Final Project:Summary of Methos of Concentration Inequalities

YiQi Xu

Advisor
Dr. Yang Yun

Abstract

In the lecture of concentration inequality, we use martingale method(see the surveys of McDiarmid [2,3]) to show the proof of centration inequalities bound tail probabilities of general functions of independent random variables. There are several other methods have been known to prove such inequalities, including information-theoretic methods (see Alhswede, GaÌĄcs and KoÌĹrner, Marton, Dembo, Massart and Rio), TalagrandâĂŹs induction method , and various problem-specific methods; see Janson, ÅĄuczak and Rucin ÌĄski for a survey. A novel way of deriving powerful inequalities, the âĂIJentropy method,âĂİ(see the surveys of Boucheron[1]) based on logarithmic Sobolev inequalities, was developed by Ledoux, Bobkov and Ledoux, Massart, Rio and Bousquet for proving sharp concentration bounds for maxima of empirical processes. Recently Boucheron, Lugosi and Massart pointed out that the methodology may be used effectively outside of the context of empirical process theory as well.

In this project, I will review the martingale method first, then discuss the entrophy-based method.

Also, by reviewing the relationship of the new results with some of the existing work, we may find some other inequalities may be recovered easily from some of the new inequalities, for example, TalagrandâĂŹs âĂIJconvex distanceâĂİ inequality.

We also will show some application of new inequalities.

CHAPTER 1

Concentration inequality

1.1 Martingale method

RCALL MARTINGALE METHOD In this section, we will recall the maqrtingale method known as a general method to proof concentration inequality.

1.1.1 *Theorem*

Consider a martingale difference sequence D_k (adapted to a filtration $\mathscr{F}_k$) that satisfies

$$E[\exp(\lambda D_k)|\mathscr{F}_k] \leq \exp(\lambda^2 \nu_k^2/2), \text{a.s.for all } |\lambda| \leq 1/b_k$$

Proof:$E[\exp(\lambda D_k)] = E[E[\exp(\lambda D_k)|\mathscr{F}_k]] \leq E[\exp(\lambda^2 \nu_k^2/2)] = \exp(\lambda^2 \nu_k^2/2)$, Then $\sum_{i=1}^n D_k$ is sub-exponential with parameters$(\nu^2, b) = (\sum_{i=1}^n \nu_k^2, max_k b_k)$,and

$$P(|\sum_{i=1}^n D_k| \leq t) \leq \begin{cases} 2\exp(-\frac{t^2}{2\nu^2}) & \text{if } 0 \leq t \geq \frac{\nu^2}{b} \\ 2\exp(-\frac{t}{2b}) & \text{if } t > \frac{\nu^2}{b} \end{cases} \tag{1.1}$$

1.1.2 *Theorem(Azuma-Hoeffding)*

Consider a martingale difference sequence D_k with $|D_k| \leq B_k$ a.s. then

$$P(|\sum_{k=1}^n D_k| \leq t) \leq 2\exp(-\frac{2t^2}{\sum_k B_k^2}) \tag{1.2}$$

Note: Since $|D_k| \leq B_k$, According to bounded random variable $P(|D_k| \leq B_k) = 1$, D_k is a sub-Gaussian r.v with $\sigma^2 = B_k^2$. According to Hoeffding Bound, $P(|\sum_{k=1}^n D_k| \leq t) \leq 2\exp(-\frac{2t^2}{\sum_k B_k^2})$

1.1.3 *Theorem(Bounded difference inequality)*

Suppose function $f : R^n \to R$ satisfies the bounded difference property:for all $x_1, ..., x_n, x_k' \in R$,

$$|f(x_1, ..., x_n) - f(x_1, ..., x_{k-1}, x_k', x_{k+1}, ..., x_n)| \leq L_k$$

Then for $X = (X_1, ..., X_n)$ with independent components,

$$P(|f(X) - E[f(X)]| \geq t) \leq 2\exp(-\frac{2t^2}{\sum_k L_k^2}) \tag{1.3}$$

Note: set $D_k = E[f(x)|x_1, ..., x_k] - E[f(x)|x_1, ..., x_{k-1}] = E[f(x_1, ..., x_n)|x_1, ..., x_k] - E[f(x_1, ..., x_k', ..., x_n)|x_1, ..., x_{k-1}, x_k] \leq L_k, \sum D_k = E[f(x)|x_1, ..., x_n] - E[f(x)] = f(x) - E[f(x)]$

1.2 ENTROPY METHOD

In this paper, Dr.Boucheron presents a novel way "entropy method" which based on logarithmic Sobolev inequalities to derive concentration inequalities. These inequalities may be considered as exponential versions of the well-known Efron-Stein inequalty.

