One Hundred Probability/Statistics Inequalities

NP Slagle

February 15, 2012

1 Basic Probability and Measure Theory Inequalities

Given events (sets) A, B, and countable $\{A_n\}_{n=1}^{\infty}$,

- $\mathbf{Prob}[A] \geq 0$
- $\mathbf{Prob}[A] \leq 1$
- If $A \subset B$, then $\mathbf{Prob}[A] \leq \mathbf{Prob}[B]$
- If $A \subset B$, then $\mathbf{Prob}[B] \leq \mathbf{Prob}[A^c]$
- **(Boole)** $\mathbf{Prob}[\cup_{n=1}^{\infty} A_n] \leq \sum_{n=1}^{\infty} \mathbf{Prob}[A_n]$ 5 p.11
- $\mathbf{Prob}[\cup_{n=1}^{\infty} A_n] \geq \sup\{\mathbf{Prob}[A_n] | n = 1, 2, \dots\}$
- $\mathbf{Prob}[\cap_{n=1}^{\infty} A_n] \leq \inf\{\mathbf{Prob}[A_n] | n = 1, 2, \dots\}$
- $\mathbf{Prob}[A \cap B] \leq \min\{\mathbf{Prob}[A], \mathbf{Prob}[B]\}$
- **(Bonferroni)** $\mathbf{Prob}[A \cap B] \geq \mathbf{Prob}[A] + \mathbf{Prob}[B] - 1$ 5 p.11
- **(Bonferroni General)** $\mathbf{Prob}[\cap_{i=1}^{n}] \geq \sum_{i=1}^{n} \mathbf{Prob}[A_i] - (n-1)$ 5 p.13
- $\mathbf{Prob}[A|B] \geq \mathbf{Prob}[A \cap B]$
- **(Karlin Ost)** Define $P_1 = \sum_{i=1}^{n} \mathbf{Prob}[A_i]$, $P_2 = \sum_{1 \leq i < j \leq n} \mathbf{Prob}[A_i \cap A_j]$, $P_3 = \sum_{1 \leq i < j < k \leq n} \mathbf{Prob}[A_i \cap A_j \cap A_k]$,...,$P_n = \mathbf{Prob}[A_1 \cap \dots \cap A_n]$. Then for $i = 1, \dots, n$,

$$P_1 - P_2 + P_3 - \dots \pm P_{i-1} \geq \mathbf{Prob}\left[\cup_{i=1}^{n} A_i\right] \geq P_1 - P_2 + P_3 - \dots \mp P_i. \text{ 5 p.45}$$

2 Means

Define the *pth power mean* of a finite set of positive numbers S to be

$$PM(S,p) = \sqrt[p]{\sum_{s \in S} \frac{s^p}{|S|}} \tag{1}$$

Notice that the arithmetic mean and harmonic mean of the set S are simply $A_M(S) = PM(S,1)$ and $H_M(S) = PM(S,-1)$, respectively. Less clear is that the geometric mean $G_M(S) = \sqrt[|S|]{\prod_{s \in S} s} = \lim_{p \to 0} PM(S,p)$, the maximum $\max(S) = \lim_{p \to \infty} PM(S,p)$, and the minimum $\min(S) = \lim_{p \to -\infty} PM(S,p)$. So, we have

- $PM(S,p_0) \leq PM(S,p_1)$ for $p_0 \leq p_1$ 5 p.204
- $H_M(S) \leq G_M(S) \leq A_M(S)$ 5 p.204

3 Expectations and Variances

Let X and Y be random variables. If an inequality includes a function f of a random variable X, assume that the expectation $\mathbb{E}f(X)$ exists.

