GLOBAL TESTING AGAINST SPARSE ALTERNATIVES
UNDER ISING MODELS

BY RAJARSHI MUKHERJEE‡, SUMIT MUKHERJEE§, AND MING YUAN¶

University of California, Berkeley‡ and Columbia University §¶

APPENDIX – PROOF OF AUXILIARY RESULTS

Proof of Lemma 1. This is a standard application of Stein’s Method for concentration inequalities (Chatterjee, 2005). The details are included here for completeness. One begins by noting that

\[ E_{Q, \mu} (X_i | X_j, j \neq i) = \tanh (m_i (X) + \mu_i), \quad m_i (X) := \sum_{j=1}^{n} Q_{ij} X_j. \]

Now let \( X \) be drawn from (1) and let \( X' \) is drawn by moving one step in the Glauber dynamics, i.e. let \( I \) be a random variable which is discrete uniform on \( \{1, 2, \cdots, n\} \), and replace the \( I^{th} \) coordinate of \( X \) by an element drawn from the conditional distribution of the \( I^{th} \) coordinate given the rest. It is not difficult to see that \( (X, X') \) is an exchangeable pair of random vectors. Further define an anti-symmetric function \( F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \) as \( F(x, y) = \sum_{i=1}^{n} (x_i - y_i) \), which ensures that

\[ E_{Q, \mu} (F(X, X') | X) = \frac{1}{n} \sum_{j=1}^{n} X_j - \tanh (m_j (X) + \mu_j) = f_{\mu}(X). \]

Denoting \( X^i \) to be \( X \) with \( X_i \) replaced by \(-X_i\), by Taylor’s series we have

\[ \tanh(m_j(X^i) + \mu_j) - \tanh(m_j(X) + \mu_j) \]
\[ = (m_j(X^i) - m_j(X))g'(m_j(X)) + \frac{1}{2}(m_j(X^i) - m_j(X))^2g''(\xi_{ij}) \]
\[ = -2Q_{ij}X_ig'(m_j(X)) + 2Q_{ij}^2g''(\xi_{ij}) \]

§The research of Sumit Mukherjee was supported in part by NSF Grant DMS-1712037.
¶The research of Ming Yuan was supported in part by NSF FRG Grant DMS-1265202, and NIH Grant 1-U54AI117924-01.
for some \( \{\xi_{ij}\}_{1 \leq i,j \leq n} \), where \( g(t) = \tanh(t) \). Thus \( f_{\mu}(X) - f_{\mu}(X^i) \) can be written as

\[
\begin{align*}
    f_{\mu}(X) - f_{\mu}(X^i) &= \frac{2X_i}{n} + \frac{1}{n} \sum_{j=1}^{n} \left\{ \tanh \left( m_j (X^i) + \mu_j \right) - \tanh \left( m_j (X) + \mu_j \right) \right\} \\
    &= \frac{2X_i}{n} - \frac{2X_i}{n} \sum_{j=1}^{n} Q_{ij} g'(m_j(X)) + \frac{2}{n} \sum_{j=1}^{n} Q_{ij}^2 g''(\xi_{ij})
\end{align*}
\]

Now setting \( p_i(X) := \mathbb{P}(Q_{ij}(X^i) = -X_i | X_k, k \neq i) \) we have

\[
\begin{align*}
    v(X) &= \frac{1}{2} \mathbb{E}_{Q_{ij}} \left( |f_{\mu}(X) - f_{\mu}(X^i)||X_I - X'_I||X \right) \\
    &= \frac{1}{n} \sum_{i=1}^{n} |f_{\mu}(X) - f_{\mu}(X^i)|p_i(X) \\
    &\leq \frac{2}{n^2} \sum_{i=1}^{n} p_i(X) - \frac{2}{n^2} \sum_{i,j=1}^{n} |Q_{ij} p_i(X) g'(m_j(X))| \\
    &\quad + \frac{2}{n^2} \sum_{i,j=1}^{n} Q_{ij}^2 g''(\xi_{ij})^2 X_i p_i(X) \\
    &\leq \frac{2}{n} + \frac{2}{n^2} \sup_{u,v \in [0,1]^n} |u^T Q v| + \frac{2}{n^2} \sum_{i,j=1}^{n} Q_{ij}^2,
\end{align*}
\]

where in the last line we use the fact that \( \max(|g'(t)|, |g''(t)|) \leq 1 \). The proof of the Lemma is then completed by an application of Theorem 3.3 of Chatterjee (2007).

