1 Stein’s method for concentration inequalities

The purpose of this lecture will be to prove the following theorem.

**Theorem 1** Suppose you have an exchangeable pair \((X, X')\) of random objects. Suppose \(f\) and \(F\) are two functions such that

1. \(F(X, X') = -F(X', X)\) a.s.; and
2. \(E[f(X, X') | X] = f(X)\) a.s.

Let \(v(X) = \frac{1}{2}E[|f(X) - f(X')|F(X, X') | X]\). Then

(a) \(E[f(X)] = 0\) and

\[
\text{Var} f(X) = \frac{1}{2}E[(f(X) - f(X'))F(X, X')] \leq E[v(X)].
\]  

(b) Suppose \(E(e^{\theta f(X)} | F(X, X'))\) is finite for all \(\theta\). If \(B\) and \(C\) are constants such that \(v(X) \leq B f(X) + C\) a.s., then

\[
P\{|f(X)| > t\} \leq 2 \exp\left(-\frac{t^2}{2Bt + 2C}\right).
\]

(c) \(E[f(X)^{2k}] \leq (2k - 1)^k E[v(X)^k]\) for all \(k \in \mathbb{N}\).

**Exercise 2** (You may be able to do this after the proof.)

Extend (c) to all real \(k > 1/2\). (We think that \((2k - 1)^k\) remains unchanged for \(k \geq 1\) but are not sure for \(1/2 < k < 1\).)
Proof: First, note that $E[f(X)] = E[F(X, X')] = 0$ since $(X, X')$ is an exchangeable pair and $F$ is antisymmetric.

Further, we will assume

$$E[v(X)] < \infty \text{ for part (a); and}$$

$$E[e^{\theta f(X)}|F(X, X')] < \infty \text{ for all } \theta \text{ in part (b).}$$

(3)

We start by showing (a).

$$\text{Var } f(X) = E[f(X)^2] - E[f(X)]^2$$

$$= E[f(X)F(X, X')]$$

$$= E[f(X')F(X', X)] \text{ since } (X, X') \text{ exchangeable}$$

$$= -E[f(X')F(X, X')] \text{ by antisymmetry of } F$$

$$= \frac{1}{2}E[(f(X) - f(X')) F(X, X')]$$

$$\leq E[v(X)],$$

which proves (a).

(b): Let $m(\theta) = E[e^{\theta f(X)}]$; then $m'(\theta) = E[f(X)e^{\theta f(X)}]$. Write this as

$$m'(\theta) = E[F(X, X')e^{\theta f(X)}]$$

$$= \frac{1}{2}E[F(X, X')\left(e^{\theta f(X)} - e^{\theta f(X')}\right)]$$

via the antisymmetry of $F$ and the exchangeability of $(X, X')$.

We’ll use the inequality

$$|e^x - e^y| \leq \frac{1}{2}|x - y|(e^x + e^y).$$

(5)

To see this, suppose $y < x$;

$$e^x - e^y = \int_0^1 \frac{d}{dt} e^{tx+(1-t)y} \, dt$$

$$= (x - y) \int_0^1 e^{tx+(1-t)y} \, dt$$

$$\leq (x - y) \int_0^1 (te^x + (1-t)e^y) \, dt \text{ by Jensen’s inequality}$$

$$= (x - y) \frac{1}{2} (e^x + e^y),$$

and similarly for $y \geq x$. 

12-2
Now

\[ |m'(\theta)| = \left| \mathbb{E}[F(X, X')e^{\theta f(X)}] \right| \]
\[ \leq \frac{1}{2} \left| \mathbb{E}[F(X, X') \left( e^{\theta f(X)} - e^{\theta f(X')} \right)] \right| \]
\[ \leq \frac{\theta}{4} \mathbb{E} \left[ |F(X, X') (f(X) - f(X')) | \left( e^{\theta f(X)} + e^{\theta f(X')} \right) \right] \]
\[ \leq \frac{\theta}{2} \mathbb{E} \left[ |F(X, X') (f(X) - f(X')) | e^{\theta f(X)} \right] \]
\[ \leq |\theta| \mathbb{E} \left[ v(X)e^{\theta f(X)} \right] \]
by exchangeability of \((X, X')\).

