Proof of the Bounded Differences Inequality
Peter Bartlett. March 9, 2006.

**Theorem** Suppose that \(X_1, \ldots, X_n \in \mathcal{X}\) are independent, and \(f : \mathcal{X}^n \to \mathbb{R}\). Let \(c_1, \ldots, c_n\) satisfy

\[
\sup_{x_1, \ldots, x_n, x'_i} |f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x'_i, x_{i+1}, \ldots, x_n)| \leq c_i,
\]

for \(i = 1, \ldots, n\). Then

\[
P(f - \mathbb{E}f \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2}\right).
\]

**Proof:** We define

\[
V_i = \mathbb{E}[f|X_1, \ldots, X_i] - \mathbb{E}[f|X_1, \ldots, X_{i-1}].
\]

These \(V_i\)s will play the same role as that played by the terms of the sum in the proof of Hoeffding’s inequality. In particular, since the sum telescopes, we have

\[
f - \mathbb{E}f = \sum_{i=1}^n V_i.
\]

Using this, and the Chernoff bounding technique, we see that

\[
P(f - \mathbb{E}f \geq t) = P\left(\sum_{i=1}^n V_i \geq t\right) \leq \inf_{s > 0} \exp(-st) \mathbb{E}\exp\left(s\sum_{i=1}^n V_i\right) = \inf_{s > 0} \exp(-st) \mathbb{E}\left(\prod_{i=1}^n e^{sV_i}\right).
\]
So we need to bound the moment generating function of this sum of (dependent) random variables. Now, we have seen that \( E(V_i | X_1, \ldots, X_{i-1}) = 0 \). If we define

\[
L_i = \inf_x E(f | X_1, \ldots, X_{i-1}, x) - E(f | X_1, \ldots, X_{i-1}),
\]

\[
U_i = \sup_x E(f | X_1, \ldots, X_{i-1}, x) - E(f | X_1, \ldots, X_{i-1}),
\]

then we have

\[
V_i - L_i = E(f | X_1, \ldots, X_i) - E(f | X_1, \ldots, X_{i-1})
\]

\[
- \inf_x E(f | X_1, \ldots, X_{i-1}, x) + E(f | X_1, \ldots, X_{i-1})
\]

\[
= E(f | X_1, \ldots, X_i) - \inf_x E(f | X_1, \ldots, X_{i-1}, x),
\]

and so \( L_i \leq V_i \) almost surely. Similarly, \( V_i \leq U_i \) a.s. Furthermore, from the independence of the \( X_i \),

\[
U_i - L_i = \sup_x E(f | X_1, \ldots, X_{i-1}, x) - \inf_x E(f | X_1, \ldots, X_{i-1}, x)
\]

\[
= \sup_x \int f(X_1, \ldots, X_{i-1}, x, x_{i+1}, \ldots, x_n) dP(x_{i+1}, \ldots, x_n)
\]

\[
- \inf_x \int f(X_1, \ldots, X_{i-1}, x, x_{i+1}, \ldots, x_n) dP(x_{i+1}, \ldots, x_n)
\]

\[
= \sup_{x,y} \int (f(X_1, \ldots, X_{i-1}, x, x_{i+1}, \ldots, x_n))
\]

\[
- f(X_1, \ldots, X_{i-1}, y, x_{i+1}, \ldots, x_n)) dP(x_{i+1}, \ldots, x_n)
\]

\[
\leq c_i,
\]

from the bounded differences assumption. Notice that the independence allows us to write the difference of the conditional expectations in terms an integral of a difference, and hence appeal to the bounded differences property. Thus, we may apply the Hoeffding lemma to the bounded, zero-mean random
variables $V_i$ conditioned on $X_1, \ldots, X_{i-1}$, as follows.

$$
\mathbb{E} \left( \prod_{i=1}^{n} e^{sV_i} \right) = \mathbb{E} \mathbb{E} \left( \prod_{i=1}^{n-1} e^{sV_i} e^{sV_n} | X_1, \ldots, X_{n-1} \right) \\
= \mathbb{E} \left( \prod_{i=1}^{n-1} e^{sV_i} \mathbb{E} \left( e^{sV_n} | X_1, \ldots, X_{n-1} \right) \right) \\
\leq \exp(s^2 c_n^2 / 8) \mathbb{E} \left( \prod_{i=1}^{n-1} e^{sV_i} \right) \\
= \exp(s^2 c_n^2 / 8) \mathbb{E} \left( \prod_{i=1}^{n-2} e^{sV_i} | X_1, \ldots, X_{n-2} \right) \\
= \exp(s^2 c_n^2 / 8) \mathbb{E} \left( \prod_{i=1}^{n-2} e^{sV_i} \mathbb{E} \left( e^{sV_{n-1}} | X_1, \ldots, X_{n-2} \right) \right) \\
\leq \exp(s^2 (c_{n-1}^2 + c_n^2) / 8) \mathbb{E} \left( \prod_{i=1}^{n-2} e^{sV_i} \right) \\
= \exp \left( s^2 \sum_{i=1}^{n} c_i^2 / 8 \right).
$$

Substituting, we have

$$
P(f - \mathbb{E} f \geq t) \leq \inf_{s>0} \exp \left( -st + \frac{s^2 \sum_{i=1}^{n} c_i^2}{8} \right) \\
= \exp \left( -\frac{2t^2}{\sum_{i=1}^{n} c_i^2} \right),
$$

where we selected the optimizing value $s = 4t / \sum_{i=1}^{n} c_i^2$. 

3