- If $g(X) \leq h(X)$, then $\mathbb{E}g(X) \leq \mathbb{E}h(X)$. 5 p.57
- If $a \leq g(X) \leq b$, then $a \leq \mathbb{E}g(X) \leq b$. 5 p.57
- **(Hölder)** If p, q satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then $|\mathbb{E}XY| \leq \mathbb{E}|XY| \leq (\mathbb{E}|X|^p)^{\frac{1}{p}}(\mathbb{E}|Y|^q)^{\frac{1}{q}}$ 5 p.187
- **(Jensen)** For a convex function g, If $X \geq Y$, then $\mathbb{E}g(X) \geq g(\mathbb{E}X)$. 5 p.190
- **(Cauchy-Schwartz)** $|\mathbb{E}XY| \leq \mathbb{E}|XY| \leq \sqrt{(\mathbb{E}|X|^2)(\mathbb{E}|Y|^2)}$ 5 p.187
- $\text{Var}(X) \geq 0$
- $\text{Cov}^2(X,Y) \leq \text{Var}(X)\text{Var}(Y)$ 5 p.188
- **(Hölder Special Case)** For $p > 1$, $\mathbb{E}|X| \leq \sqrt[p]{\mathbb{E}|X|^p}$ 5 p.188
- **(Liapounov)** For $s > r > 1$, $\sqrt[r]{\mathbb{E}|X|^r} \leq \sqrt[s]{\mathbb{E}|X|^s}$ 5 p.188
- **(Minkowski)** For $p \geq 1$, $\sqrt[p]{\mathbb{E}|X+Y|^p} \leq \sqrt[p]{\mathbb{E}|X|^p} + \sqrt[p]{\mathbb{E}|Y|^p}$ 5 p. 188
- **(Triangle)** As a special case of Minkowski's inequality, $\mathbb{E}|X+Y| \leq \mathbb{E}|X| + \mathbb{E}|Y|$. 5 p.203
- If g is nondecreasing and h is nonincreasing, then $\mathbb{E}(g(X)h(X)) \leq (\mathbb{E}g(X))(\mathbb{E}h(X))$. 5 p.192
- If g and h are both nondecreasing or both nonincreasing, then $\mathbb{E}(g(X)h(X)) \geq (\mathbb{E}g(X))(\mathbb{E}h(X))$. 5 p.192
- **(Cramér-Rao)** Suppose $X_1, \dots, X_n$ is a sample with joint pdf $f(\mathbf{x}|\theta)$ and $W(\mathbf{X})$ is any estimator of θ such that $\frac{d}{d\theta}\mathbb{E}_\theta W(\mathbf{X}) = \int_\chi \frac{\partial}{\partial\theta}[W(\mathbf{x})]f(\mathbf{x}|\theta)d\mathbf{x}$ and $\text{Var}_\theta(W(\mathbf{X})) < \infty$. Then

$$\text{Var}_\theta(W(\mathbf{X})) \geq \frac{(\frac{d}{d\theta}\mathbb{E}_\theta W(\mathbf{X}))^2}{\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log(f(\mathbf{x}|\theta))\right)^2\right]}.$$ 5 p.335

- **(Cramér-Rao IID)** Suppose $X_1, \dots, X_n$ is a sample iid with marginal pdf $f(x|\theta)$ and $W(\mathbf{X})$ is any estimator of θ such that $\frac{d}{d\theta}\mathbb{E}_\theta W(\mathbf{X}) = \int_\chi \frac{\partial}{\partial\theta}[W(\mathbf{x})]f(\mathbf{x}|\theta)d\mathbf{x}$ and $\text{Var}_\theta(W(\mathbf{X})) < \infty$. Then

$$\text{Var}_\theta(W(\mathbf{X})) \geq \frac{(\frac{d}{d\theta}\mathbb{E}_\theta W(\mathbf{X}))^2}{nE_\theta\left[\left(\frac{\partial}{\partial\theta}\log(f(x|\theta))\right)^2\right]}.$$ 5 p.337

- **(Rao-Blackwell)** Let U be an unbiased estimator of $\tau(\theta)$, and let T be a sufficient statistic for θ. Define $\phi(T) = \mathbb{E}(U|T)$. Then $\mathbb{E}\phi(T) = \tau(\theta)$, and

$$\text{Var}_\theta\phi(T) \leq \text{Var}_\theta W \text{ for all } \theta.$$ 5 p.342

