0-3 Martingale Methods

1. The independent bounded differences inequality

[Theorem 3.1] Let $X = (X_1, \cdots, X_n)$ be a family of independent random variables with X_k taking values in A_k for each k. Suppose that the real-valued function f defined on $\prod A_k$ s.t.

$$|f(\vec{x}) - f(\vec{x}')| \leq c_k$$

whenever $\vec{x}$ and $\vec{x}'$ differ only in the k-th coordinate. Let $\mu = E[f(X)]$, then for any $t \geq 0$,

$$Pr(f(X) - \mu \geq t) \leq e^{-2t^2/\sum c_k^2}.$$

If we apply it to $-f$ we obtain

$$Pr(f(X) - \mu \leq -t) \leq e^{-2t^2/\sum c_k^2}$$

and we get the 'two-sided' inequlity

$$Pr(|f(X) - \mu| \geq t) \leq 2e^{-2t^2/\sum c_k^2}$$

Note: ① If we let $A_k = \{0,1\}$ and $f(\vec{x}) = \sum x_k$, we obtain Th 2.1

② If A_k is a bounded set, we obtain Th 2.5

We don't give the proof of Th. 3.1.

One application is given:

[Random Graphs], we may take A_k as a set of edges in a graph. $G_{n,p}$ has vertices $1, \cdots, n$ and the possible edges appear independently with prob. p.

[Lemma 3.2] Let $(A_1, \cdots, A_m)$ be a partition of the edge set of the complete graph K_n into m blocks; suppose that the graph function f s.t. $|f(G) - f(G')| \leq 1$ whenever the symmetric difference $E(G) \Delta E(G')$ of the edge-sets is contained in a single block A_k. Then the r.v. $X = f(G_{n,p})$ satisfies in a single block A_k. Then the r.v. $Y = f(G_{n,p})$ satisfies

$$Pr(Y - E(Y) \geq t) \leq e^{-2t^2/m} \quad \text{for} \quad t \geq 0.$$

Note: Lemma 3.2 follows directly from Theorem 3.1 with $C_k=1$.

2. Extensions.

Let $X = (X_1, \cdots, X_n)$, if $E(X) = 0$ and $a \leq X \leq b$,

$$var(X) = E(X^2) = E(X(X-a)) \leq E(b(X-a))$$
$$= |ab| \leq \frac{1}{4}(b-a)^2 \quad (3.13)$$

Let $x_i \in A_i$ for $i = 1, \cdots, k-1$, and let B denote the event that $X_i = x_i$ for each $i = 1, \cdots, k-1$. Let r.v. $P(Y) = p(X_k | B)$

For $x \in A_k$, let

$$g(x) = E(f(X) | B, X_k = x) - E(f(X)|B)$$

If X_k independent then rewrite $g(x)$ as

$$E(f(x_1, \cdots, x_{k-1}, x, X_{k+1}, \cdots, X_n)) - E(f(x_1, \cdots, x_{k-1}, X_k, \cdots, X_n))$$

Observe that $E(g(Y)) = 0$.

Let $dev^+(x_1, \cdots, x_{k-1})$ be the positive deviation of $g(Y)$ as

$$dev^+(x_1, \cdots, x_{k-1}) = \sup\{g(x) : x \in A_k\}$$

Similarily, let

$$dev(x_1, \cdots, x_{k-1}) = \sup\{|g(x)| : x \in A_k\}.$$

If denote $E(f(X))$ by μ, then

$$|f(\vec{x}) - \mu| \leq \sum dev(x_1, \cdots, x_{k-1})$$

Let $ran(x_1, \cdots, x_{k-1})$ denote the range of $g(Y)$ as

$$ran(x_1, \cdots, x_{k-1}) = \sup\{|g(x) - g(y)| : x, y \in A_k\}$$

Also, denote the variance of $g(Y)$ by $var(x_1, \cdots, x_{k-1})$

For $\vec{x} \in \prod A_k$, let

$$R^2(\vec{x}) = \sum_{k=1}^{n} (\mathrm{ran}(x_1, \cdots, x_{k-1}))^2$$
$$\hat{r}^2 = \sup_{\vec{x}} R^2(\vec{x})$$

Similarily, let

$$V(\vec{x}) = \sum_{k=1}^{n} \mathrm{var}(x_1, \cdots, x_{k-1})$$
$$\vec{v} = \sup_{\vec{x}} V(\vec{x})$$

Observe that $V(\vec{x}) \leq R^2(\vec{x})/4$ for each $\vec{x}$ by (3.15), and so $\hat{v} \leq \hat{r}^2/4$.

Finally, let max dev^+ be the maximum of all the positive deviation values $\mathrm{dev}(x_1, \cdots, x_{k-1})$ over all k and x_i.

Similar to denote max dev.

A extension of Th. 3.1

[Th 3.7] Let $X = (X_1, \cdots, X_n)$ with X_k taking values in A_k, and let f be a bounded real-valued function defined on $\prod A_k$. Let μ denote the mean of $f(X)$, and let $\hat{r}^2$ denote the max sum of squared ranges. Then for any $t \geq 0$,

$$\Pr(f(X) - \mu \geq t) \leq e^{-2t^2/\hat{r}^2}$$

More generally, let B be any 'bad' subset of $\prod A_k$ s.t. $R^2(\vec{x}) \leq r^2$ for $\vec{x} \notin B$. Then

$$\Pr(f(X) - \mu \geq t) \leq e^{-2t^2/r^2} + \Pr(X \in B)$$

The next theorem extends the Bernstein theorem.

[Th 3.8] Let $X = (X_1, \cdots, X_n)$ be a family of r.v. with X_k taking values in a set A_k. f defined on $\prod A_k$. Let $\mu = E(f(X))$, Let $b = \max \mathrm{dev}^+$, $\hat{v}$ = max sum of variance, both of which we assume to be finite. Then for any $t \geq 0$

$$\Pr(f(X) - \mu \geq t) \leq e^{-\frac{t^2}{2\hat{v}(1+(bt/(3\hat{v}))}}$$

3. Martingales.

[Martingales] $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ of σ-fields in $\mathcal{F}$. $X_0, X_1, \cdots$ called Martingale if $E(X_{k+1} | \mathcal{F}_k) = X_k$.

[martingale difference sequence] $E(X_{k+1} | \mathcal{F}_k) = 0$.

Note: From a martingale $X_0, X_1, \cdots$, we obtain a martingale difference sequence by setting $Y_k = X_k - X_{k-1}$

The Hoeffding - Azuma Inequality

[Theorem 3.10] Let $c_1, \cdots, c_n$ be constants and let $Y_1, \cdots, Y_n$ be a martingale difference sequence with $|Y_k| \leq c_k$. Then for any $t \geq 0$, $\Pr(|\sum Y_k| \geq t) \leq 2e^{-t^2/2\sum c_k^2}$.

An extension is Th3.13.

[Th 3.13] Let $Y_1, \cdots, Y_n$ be MDS with $a_k \leq Y_k \leq b_k$. Then for any $t \geq 0$ $\Pr(|\sum Y_k| \geq t) \leq 2e^{-2t^2/\sum(b_k - a_k)^2}$