**Proof of Lemma 2.** Let \( Y := (Y_1, \cdots, Y_n) \) be i.i.d. random variables on \( \{-1, 1\} \) with \( \mathbb{P}(Y_i = \pm 1) = \frac{1}{2} \), and let \( W := (W_1, \cdots, W_n) \) i.i.d. \( N(0,1) \). Also, for any \( t > 0 \) let \( Z(tQ, \mu) \) denote the normalizing constant of the p.m.f.

\[
\frac{1}{Z(tQ, \mu)} \exp \left( \frac{1}{2} x^T tQx + \mu^T x \right)
\]

Thus we have

\[
2^{-n} Z(tQ, \mu) = \mathbb{E} \exp \left( \frac{t}{2} Y^T QY + \sum_{i=1}^{n} \mu_i Y_i \right) \leq \mathbb{E} \exp \left( \frac{t}{2} W^T QW + \sum_{i=1}^{n} \mu_i W_i \right),
\]

where we use the fact that \( \mathbb{E} Y_i^k \leq \mathbb{E} W_i^k \) for all positive integers \( k \). Using spectral decomposition write \( Q = P^T \Lambda P \) and set \( \nu := P^T \mu, \tilde{W} = PW \) to
note that

\[\text{Exp}\left(\frac{t}{2}W^\top QW + \sum_{i=1}^{n} \mu_i W_i\right) = \text{Exp}\left(\frac{t}{2} \sum_{i=1}^{n} \lambda_i \tilde{W}_i^2 + \sum_{i=1}^{n} \nu_i \tilde{W}_i\right) = \prod_{i=1}^{n} e^{\frac{\nu_i^2}{1-t\lambda_i}}.\]

Combining for any \(t > 1\) we have the bounds

\[(1) \quad 2^n \prod_{i=1}^{n} \cosh(\mu_i) = Z(0, \mu) \leq Z(Q, \mu) \leq Z(tQ, \mu) \leq 2^n \prod_{i=1}^{n} e^{\sum_{i=1}^{n} \frac{\nu_i^2}{1-t\lambda_i}},\]

where the lower bound follows from noting that \(\log Z(tQ, \mu)\) is monotone non-decreasing in \(t\), using results about exponential families. Thus invoking convexity of the function \(t \mapsto \log Z(tQ, \mu)\) we have

\[E_{Q, \mu} \frac{1}{2} X^\top QX = \left. \frac{\partial \log Z(tQ, \mu)}{\partial t} \right|_{t=1} \leq \frac{\log Z(tQ, \mu) - \log Z(Q, \mu)}{t-1} \leq \sum_{i=1}^{n} \left\{ \frac{\nu_i^2}{2(1-t\lambda_i)} - \log \cosh(\mu_i) \right\} - \sum_{i=1}^{n} \frac{1}{2} \log(1-t\lambda_i),\]

where we use the bounds obtained in (1). Proceeding to bound the rightmost hand side above, set \(t = \frac{1+\rho}{2}\rho > 1\) and note that

\[|t\lambda_i| \leq \frac{1+\rho}{2} < 1.\]

For \(x \in \frac{1}{2}[-(1+\rho), (1+\rho)] \subseteq (-1, 1)\) there exists a constant \(\gamma_\rho < \infty\) such that

\[\frac{1}{1-x} \leq 1 + x + 2\gamma_\rho x^2, \quad -\log(1-x) \leq x + 2\gamma_\rho x^2.\]