- **(Han)** Let $X_1, \dots, X_n$ be independent discrete random variables. Let $H(X_{\pi_1}, \dots, X_{\pi_k})$ be the joint entropy of a subset of the $\{X_i\}$. Then

$$H(X_1, \dots, X_n) \leq \sum_{i=1}^n H(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n).$$ 4 p.230

- Let $X_1, \dots, X_n$ be independent random variables. Let $g : \mathit{Domain}(X_1, \dots, X_n) \to \mathbb{R}$ be Lesbegue measurable, and $Z = g(X_1, \dots, X_n)$. Then

$$\text{Var}(Z) \leq \sum_{i=1}^n \mathbb{E}[(Z - \mathbb{E}(Z|X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)^2].$$ 4 p.219

- **(Efron-Stein)** Let $X_1, \ldots, X_n$ be independent random variables. Let $g : Domain(X_1, \ldots, X_n) \to \mathbb{R}$ be Lesbegue measurable, and $Z = g(X_1, \ldots, X_n)$. Let $Y_1, \ldots, Y_n$ be an independent copy of $X_1, \ldots, X_n$, and let $Z_i = g(X_1, \ldots, Y_i, \ldots, X_n)$. Then

$$\mathrm{Var}(Z) \leq \sum_{i=1}^{n} \mathbb{E}[(Z - Z_i)^2]. \text{ 4 p.220}$$

- **(Logarithmic Sobolev)** Let $X_1, \ldots, X_n$ be independent random variables. Let $g_i : Domain(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) \to \mathbb{R}$ be Lesbegue measurable, $Z_i = g_i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$, $g : Domain(X_1, \ldots, X_n) \to \mathbb{R}$ be Lesbegue measurable, and $Z = g(X_1, \ldots, X_n)$. Let $\psi(t) = e^t - t - 1$ and $s > 0$. Then

$$sE(Ze^{sZ}) - \mathbb{E}(e^{sZ}) \log[\mathbb{E}(e^{sZ})] \leq \sum_{i=1}^{n} \mathbb{E}[e^{sZ} \psi(-s(Z - Z_i))]. \text{ 4 p.233}$$

- **(Symmetrized Logarithmic Sobolev)** Let $X_1, \ldots, X_n$ be independent random variables. Let $g : Domain(X_1, \ldots, X_n) \to \mathbb{R}$ be Lesbegue measurable, and $Z = g(X_1, \ldots, X_n)$. Let $Y_1, \ldots, Y_n$ be an independent copy of $X_1, \ldots, X_n$, and let $Z_i = g(X_1, \ldots, Y_i, \ldots, X_n)$. Let $\psi(t) = e^t - t - 1$ and $s > 0$. Then

$$sE(Ze^{sZ}) - \mathbb{E}(e^{sZ}) \log[\mathbb{E}(e^{sZ})]] \leq \sum_{i=1}^{n} \mathbb{E}[e^{sZ} \psi(-s(Z - Z_i))]. \text{ 4 p.234}$$

- Suppose $\{X_n\}$ is a sequence of random variables such that for all n, $X_n \geq 0$, and for all $\epsilon > 0$, there exist $c_1 > 0$ and $c_2 > e^{-1}$ such that $\mathbf{Prob}[X_n > \epsilon] \leq c_1 e^{-c_2 n \epsilon^2}$. Then

$$\mathbb{E}X_n \leq \sqrt{\frac{1 + \log(c_1)}{nc_2}}. \text{ 25}$$

- **(Kannan Strong Negative Correlation)** Suppose m is an even positive integer, and $X_1, \ldots, X_n$ are real-valued random observations satisfying the *strong negative correlation principle*. That is, for all i, $\mathbb{E}X_i(X_1 + \cdots + X_{i-1})^l < 0$ when $l < m$ is odd and $\mathbb{E}(X_i^l | X_1 + \cdots + X_{i-1}) \leq \left(\frac{n}{m}\right)^{\frac{l-2}{2}} l!$ for $l \leq m$ even. Then

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)^m \leq (24mn)^{\frac{m}{2}}. \text{ 13 p.2}$$