1.2.1 *Notation*

First, I will introduce the notation used in this paper. $X_1, ..., X_n$ are random variables which belong to a measurable space χ, and X_1^n represent vector of these n random variables, $f : \chi^n \to R$ is measurable function

$$Z = f(X_1, ..., X_n)$$

$$Z^{(i)} = f(X_1, ..., X_{i-1}, X_{i'}, X_{i+1}, ..., X_n)$$

1.2.2 *Efron-Stein inequality*

These inequalities may be considered as exponential versions of the well-known Efron-Stein inequality. And our goal is to find out under which conditions the Efron-Stein inequality can be converted into exponential upper bounds for either $P[Z > EZ + t]$ or $P[Z < EZ - t]$ where $t > 0$

$$Var(Z) \leq \frac{1}{2} E\left[\sum_{i=1}^{n} (Z - Z^{(i)})^2\right] \tag{1.4}$$

Define the random variables V_+ and V_- by:

$$V_+ = E\left[\sum_{i=1}^{n} (Z - Z^i)^2 I_{Z > Z^{(i)}} | X_1^n\right]$$

$$V_- = E\left[\sum_{i=1}^{n}(Z-Z^i)^2 I_{Z<Z^{(i)}}|X_1^n\right]$$

Then Efron-Stein inequality may be rewritten as

$$Var(Z) \leq E[V_+] = E[V_-] \tag{1.5}$$

1.2.3 *Theorem and Corollary*

In this paper, Dr.Boucheron presents 6 theorems and 2 corollaries.

1.2.3.1 Theorem 1

For all $\theta > 0$ and $\lambda \in (0, 1/\theta)$

$$\lambda E[Z\exp^{\lambda Z}] - E[\exp^{\lambda Z}]\log E[\exp^{\lambda Z}] \leq \frac{1}{2}\sum_{i=1}^{n} E[\exp^{\lambda Z}\,\psi(-\lambda(Z-Z^{(i)}))1_{Z>Z^{(i)}}] \tag{1.6}$$

On the other hand, we have for all $\theta > 0$ and $\lambda \in (0, 1/\theta)$

$$\lambda E[Z\exp^{\lambda Z}] - E[\exp^{\lambda Z}]\log E[\exp^{\lambda Z}] \leq \frac{1}{2}\sum_{i=1}^{n} E[\exp^{\lambda Z}\,\psi(\lambda(Z^{(i)}-Z))1_{Z<Z^{(i)}}] \tag{1.7}$$

Proof: Using Logarithmic Sobolev inequalities to derive inequality in Theorem 1.

Logarithmic Sobolev inequalities: For any function $f : \chi^n \to R$, nothing $Z = f(X_1, ..., X_n)$ and for all $\lambda \in R$

$$\lambda E[Ze^{\lambda Z}] - E[e^{\lambda Z}]\log E[e^{\lambda Z}] \leq \sum_{i=1}^{n} E\left[e^{\lambda Z}\psi(-\lambda(Z-Z^{(i)}))I_{Z>Z^{(i)}}\right] \tag{1.8}$$

$$\lambda E[Ze^{\lambda Z}] - E[e^{\lambda Z}]\log E[e^{\lambda Z}] \leq \sum_{i=1}^{n} E\left[e^{\lambda Z}\psi(\lambda(Z^{(i)}-Z))I_{Z<Z^{(i)}}\right] \tag{1.9}$$

where $\psi(x) = x(e^x - 1)$ Let $\lambda \geq 0$ and introduce $F(\lambda) = E[\exp(\lambda Z)]$, Inequality(1.8) implies:

$$\lambda F'(\lambda) - F(\lambda)\log F(\lambda) \leq \sum_{i=1}^{n}\lambda^2 E\left[e^{\lambda Z}(Z-Z^{(i)})^2 I_{Z>Z^{(i)}}\right] = \lambda^2 E[V_+ e^{\lambda Z}]$$

Lemma:

$$\frac{E[\lambda W e^{\lambda Z}}{E[e^{\lambda Z}]} \leq \frac{E[\lambda Z e^{\lambda Z}]}{E[e^{\lambda Z}]} - \log E[e^{\lambda Z}] + \log E[e^{\lambda W}] \tag{1.10}$$