Also a Taylor’s expansion gives

\[-\log \cosh(x) \leq -\frac{x^2}{2} + x^4,\]

where we have used the fact that \(||(\log \cosh(x))^{(4)}||_\infty \leq 1.\) These, along with the observations that

\[\sum_{i=1}^{n} \lambda_i = tr(Q) = 0, \quad \sum_{i=1}^{n} \nu_i^2 = ||P\mu||^2 = ||\mu||^2\]
give the bound
\[
\sum_{i=1}^{n} \left\{ \frac{\nu_i^2}{2(1-t\lambda_i)} - \log \cosh(\mu_i) \right\} - \sum_{i=1}^{n} \frac{1}{2} \log(1-t\lambda_i)
\leq \left\{ \frac{1}{2} \sum_{i=1}^{n} \mu_i^2 + \frac{t}{2} \sum_{i=1}^{n} \nu_i^2 \lambda_i \right\} + \left\{ -\frac{1}{2} \sum_{i=1}^{n} \mu_i^2 + \sum_{i=1}^{n} \mu_i^4 \right\} + \gamma_\rho t^2 \sum_{i=1}^{n} \lambda_i^2
\]
\[
= \frac{t}{2} \mu^\top Q \mu + t^2 \gamma_\rho \mu^\top Q^2 \mu + \sum_{i=1}^{n} \mu_i^4 + \gamma_\rho t^2 \sum_{i,j=1}^{n} Q_{ij}^2
\leq \frac{t}{2} C_\rho \sqrt{n} + t^2 \gamma_\rho C_\rho^2 \sqrt{n} + C \sqrt{n} + \gamma_\rho t^2 D \sqrt{n},
\]
where \( D > 0 \) is such that \( \sum_{i,j=1}^{n} Q_{ij}^2 \leq D \sqrt{n} \). This along with (1) gives
\[
\left[ \frac{1}{2} C(1 + t\rho) + t^2 \gamma_\rho C_\rho^2 + C + \gamma_\rho t^2 D \right] \sqrt{n} \geq \frac{1}{2} E_{Q,\mu} X^\top Q X = \frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} X_i m_i(X)
\]
But, for some random \((\xi_i, i = 1, \ldots, n)\)
\[
\frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} X_i m_i(X)
\]
\[
= \frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} \tanh(m_i(X)) + \mu_i m_i(X)
\]
\[
= \frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} \tanh(m_i(X))) m_i(X) + \frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} \mu_i m_i(X) \sech^2(\xi_i).
\]
Now,
\[
\frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} \tanh(m_i(X))) m_i(X) \geq \eta \frac{1}{2} E_{Q,\mu} \sum_{i=1}^{n} m_i(X)^2,
\]
where
\[
\eta := \inf_{|x| \leq 1} \frac{\tanh(x)}{x} > 0.
\]
The desired conclusion of the lemma follows by noting that
\[
\left| E_{Q,\mu} \sum_{i=1}^{n} \mu_i m_i(X) \sech^2(\xi_i) \right| \leq C \sqrt{n}.
\]
PROOF OF LEMMA 3. We begin with Part (a). By a simple algebra, the p.m.f. of $X$ can be written as
\[
\mathbb{P}_{\theta,\mu}(X = x) \propto \exp \left\{ \frac{n\theta}{2} x^2 + \sum_{i=1}^{n} x_i \mu_i \right\}.
\]
Consequently, the joint density of $(X, Z_n)$ with respect to the product measure of counting measure on $\{-1, 1\}^n$ and Lebesgue measure on $\mathbb{R}$ is proportional to
\[
\exp \left\{ \frac{n\theta}{2} (z - \bar{x})^2 - \frac{n\theta}{2} z^2 + \sum_{i=1}^{n} x_i (\mu_i + z \theta) \right\}.
\]
Part (a) follows from the expression above.

Now consider Part (b). Using the joint density of Part (a), the marginal density of $Z_n$ is proportional to
\[
\sum_{x \in \{-1, 1\}^n} \exp \left\{ -\frac{n\theta}{2} z^2 + \sum_{i=1}^{n} \log \cosh(\mu_i + z \theta) \right\} = e^{-f_{\theta,\mu}(z)},
\]
thus completing the proof of Part (b).