- **(Kannan Hamiltonian Tour)** Suppose $Y_1, \ldots, Y_n$ are sets of points generated independently and respectively from n subsquares of size $\frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{n}}$ of the unit square, and there exists a constant $c \in (0, 1)$ such that $\mathbf{Prob}[|Y_i|] \leq c$ for all i. Suppose further that for $\epsilon > 0$, and $l \in \{1, \ldots, \frac{m}{2}\}$, $\mathbb{E}|Y_i|^l \leq [O(l)]^{(2-\epsilon)l}$. Finally, suppose $f(Y_1, \ldots, Y_n)$ is the length of the shortest Hamiltonian tour through $Y_1 \cup \cdots \cup Y_n$. Then

$$\mathbb{E}[f(Y_1, \ldots, Y_n) - \mathbb{E}f(Y_1 \ldots, Y_n)]^m \leq (cm)^{\frac{m}{2}}. \text{ 13 p.4}$$

- **(Kannan MST)** Suppose $Y_1, \ldots, Y_n$ are sets of points generated independently and respectively from n subsquares of size $\frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{n}}$ of the unit square, and there exists a constant $c \in (0, 1)$ such that $\mathbf{Prob}[|Y_i|] \leq c$ for all i. Suppose further that for $\epsilon > 0$, and $l \in \{1, \ldots, \frac{m}{2}\}$, $\mathbb{E}|Y_i|^l \leq [O(l)]^{(2-\epsilon)l}$. Finally, suppose $f(Y_1, \ldots, Y_n)$ is the length of a minimum spanning tree of $Y_1 \cup \cdots \cup Y_n$. Then

$$\mathbb{E}[f(Y_1, \ldots, Y_n) - \mathbb{E}f(Y_1 \ldots, Y_n)]^m \leq (cm)^{\frac{m}{2}}. \text{ 13 p.5}$$

- **(Kannan Random Vector)** Suppose $\mathbf{Y} = (Y_1, \ldots, Y_n)$ is a random vector such that for a fixed $k \leq n$, $\mathbb{E}(Y_i^2 | Y_1^2, \ldots, Y_i^2)$ is a nondecreasing function of $Y_1^2 + \cdots + Y_{i-1}^2$ for $i = 1, \ldots, k$ and for even $l \leq k$, there exists a $c > 0$ such that $\mathbb{E}(Y_i^l | Y_1^2, \ldots, Y_{i-1}^2) \leq \left(\frac{cl}{n}\right)^{\frac{l}{2}}$. Then for any even $m \leq k$,
$$\mathbb{E}\left(\sum_{i=1}^{k} Y_i^2 - \mathbb{E}Y_i^2\right)^m \leq \left(\frac{\sqrt{cmk}}{n}\right)^m.$$ 13 p.5

- **(Ledoux-Talagrand Contraction)** Suppose $X_i, \ldots, X_n$ are iid *Rademacher* variables ($\mathbf{Prob}[X_i = 1] = \mathbf{Prob}[X_i = -1] = \frac{1}{2}$). Suppose $f : \mathbb{R}^+ \to \mathbb{R}^+$ be convex and increasing, and $\phi_i : \mathbb{R} \to \mathbb{R}$ be Lipschitz with constant L for $i = 1, \ldots, n$. Then for $T \subset \mathbb{R}^n$,
$$\mathbb{E}f\left(\frac{1}{2}\sup_{\mathbf{t}\in T}\left|\sum_{i=1}^{n} X_i \phi_i(t_i)\right|\right) \leq \mathbb{E}f\left(L \sup_{\mathbf{t}\in T}\left|\sum_{i=1}^{n} X_i t_i\right|\right) \quad 9p.9$$

- **(Bhatia-Davis)** If a univariate probability distribution F has minimum m, maximum M, and mean μ, then for any X following F, $\mathrm{Var}(X) \leq (M-\mu)(\mu-m)$. 2 p.353–357

- **(Popoviciu)** If a univariate probability distribution F has minimum m and maximum M, then for any X following F, $\mathrm{Var}(X) \leq \frac{1}{4}(M-m)^2$. 23 p.313–318