Let $G(\lambda) = \log E[\exp(\lambda V_+)]$ and $F'(\lambda) = E[Z\exp(\lambda Z)]$, applying lemma with $W = V_+/\theta$, obtain

$$\lambda F'(\lambda) - F(\lambda)\log F(\lambda) \le \lambda^2\theta\Big(F'(\lambda) + \frac{1}{\lambda}F(\lambda)G(\lambda/\theta) - \frac{1}{\lambda}F(\lambda)\log F(\lambda)\Big)$$

Deviding both sides by $\lambda^2 F(\lambda)$

$$\frac{1}{\lambda}\frac{F'(\lambda)}{F(\lambda)} - \frac{1}{\lambda^2}\log F(\lambda) \le \frac{\theta G(\lambda/\theta)}{\lambda(1-\lambda\theta)}$$

Since left-hand side is the derivative of $H(\lambda) = (1/\lambda)\log F(\lambda)$ and $H(\lambda) \to E[Z]$ as $\lambda \to 0$, then

$$H(\lambda) \le E[Z] + \int_0^\lambda \frac{\theta G(s/\theta)}{s(1-s\theta)}ds$$

which proving Inequality(1.6)

$$\log F(\lambda) \le \lambda E[Z] + \frac{\lambda\theta G(\lambda/\theta)}{1-\lambda\theta}$$

Inequality(1.7) follows similarly by replacing Z by -Z and noting that a bound on V_- for Z is equivalent to a bound on V_+ for -Z.

1.2.3.2 Corollary 1

Assume that three exists a positive constant c such that, almost surely, $V_+ \le c$. Then, for all $t > 0$

$$P[Z > E[Z] + t] \le e^{-t^2/4c} \tag{1.11}$$

Moreover, if $V_- \le c$ almost surely, then for all $t > 0$,

$$P[Z < E[Z] - t] \le e^{-t^2/4c} \tag{1.12}$$

Proof: Inequality(1.6) implies

$$\log E\left[\exp(\lambda(Z - E[Z]))\right] \le \frac{\lambda\theta}{1-\lambda\theta}\frac{\lambda c}{\theta}$$

Thus, letting θ approach zero, we obtain

$$E\left[\exp(\lambda(Z - E[Z]))\right] \le e^{\lambda^2 c}$$

Hence, by Markov's inequality, for all $\lambda > 0$ and $t > 0$

$$P[Z > E[Z] + t] \le e^{\lambda^2 c - \lambda t}$$

1.2.3.3 Corollary 2

Assume that the random variable V_+ is such that there exists a positive constant a such that for $\lambda \in (0, 1/a)$,

$$\log E[e^{\lambda(V_+ - E[V_+])}] \leq \frac{\lambda^2 a E[V_+]}{1 - a\lambda}$$

Then

$$\log E\left[\exp(\lambda(Z - E[Z]))\right] \leq \frac{\lambda^2 E[V_+]}{1 - (a+1)\lambda} \tag{1.13}$$

in particular

$$P[Z > E[Z] + t] \leq \exp\left(\frac{-t^2}{4E[V_+] + 2(a+1)t/3}\right) \tag{1.14}$$

Proof: Taking $\theta = 1$ in Inequality(1.6) and using the condition on the moment generating function of V_+, we obtain

$$\log E\left[\exp(\lambda(Z - E[Z]))\right] \leq \frac{\lambda}{1-\lambda}\left(\lambda E[V_+] + \frac{\lambda^2 a E[V_+]}{1 - a\lambda}\right) = \frac{\lambda^2 E[V_+]}{(1-\lambda)(1-a\lambda)} \leq \frac{\lambda^2 E[V_+]}{1 - (a+1)\lambda}$$

1.2.3.4 Theorem 2

Assume that there exist positive constants a and b such that

$$V_+ \leq aZ + b \tag{1.15}$$

Then, for $\lambda \in (0, 1/a)$,

$$\log E\left[\exp(\lambda(Z - E[Z]))\right] \leq \frac{\lambda^2}{1 - a\lambda}(aE[Z] + b) \tag{1.16}$$

and for all $t > 0$,

$$P[Z > E[Z] + t] \leq \exp\left(\frac{-t^2}{4aE[Z] + 4b + 2at}\right) \tag{1.17}$$