Finally, consider Part (c). By Part (a) given $Z_n = z$ the random variables $(X_1, \cdots, X_n)$ are independent, with
\[
\mathbb{P}_{\theta,\mu}(X_i = 1|Z_n = z) = \frac{e^{\mu_i + \theta z}}{e^{\mu_i + \theta z} + e^{-\mu_i - \theta z}},
\]
and so
\[
\mathbb{E}_{\theta,\mu}(X_i|Z_n = z) = \tanh(\mu_i + \theta z), \quad \mathbb{V}_{\theta,\mu}(X_i|Z_n = n) = \text{sech}^2(\mu_i + \theta z).
\]
Thus for any $\mu \in [0, \infty)^n$ we have
\[
\mathbb{E}_{\theta,\mu} \left\{ \left( \sum_{i=1}^{n} (X_i - \tanh(\mu_i + \theta Z_n)) \right)^2 \right\} = \mathbb{E} \sum_{i=1}^{n} \text{sech}^2(\mu_i + \theta Z_n) \leq n.
\]
\[\Box\]
Proof of Lemma 4. We begin with Part (a). Since
\[ f''_{n,\mu}(z) = \sum_{i=1}^{n} \tanh^2(z + \mu_i) \]
is strictly positive for all but at most one \( z \in \mathbb{R} \), the function \( z \mapsto f_{n,\mu}(z) \) is strictly convex with \( f_{n,\mu}(\pm \infty) = \infty \), it follows that \( z \mapsto f_{n,\mu}(z) \) has a unique minima \( m_n \) which is the unique root of the equation \( f'_{n,\mu}(z) = 0 \). The fact that \( m_n \) is positive follows on noting that
\[ f'_{n,\mu}(0) = -\sum_{i=1}^{n} \tanh(\mu_i) < 0, \quad f'_{n,\mu}(+\infty) = \infty. \]
Also \( f'(m_n) = 0 \) gives
\[ m_n = \frac{1}{n} \sum_{i=1}^{n} \tanh(m_n + \mu_i) \leq 1, \]
and so \( m_n \in (0,1] \). Finally, \( f'_{n,\mu}(m_n) = 0 \) can be written as
\[ m_n - \tanh(m_n) = \frac{s}{n} \left[ \tanh(m_n + B) - \tanh(m_n) \right] \geq C \frac{s}{n} \tanh(B), \]
for some \( C > 0 \), which proves Part (a).

Now consider Part (b). By a Taylor’s series expansion around \( m_n \) and using the fact that \( f''(z) \) is strictly increasing on \((0,\infty)\) gives
\[ f_n(z) \geq f_n(m_n) + \frac{1}{2} (z - m_n)^2 f''_n(m_n + Kn^{-1/4}) \text{ for all } z \in [m_n + Kn^{-1/4}, \infty) \]
\[ f_n(z) \leq f_n(m_n) + \frac{1}{2} (z - m_n)^2 f''_n(m_n + Kn^{-1/4}) \text{ for all } z \in [m_n, m_n + Kn^{-1/4}]. \]
Setting \( b_n := f''_n(m_n + Kn^{-1/4}) \) this gives
\[ \mathbb{P}_{\theta,\mu}(Z_n > m_n + Kn^{-1/4}) \]
\[ = \frac{\int_{m_n+Kn^{-1/4}}^{\infty} e^{-f_n(z)} dz}{\int_{\mathbb{R}} e^{-f_n(z)} dz} \]
\[ \leq \frac{\int_{m_n+Kn^{-1/4}}^{\infty} e^{-\frac{b_n}{2}(z-m_n)^2} dz}{\int_{m_n}^{m_n+Kn^{-1/4}} e^{-\frac{b_n}{2}(z-m_n)^2} dz} \]
\[ = \frac{\mathbb{P}(N(0,1) > Kn^{-1/4} \sqrt{b_n})}{\mathbb{P}(0 < N(0,1) < Kn^{-1/4} \sqrt{b_n})}. \]
from which the desired conclusion will follow if we can show that \( \lim \inf_{n \to \infty} n^{-1/2}b_n > 0 \). But this follows on noting that

\[
n^{-1/2}b_n = n^{-1/2}f''_n(m_n + Kn^{-1/4}) \geq \sqrt{n} \tanh^2(Kn^{-1/4}) = K^2\Theta(1).
\]

