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 \( A \subseteq \mathcal{V} \)

\[
\psi_A(x_A) = \log p(x_A, x_A^*),
\]

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

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

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

and in particular,

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

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 \( \alpha, \beta \) 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

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

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

\[
\psi_B(x_B) - \psi_{B \cup \{\alpha\}}(x_{B \cup \{\alpha\}}) - \psi_{B \cup \{\beta\}}(x_{B \cup \{\beta\}}) + \psi_{B \cup \{\alpha, \beta\}}(x_{B \cup \{\alpha, \beta\}}) = \log \left( \frac{p(x_B, x_{\alpha}, x_{\beta}, x_D') p(x_B, x_{\alpha}', x_{\beta}', x'_D)}{p(x_B, x_{\alpha}, x_{\beta}, x_D') p(x_B, x_{\alpha}', x_{\beta}', x_D')} \right).
\]

Now the last expression is the logarithm of the ratio of the conditional odds on \( X_\alpha = x_\alpha \) against \( X_\beta = x_\beta \), 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 \( X_\alpha \) is independent of \( X_\beta \) conditional on \( X_{B \cup D} \), this log odds ratio is zero.