Proof: Using first steps of proof of Theorem 1 and Assumption(1.15)

$$\lambda E[Ze^{\lambda Z}] - E[e^{\lambda Z}]\log E[e^{\lambda Z}] \leq E[e^{\lambda Z} V_+] \leq \lambda^2(aE[Ze^{\lambda Z}] + bE[\lambda Z])$$

$$which is \lambda F'(\lambda) - F(\lambda)\log F(\lambda) \leq a\lambda^2 F'(\lambda) + b\lambda^2 F(\lambda)$$

Dividing both sides by $\lambda^2 F(\lambda)$

$$H'(\lambda) \leq a(\log F(\lambda))' + b$$

Since $H(0) = F'(0)/F(0) = E[Z]$ and $\log F(0) = 0$, we obtain

$$H(\lambda) \leq E[Z] + a\log F(\lambda) + b\lambda$$

or, if $\lambda < 1/a$,

$$\log E\left[\lambda(Z - E[Z])\right] \leq \frac{\lambda^2}{1 - a\lambda}(aE[Z] + b)$$

1.2.3.5 Theorem 3

Assuming that for some non-decreasing function g,

$$V_- \leq g(Z) \tag{1.18}$$

Then, for all $t > 0$,

$$P[Z < E[Z] - t] \leq \exp\left(\frac{-t^2}{4E[g(Z)]}\right) \tag{1.19}$$

Proof: By the Inequality(1.8),

$$\lambda E[Ze^{\lambda Z}] - E[e^{\lambda Z}]\log E[e^{\lambda Z}] \leq \sum_{i=1}^{n} E\left[e^{\lambda Z}\psi(\lambda(Z^{(i)} - Z))I_{Z<Z^{(i)}}\right]$$
$$\leq \sum_{i=1}^{n} E\left[e^{\lambda Z}\lambda^2(Z^{(i)} - Z)^2 I_{Z<Z(i)}\right] = \lambda^2 E[e^{\lambda Z}V_-] \leq \lambda^2 E[e^{\lambda Z}g(Z)]$$

Chebyshev's association inequality implies that

$$E[e^{\lambda Z}g(Z)] \leq E[e^{\lambda Z}]E[g(Z)]$$

Deviding both sides of inequality by $\lambda^2 F(\lambda)$, we obtain

$$F(\lambda) \leq \exp(\lambda^2 E[g(Z)] + \lambda E[Z])$$

1.2.3.6 Theorem 4

Assume that there exists a nondecreasing function g shuch that $V_+ \leq g(Z)$ and for any value of X_1^n and X_i', $|Z - Z^{(i)}| \leq 1$. Then, for all $K > 0$, $\lambda \in [0.1/K]$

$$\log E\left[\exp(-\lambda(Z - E[Z]))\right] \leq \lambda^2 \frac{\psi(K)}{K^2} E[g(Z)] \tag{1.20}$$

and for all $t > o$, with$t \leq (e-1)E[g(Z)]$ we have

$$P[Z < E[Z] - t] \leq \exp\left[-\frac{t^2}{4(e-1)E[g(Z)]}\right] \tag{1.21}$$

Proof: Inequality(1.8) implies that

$$\lambda E[Ze^{\lambda Z}] - E[e^{\lambda Z}]\log E[e^{\lambda Z}] \leq \sum_{i=1}^{n} E\left[e^{\lambda Z}\psi(-\lambda(Z - Z^{(i)}))I_{Z>Z^{(i)}}\right]$$
$$\leq \lambda^2 \frac{\psi(K)}{k^2} E\left[e^{\lambda Z}\sum_{i=1}^{n}(Z - Z^{(i)})^2 I_{Z>Z^{(i)}}\right] \leq \lambda^2 \frac{\psi(K)}{K^2} E[g(Z)e^{\lambda Z}]$$

1.2.3.7 Theorem 5

Assuming that f is non-negative. Assume that there exists a random variable W, such that

$$V_+ \leq WZ \tag{1.22}$$

Then, for all $\theta > 0$ and $\lambda \in (0.1/\theta)$,

$$\log E\left[\exp(\lambda(\sqrt{Z} - E[\sqrt{Z}]))\right] \leq \frac{\lambda\theta}{1-\lambda\theta} \log E\left[\exp(\frac{\lambda W}{\theta})\right] \tag{1.23}$$