- **(Chapman-Robbins)** Suppose $\mathbf{X}$ is a random variable in $\mathbb{R}^k$ with an unknown parameter θ. If $\delta(\mathbf{X})$ is an unbiased estimator for $\tau(\theta)$, then
$$\mathrm{Var}(\delta(\mathbf{X})) \geq \sup_{\Delta} \frac{[\tau(\theta+\Delta) - \tau(\theta)]^2}{\mathbb{E}_\theta\left[\frac{p(\mathbf{X}.\theta+\Delta)}{p(\mathbf{X},\theta)} - 1\right]^2}.$$ 6 p.581–586

- **(Entropy Power)** Define the *entropy* of X to be $h(X) = -\mathbb{E}\log f_X(X)$, where $f_X(x)$ is the pdf or pmf of X. Define the *entropypower* of X to be $N(X) = \frac{1}{2\pi e}e^{\frac{2}{n}h(X)}$. Then for random variables X and Y, we have $N(X+Y) \geq N(X) + N(Y)$. 8 p.1501–1518

- **(Marcinkiewicz Zygmund)** Let $X_1, \ldots, X_n$ be independent random variables with common support such that $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^p < \infty$ for all $p \geq 1$. Then there exist constants $A(p)$ and $B(p)$, dependent only on p, such that
$$A(p)\mathbb{E}\left(\sum_{i=1}^{n}|X_i|^2\right)^{\frac{p}{2}} \leq \mathbb{E}\left(\sum_{i=1}^{n}|X_i|\right)^p \leq B(p)\mathbb{E}\left(\sum_{i=1}^{n}|X_i|^2\right)^{\frac{p}{2}}.$$ 15 p.233–259

- **(Khintchine)** Let $X_1, \ldots, X_n$ be iid Rademacher random variables. Then for any $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and $p > 0$, there exist constants $A(p)$ and $B(p)$, dependent only on p, such that
$$A(p)\left(\sum_{i=1}^{n}|\lambda_i|^2\right)^{\frac{1}{2}} \leq \left(\mathbb{E}\left(\sum_{i=1}^{n}|\lambda_i X_i|\right)^p\right)^{\frac{1}{p}} \leq B(p)\left(\sum_{i=1}^{n}|\lambda_i|^2\right)^{\frac{1}{2}}.$$ 26

- **(Rosenthal I)** Let $X_1, \ldots, X_n$ be independent nonnegative random variables such that $\mathbb{E}X_i^p < \infty$ for a fixed $p \geq 1$, $i = 1, \ldots, n$. Then there exist constants $A(p)$ and $B(p)$ dependent only on p such that
$$A(p)\max\left\{\sum_{i=1}^{n}\mathbb{E}X_i^p, \left(\sum_{i=1}^{n}\mathbb{E}X_i\right)^p\right\} \leq \left(\sum_{i=1}^{n}\mathbb{E}X_i\right)^p \leq B(p)\max\left\{\sum_{i=1}^{n}\mathbb{E}X_i^p, \left(\sum_{i=1}^{n}\mathbb{E}X_i\right)^p\right\}.$$ 19 p.273–303

- **(Rosenthal II)** Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^p < \infty$ for a fixed $p \geq 1$, $i = 1, \ldots, n$. Then there exist constants $A(p)$ and $B(p)$ dependent only on p such that
$$A(p)\max\left\{\sum_{i=1}^{n}\mathbb{E}|X_i|^p, \left(\sum_{i=1}^{n}\mathbb{E}X_i^2\right)^{\frac{p}{2}}\right\} \leq \left|\sum_{i=1}^{n}\mathbb{E}X_i\right|^p \leq B(p)\max\left\{\sum_{i=1}^{n}\mathbb{E}|X_i|^p, \left(\sum_{i=1}^{n}\mathbb{E}X_i^2\right)^{\frac{p}{2}}\right\}.$$ 19 p.273–303