If \(v(X) \leq Bf(X) + C\), then the above gives

\[ |m'(\theta)| \leq |\theta| \left( B \mathbb{E} \left[ f(X)e^{\theta f(X)} \right] + C \mathbb{E} \left[ e^{\theta f(X)} \right] \right) \]
\[ = B|\theta|m'(\theta) + C|\theta|m(\theta). \]

Now, \(m\) is a convex function, with \(\mathbb{E}[f(X)] = m'(0) = 0\) and taking the value \(m(0) = 1\) at its minimum. Suppose \(0 < \theta < 1/B\); then

\[ m'(\theta)(1 - B\theta) \leq C\theta m(\theta). \]

Thus

\[ \frac{d}{d\theta} \log m(\theta) = \frac{m'(\theta)}{m(\theta)} \leq \frac{C\theta}{1 - B\theta} \]
for \(0 < \theta < 1/B\). Therefore,

\[ \log m(\theta) = \int_0^\theta \frac{d}{dt} \log m(t) \, dt \]
\[ \leq \int_0^\theta \frac{Ct}{1 - Bt} \, dt \]
\[ \leq \frac{1}{1 - B\theta} \int_0^\theta Ct \, dt \]
\[ = \frac{C\theta^2}{2(1 - B\theta)}. \]

So, for \(\theta > 0\),

\[ \mathbb{P}\{f(X) \geq t\} = \mathbb{P}\left\{ e^{\theta f(X)} \geq e^{\theta t} \right\} \]
\[ \leq e^{-\theta t}m(\theta) \]
\[ \leq \exp \left( -\theta t + \frac{C\theta^2}{2(1 - B\theta)} \right) \text{ if } 0 < \theta < 1/B. \]
Taking \( \theta = \frac{t}{C + Bt} \in \left( 0, \frac{1}{B} \right) \), we get the desired upper bound, and \( P(f(X) \leq -t) \) can be bounded similarly.

(c): Using a similar manipulation to that in the proof of (a),

\[
E\left[ f(X)^{2k} \right] = E\left[ f(X)^{2k-1} F(X, X') \right] = \frac{1}{2} E\left[ \left( f(X)^{2k-1} - f(X')^{2k-1} \right) F(X, X') \right].
\]

Also, similarly to (5), we can show

\[
\left| x^{2k-1} - y^{2k-1} \right| \leq \frac{2k - 1}{2} \left| x - y \right| \left| x^{2k-2} + y^{2k-2} \right|.
\]

Plugging this into (6), we get

\[
E\left[ f(X)^{2k} \right] \leq (2k - 1) E\left[ v(X) f(X)^{2k-2} \right].
\]

Applying Hölder’s inequality with \( 1/p = 1 - 1/k \) and \( 1/q = 1/k \) gives

\[
E\left[ f(X)^{2k} \right] \leq (2k - 1) \left( E\left[ f(X)^{2k} \right] \right)^{(k-1)/k} \left( E\left[ v(X)^k \right] \right)^{1/k}
\]

and so

\[
\left( E\left[ f(X)^{2k} \right] \right)^{1/k} \leq (2k - 1) \left( E\left[ v(X)^k \right] \right)^{1/k}.
\]

This completes the proof of (c).

**Example 3** Suppose we have the Curie-Weiss model on \( n \) spins: for \( \sigma = (\sigma_1, \ldots, \sigma_n) \),

\[
P\{\sigma\} = \frac{1}{Z_\beta} \exp\left( \frac{\beta}{n} \sum_{1 \leq i < j \leq n} \sigma_i \sigma_j \right).
\]

Let

\[
m(\sigma) = \frac{1}{n} \sum_i \sigma_i.
\]

Construct \( X' \) by taking one step in the Gibbs sampler (also known as the Glauber dynamics). Set \( F(X, X') = \sigma_I - \sigma'_I \) where \( I \) is the updated index. Then

\[
f(X) = E[F(X, X') \mid X] \approx m(\sigma) - \tanh(\beta m(\sigma)).
\]

We’ll do this in the next lecture.