Writing $x = \sqrt{E[Z]+t} - \sqrt{E[Z]}$, we have. for $\lambda > 0$,

$$P[Z > E[Z]+t] \leq P\left[\sqrt{Z} > E[\sqrt{Z}] + x\right] \leq E[\exp(\lambda(\sqrt{Z} - E[\sqrt{Z}]))]e^{-\lambda x} \tag{1.24}$$

Proof: Define $Y = \sqrt{Z}$ and $Y^{(i)} = \sqrt{Z^{(i)}}$, then,

$$E\left[\sum_i (Y - Y^{(i)})^2 I_{Y > Y^{(i)}} | X_1^n\right] = E\left[\sum_i (\sqrt{Z} - \sqrt{Z^{(i)}})^2 I_{Z > Z^{(i)}} | X_1^n\right]$$
$$\leq E\left[\sum_i \left(\frac{Z - Z^{(i)}}{\sqrt{Z}}\right)^2 I_{Z > Z^{(i)}} | X_1^n\right] \leq \frac{1}{Z} E\left[\sum_n (Z - Z^{(i)})^2 I_{Z > Z(i)} | X_1^n\right] \leq W$$

1.2.3.8 Theorem 6

Assuming that f is nonnegative. Assume that there exist constants $a > 0$ and $\alpha \in (0, 2)$ such that $V_+ \leq \alpha Z^\alpha$. Then, for all $\lambda > 0$

$$\log E\left[\exp(\lambda(Z^{(2-\alpha)/2} - E[Z^{(2-\alpha)/2}]))\right] \leq \lambda^2 a \tag{1.25}$$

and for all $\lambda \in (0.1/a)$,

$$\log E\left[\exp(\lambda(Z^{2-\alpha}) - E[Z^{2-\alpha}])\right] \leq \frac{\lambda^2 a E[Z^{2-\alpha}]}{1 - \alpha\lambda} \tag{1.26}$$

Proof: For any $p > 0$

$$E\left[\sum_i (Z^p - (Z^{(i)})^p)^2 I_{Z > Z^{(i)}} | X_1^n\right] = E\left[\sum_i \left(\frac{Z}{Z^{1-p}} - \frac{Z^{(i)}}{Z^{(i)1-p}}\right)^2 I_{Z > Z^{(i)}} | X_1^n\right]$$
$$\leq \frac{1}{Z^{2-2p}} E\left[\sum_i (Z - Z^{(i)})^2 I_{Z > Z^{(i)}} | X_1^n\right] \leq aZ^{\alpha+2p-2}$$

1.3 RELATION TO PREVIOUS RESULT

1.3.1 *Corollary 1 and Bounded difference inequality*

we will see that Corollary 1 implies the bounded difference inequality

1.3.2 *Theorem 2 and Configuration function bound*

(Configuration function bound) Assume that $f : \chi \to R^+$ is a configuration function, that is, for all $\mathbf{x} = (x_1, .., x_n) \in \chi^n$, there exists a set $I \in 1, ..., n$ of indices such that $f(\mathbf{x}) = |I|$ and for all $\mathbf{y} = (y_1, ..., y_n) \in \chi^n$, $f(\mathbf{y}) \geq \sum_{i \in I} I_{x_i = y_i}$. Then, for all $t > 0$, if MZ denote the median of Z,

$$P[Z \geq MZ + t] \leq 2\exp\left[-\frac{t^2}{4MZ + 4t}\right] \tag{1.27}$$

$$P[Z \geq MZ - t] \leq 2\exp\left[-\frac{t^2}{4MZ}\right] \tag{1.28}$$

(Corollary 3) Assume that f is nonnegative and that there exists a function g: $\chi^{n-1} \to R^+$ such that, for all $x_1, ..., x_n \in \chi$:
(i) $0 \leq f(x_1, ..., x_n) - g(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n) \leq 1$ for all $i = 1, ..., n$;
(ii) $\sum_{i=1}^{n} [f(x_1, ..., x_n) - g(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)] \leq f(x_1, ..., x_n)$
Then, for any $t > 0$,

$$P[Z \geq E[Z] + t] \leq \exp\left[-\frac{t^2}{2E[Z] + 2t/3}\right] \tag{1.29}$$

$$P[Z \leq E[Z] - t] \leq \exp\left[-\frac{t^2}{2E[Z]}\right] \tag{1.30}$$

Easy to see that configuration functions satisfy the condition of Corollary 3, and so Corollary 3 is a generalization of the configuration function bound. Also, a function f satisfies the condition of Corollary 3 satisfies the condition of Theorem 2 as well with $b = 0$ and $a = 1$. Thus, Theorem 2 may be considered as a generalization of Corollary.