Finally, let us prove Part (c). By a Taylor’s series expansion about \( \delta m_n \) and using the fact that \( f_n(\cdot) \) is convex with unique global minima at \( m_n \) we have

\[
f_n(z) \geq f_n(m_n) + (z - \delta m_n)f'_n(\delta m_n), \quad \forall z \in (-\infty, \delta m_n].
\]

Also, as before we have

\[
f_n(z) \leq f_n(m_n) + \frac{1}{2}(z - m_n)^2f''_n(m_n), \quad \forall z \in [m_n, 2m_n].
\]

Thus with \( c_n := f''_n(2m_n) \) for any \( \delta > 0 \) we have

\[
\mathbb{P}_{\theta, \mu}(Z_n \leq \delta m_n) = \frac{\int_{-\infty}^{\delta m_n} e^{-f_n(z)}dz}{\int_{\mathbb{R}} e^{-f_n(z)}dz} \leq \frac{\int_{-\infty}^{\delta m_n} e^{-f_n(z)}dz}{\int_{m_n}^{2m_n} e^{-f_n'(\delta m_n)(z - m_n)^2}dz} \leq \frac{\sqrt{2\pi c_n}}{|f'_n(\delta m_n)|}\mathbb{P}(0 < Z < m_n \sqrt{c_n}).
\]

To bound the rightmost hand side of (2), we claim that the following estimates hold:

\[
(3) \quad c_n = \Theta(nm_n^2), \quad nm_n^3 = O(|f'_n(\delta m_n)|).
\]

Given these two estimates, we immediately have

\[
(5) \quad m_n \sqrt{c_n} = \Theta(m_n^2 \sqrt{n}) \geq \Theta(A_n^{2/3} \sqrt{n}) \to \infty,
\]

as \( A_n \gg n^{-3/4} \) by assumption. Thus the rightmost hand side of (2) can be bounded by

\[
\frac{m_n \sqrt{n}}{nm_n^3} = \frac{1}{m_n^2 \sqrt{n}} \to 0,
\]

where the last conclusion uses (5). This completes the proof of Part (c).
It thus remains to prove the estimates (3) and (4). To this effect, note that

\[ f''_n(2m_n) = \sum_{i=1}^{n} \tanh^2(2m_n + \mu_i) \leq \sum_{i=1}^{n} \left( \tanh(2m_n) + C_1 \tanh(\mu_i) \right)^2 \leq 2n \tanh^2(2m_n) + 2C_1^2 \sum_{i=1}^{n} \tanh^2(\mu_i) \lesssim n m_n^2 + nA(\mu_n) \lesssim n m_n^2, \]

where the last step uses part (a), and \(C_1 < \infty\) is a universal constant. This gives the upper bound in (3). For the lower bound of (3) we have

\[ f''_n(m_n) = \sum_{i=1}^{n} \tanh(2m_n + \mu_i) \geq n \tanh^2(2m_n) \gtrsim n m_n^3. \]

Turning to prove (4) we have

\[ |f'_n(\delta m_n)| = \sum_{i=1}^{n} \tanh(\delta m_n + \mu_i) - n \delta m_n \]
\[ = \left[ \sum_{i=1}^{n} \tanh(\delta m_n + \mu_i) - \tanh(\delta m_n) \right] - n[\delta m_n - \tanh(\delta m_n)] \geq C_2 nA(\mu_n) - C_3 n \delta^3 m_n^3 \gtrsim n m_n^3, \]

where \(\delta\) is chosen small enough, and \(C_2 > 0, C_3 < \infty\) are universal constants. This completes the proof of (4), and hence completes the proof of the lemma.

\[ \square \]

REFERENCES


Division of Biostatistics, Haviland Hall, Berkeley, CA- 94720. E-mail: rmukherj@berkeley.edu

Department of Statistics 1255 Amsterdam Avenue New York, NY-10027. E-mail: sm3949@columbia.edu E-mail: ming.yuan@columbia.edu