Advanced probability and applications

Azuma’s & McDiarmid’s inequalities

Azuma’s inequality

This inequality is a direct generalization of Hoeffding’s inequality for martingales. The statement is the following:

Let \((M_n, n \in \mathbb{N})\) be a martingale w.r.t. a filtration \((\mathcal{F}_n, n \in \mathbb{N})\) such that \(|M_{n+1}(\omega) - M_n(\omega)| \leq 1\) for all \(n \in \mathbb{N}\) and \(\omega \in \Omega\).

Then
\[
P(\{|M_n - M_0| > n\epsilon\}) \leq 2 \exp\left(-n \epsilon^2/2\right) \quad \forall \epsilon > 0
\]

(This is exactly Hoeffding’s inequality when \(M_n\) is a sum of iid, r.v. with zero mean.)
Proof

Let $X_n = \Pi_n - \Pi_{n-1}$ for $n \geq 1$. Then $\Pi_n - \Pi_0 = \sum_{j=1}^{n} X_j$, but the $X_j$ are not necessarily independent here.

Because of the martingale property and the extra assumption, we know that $\mathbb{E}(X_j | \mathcal{F}_{j-1}) = 0$ & $|X_j| \leq 1$ $\forall j \geq 1$.

Proceeding as in the proof of Hoeffding's inequality, let us compute:

\[
P\left( \sum_{j=1}^{n} X_j > nt \right) \leq \frac{\mathbb{E} \left( \exp \left( s (\Pi_n - \Pi_0) \right) \right)}{\exp(snt)}
\]

\[
e^{-snt} \cdot \mathbb{E} \left( \exp \left( s \sum_{j=1}^{n} X_j \right) \right) = e^{-snt} \cdot \mathbb{E} \left( \mathbb{E} \left( \exp \left( s \sum_{j=1}^{n} X_j \right) | \mathcal{F}_{n-1} \right) \right)
\]

\[
e^{-snt} \cdot \mathbb{E} \left( \exp \left( s \sum_{j=1}^{n-1} X_j \right) \cdot \mathbb{E} \left( \exp(sX_n) | \mathcal{F}_{n-1} \right) \right)
\]

\[
\text{measurable}
\]

\[
(1)
\]
What do we know about the term (1)?
As said before, \( E(X_n | \mathcal{F}_{n-1}) = 0 \) & \( |X_n(\omega)| \leq 1 \ \forall \omega \in \mathcal{\Omega} \nabla \\
We can therefore apply the lemma seen in the proof of Hoeffding's inequality to conclude that
\[
E(e^{sX_n} | \mathcal{F}_{n-1}) \leq \exp\left(\frac{s^2}{2}\right)
\]
(the only change here is the conditioning w.r.t. \( \mathcal{F}_{n-1} \))
We therefore obtain:
\[
P(\{ \Pi_n - \Pi_0 > n/3 \}) \leq e^{-sn/3} \cdot E\left(\exp\left(\frac{s \sum_{j=1}^{n-1} X_j}{n}\right)\right) \cdot \exp\left(\frac{s^2}{2}\right)
\]
Proceeding recursively in the same manner, we obtain
\[
P(\{ \Pi_n - \Pi_0 > n/3 \}) \leq e^{-sn/3} \cdot \left(\exp\left(\frac{s^2}{2}\right)\right)^n = \exp\left(-n(s/n - \frac{s^2}{2})\right)
\]
leading, after optimization over \( s \geq 0 \), to the same conclusion as before.
McDiarmid's inequality

The previous inequality, which looks like an innocent and direct generalization of an inequality about sums of iid r.v., allows in fact to derive a much more general inequality than Hoeffding's inequality as a corollary:

Let $n \geq 1$ be fixed, $X_1, \ldots, X_n$ be iid random variables and $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be such that

$$|F(x_1, \ldots, x_j, \ldots, x_n) - F(x_1, \ldots, x_j', \ldots, x_n)| \leq 1 \quad \forall x_1, \ldots, x_j, x_j', \ldots, x_n \in \mathbb{R}$$

Then

$$P\left(\left|F(x_1, \ldots, x_n) - E(F(x_1, \ldots, x_n))\right| > n^{1/2}\right) \leq 2 \exp\left(-n^{1/2}\right) \quad \forall n$$
Proof

Let \( F_0 = \{ \emptyset, \Omega \} \), \( F_j = \sigma(\mathbf{X}_1, \ldots, \mathbf{X}_j) \) \( 1 \leq j \leq n \)

\[ M_j = \mathbb{E}(F(\mathbf{X}_1, \ldots, \mathbf{X}_n) | F_j) \quad 0 \leq j \leq n \]

By definition, \( \Pi \) is a (Doob) martingale and

\[ \Pi_n = F(\mathbf{X}_1, \ldots, \mathbf{X}_n) \quad \Pi_0 = E(F(\mathbf{X}_1, \ldots, \mathbf{X}_n)) \]

Moreover, we have:

\[ |\Pi_j - \Pi_{j-1}| = |\mathbb{E}(F(\mathbf{X}_1, \ldots, \mathbf{X}_n) | F_j) - \mathbb{E}(F(\mathbf{X}_1, \ldots, \mathbf{X}_n) | F_{j-1})| \]

\[ = |g(\mathbf{X}_1, \ldots, \mathbf{X}_j) - h(\mathbf{X}_1, \ldots, \mathbf{X}_{j-1})| \]

where \( g(\mathbf{X}_1, \ldots, \mathbf{X}_j) = \mathbb{E}(F(\mathbf{X}_1, \ldots, \mathbf{X}_{j-1}, \mathbf{X}_j, \mathbf{X}_{j+1}, \ldots, \mathbf{X}_n)) \)

and \( h(\mathbf{X}_1, \ldots, \mathbf{X}_{j-1}) = \mathbb{E}(F(\mathbf{X}_1, \ldots, \mathbf{X}_{j-1}, \mathbf{X}_j, \mathbf{X}_{j+1}, \ldots, \mathbf{X}_n)) \)
Given the assumption made on the function $F$, it holds that

$$
\left| q(x_1, \ldots, x_j) - h(x_1, \ldots, x_{j-1}) \right|
\leq \mathbb{E} \left( | F(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) - F(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) | \right)
\leq \mathbb{E}(1) = 1 \quad \forall x_1, \ldots, x_{j-1}, x_j \in \mathbb{R}
$$

So $| \Pi_j(w) - \Pi_{j-1}(w) | \leq 1$ and by Azuma's inequality, we obtain:

$$
P \left( \left\{ \left| \frac{\Pi_j(w) - \Pi_0(w)}{\Pi_j(w) - \Pi_0(w)} \right| > n \varepsilon \right\} \right) \leq 2 \exp \left( - \frac{n \varepsilon^2}{2} \right)
$$