1.4 APPLICATION

1.4.1 *Rademacher averages and chaos*

(Rademacher average):

$$Z = E\left[||\sum_{i=1}^{n} \varepsilon_i X_i|| \mid X_1^n\right]$$

where the ε_i are the independent centered $1. - 1$-valued random variables.

1.4.1.1 Theorem 7

For any $t > 0$,

$$P[Z \geq E[Z] + t] \leq \exp\left[-\frac{t^2}{2E[Z] + 2t/3}\right] \tag{1.31}$$

and

$$P[Z \leq E[Z] - t] \leq \exp\left[-\frac{t^2}{2E[Z]}\right] \tag{1.32}$$

(Rademacher chaos):
First, define random variable

$$Z = \sup_{M \in \mathscr{F}} \sum_{i,j \leq n} \varepsilon_i \varepsilon_j M(i,j)$$

where $\mathscr{F}$ denotes a collection $n * n$ symmetric matrices M, and $\varepsilon_i, \ldots, \varepsilon_n$ are i.i.d Rademacher variables.

1.4.1.2 Theorem 8

For all $t > 0$

$$PZ \geq E[Z] + t \leq \exp\left(-\frac{t^2}{32E[Y^2] + 65t/3}\right) \tag{1.33}$$

where the random variable Y is defined as

$$Y = \sup_{M \in \mathscr{F}} \left(\sum_{i=1}^{n} (\sum_{j=1}^{n} \varepsilon_j M(i,j))^2\right)^{1/2} \tag{1.34}$$

1.4.1.3 Theorem 9

There exists a universal constant K such that

$$P[Z + E[Z] + t] \leq 2\exp\left(-\frac{1}{K} min(t, \frac{t^2}{E[Z] + E^2[Y]})\right) \tag{1.35}$$

Note: Easy to attain the theorem by consequence of Theorem 4.

1.4.2 *Minimum of the empirical risk*

Define the empirical risk:

$$L_n(f) = \frac{1}{n} \sum_{i=1}^{n} \ell(f(X_i), Y_i)$$

and minimal empirical risk:

$$\widehat{L} = \inf_{f \in \mathscr{F}} L_n(f)$$

supremum of the empirical process

$$W = n \sup_{f \in \mathscr{F}} |L_n(f) - E[L_n(f)]|$$

and introduce

$$v = E\left[\sup_{f \in \mathscr{F}} \sum_{i=1}^{n} (\ell(f(X_i), Y_i) - \ell(f(X_i'), Y_i'))^2\right]$$

1.4.2.1 Theorem 10

Let $a = \frac{33}{8} + \frac{(e-1)v}{8nE[\widehat{L}]+4E[W]}$, Then, for all $t > 0$

$$P[\widehat{L} > E[widehatL] + t] \leq \exp\left(-\frac{nt^2}{25E[\widehat{L}]/2 + 25E[W]/(4n) + 2at/3}\right) \tag{1.36}$$

and for $t \in [0, E[\widehat{L}])$

$$P[\widehat{L} > E[widehatL] - t] \leq \exp\left(-\frac{nt^2}{4E[\widehat{L}] + 4E[W]/n + 4tv(e-1)/(nE[\widehat{L}] + E[W])}\right) \tag{1.37}$$

1.5 SMMARY

Dr.Boucheron uses "entropy method" to generate concentration inequality and applies the results into application like rademacher averages and chaos and minimum of the empirical risk. We find it can be some previous work' generalization, although in some cases, other method's results are better.

Bibliography

[1] STÉPHANE BOUCHERON, GÁBOR LUGOSI, and PASCAL MASSART. "CONCENTRATION INEQUALITIES USING THE ENTROPY METHOD". In: *The Annals of Probability* 31.3 (2003), pp. 1583–1614.

[2] MCDIARMID. "On the method of bounded differences." In: *In Surveys in Combinatorics (J. Siemons, ed.) Cambridge Univ. Press.* (1989), pp. 148–188.

[3] MCDIARMID and C. "Concentration. 195–248." In: *In Probabilistic Methods for Algorithmic Discrete Mathematics (M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed, eds.)* (Springer 1998), pp. 195–248.