subseteq \mathcal{V}$

$$\psi_A(x_A) = \log p(x_A, x^*_{A^c}),$$

$$\phi_A(x_A) = \sum_{B:B\subseteq A} (-1)^{|A\setminus B|} \psi_B(x_B).$$

By the Möbius inversion lemma (please prove it yourself!), we can invert the relationship between the ϕ and the ψ functions evaluated at x to obtain for all B

$$\psi_B(x_B) = \sum_{A\subseteq B} \phi_A(x_A),$$

and in particular,

$$\log p(x) = \psi_{\mathcal{V}}(x) = \sum_{A\subseteq \mathcal{V}} \phi_A(x_A).$$

We will show that under the pairwise local Markov property, $\phi_A = 0$ if A is not a complete subset of $\mathcal{V}$. If A is not complete, there exist points α, β in A such that $\alpha \not\sim \beta$. Recall that $\phi_A(x_A) = \sum_{B:B\subseteq A}(-1)^{|A\setminus B|}\psi_B(x_B)$. Define $C = A \setminus \{\alpha, \beta\}$. We can now write

$$\phi_A(x_A) = \sum_{B:B\subseteq C} (-1)^{|A\setminus B|} \Big(\psi_B(x_B) - \psi_{B\cup\{\alpha\}}(x_{B\cup\{\alpha\}}) - \psi_{B\cup\{\beta\}}(x_{B\cup\{\beta\}}) + \psi_{B\cup\{\alpha,\beta\}}(x_{B\cup\{\alpha,\beta\}})\Big).$$

Now, for given B define $D = \mathcal{V} \setminus (B \cup \{\alpha, \beta\})$. It follows that

$$\begin{aligned}
&\psi_B(x_B) - \psi_{B\cup\{\alpha\}}(x_{B\cup\{\alpha\}}) - \psi_{B\cup\{\beta\}}(x_{B\cup\{\beta\}}) + \psi_{B\cup\{\alpha,\beta\}}(x_{B\cup\{\alpha,\beta\}}) \\
&= \log\Big(\frac{p(x_B, x_\alpha, x_\beta, x^*_D)p(x_B, x^*_\alpha, x^*_\beta, x^*_D)}{p(x_B, x^*_\alpha, x_\beta, x^*_D)p(x_B, x_\alpha, x^*_\beta, x^*_D)}\Big) \\
&= \log\Big(\frac{p(x_B, x_\alpha, x_\beta, x^*_D)/p(x_B, x^*_\alpha, x_\beta, x^*_D)}{p(x_B, x_\alpha, x^*_\beta, x^*_D)/p(x_B, x^*_\alpha, x^*_\beta, x^*_D)}\Big).
\end{aligned}$$

Now the last expression is the logarithm of the ratio of the conditional odds on $X_\alpha = x_\alpha$ against $X_\alpha = x^*_\alpha$, under the conditions $X_B = x_B$, $X_D = x^*_D$ and $X_\beta = x_\beta$ and under the conditions $X_B = x_B$, $X_D = x^*_D$ and $X_\beta = x^*_\beta$. Thus if X_α is independent of X_β conditional on $X_{B\cup D}$, this log odds ratio is zero.