- **(Papadatos)** Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of iid random variables $X_1, \dots, X_n$ with variance σ^2. Define $G(x) = I_x(k, n+1-k)$ and $\sigma_n^2(k) = \sup_{0<x<1}\left[\frac{G(x)(1-G(x))}{x(1-x)}\right]$. Then

$$\mathrm{Var}(X_{(k)}) \le \sigma_n^2(k)\sigma^2. \text{ 17 p.5}$$

- **(Hürlimann Upper *n*–*r*)** Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of iid random variables $X_1, \dots, X_n$. Define $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and the *biased observed variance* $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Then for $r = 0, \dots, n-1$, the *average of the upper n-r order statistics* satisfies

$$\frac{1}{n-r}\sum_{i=r+1}^n X_{(i)} \le \bar{X} + S\sqrt{\frac{r}{n-r}}. \text{ 12 p.4}$$

- **(Hürlimann Average Excess)** Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of iid random variables $X_1, \dots, X_n$. Define $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and the *biased observed variance* $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Then for $r = 0, \dots, n-1$, the *average excess of the upper $n-r$ order statistics conditioned on the rth order statistic* satisfies

$$\frac{1}{n-r}\sum_{i=r+1}^n (X_{(i)} - X_{(r)}) \le S\frac{n}{\sqrt{r(n-r)}}. \text{ 12 p.7}$$

- **(Hürlimann Stop-Loss Excess)** Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of iid random variables $X_1, \dots, X_n$. Define $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and the *biased observed variance* $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Define the *rth stop-loss statistics* to be $SL(d_r) = \sum_{i=1}^n (X_{(i)} - (n-r)d_r)$ for $d_r \in [X_{(r)}, X_{(r+1)}]$. Then for $r = 0, \dots, n-1$,

$$SL(d_r) \le (n-r)\left[\bar{X} - d_r + S\sqrt{\frac{r}{n-r}}\right]. \text{ 12 p.4}$$

4 Concentration Inequalities

Let X be a random variable.

- **(Chebychev General)** For $r > 0$, g a nonnegative function, $\mathbf{Prob}[g(X) \ge r] \le \frac{\mathbb{E}g(X)}{r}$. 5 p.122
- **(Chebychev)** For $t > 0$, $\mathbf{Prob}[|X - \mathbb{E}X| \ge t] \le \frac{\mathrm{Var}(X)}{t^2}$. 5 p.122
- **(Normal I, Mill)** For Z a standard normal, $\mathbf{Prob}[|Z| \ge t] \le \sqrt{\frac{2}{\pi}}\frac{e^{-t^2/2}}{t}$. 5 p.122 24 p.65
- **(Normal II)** For Z a standard normal, $\mathbf{Prob}[|Z| \ge t] \ge \sqrt{\frac{2}{\pi}}e^{-t^2/2}\frac{t}{1+t^2}$. 5 p.135
- **(Chernoff I)** Let $M_X(t)$,$-h \le t \le h$ be the moment-generating function of X. Then $\mathbf{Prob}[X > a] \le e^{-at}M_X(t)$, $-h \le t \le h$. 5 p.134 16 p.65
- **(Chernoff II)** Let $M_X(t)$,$-h \le t \le h$ be the moment-generating function of X. Then $\mathbf{Prob}[X \le a] \le e^{-at}M_X(t)$, $-h \le t \le 0$. 5 p.134 16 p.65
- **(Chernoff Sum I)** Let $X_1, \dots, X_n$ be iid, $X = \sum_{i=1}^n X_i$, and $M_X(t)$,$-h \le t \le h$ be the moment-generating function of X_1. . Then $\mathbf{Prob}[S > a] \le e^{-at}[M_X(t)]^n$ for $0 \le t \le h$. 5 p.262
- **(Chernoff Sum II)** Let $X_1, \dots, X_n$ be iid, $X = \sum_{i=1}^n X_i$, and $M_X(t)$,$-h \le t \le h$ be the moment-generating function of X_1. Then $\mathbf{Prob}[S \le a] \le e^{-at}[M_X(t)]^n$ for $-h \le t \le 0$. 5 p.262
- **(Chernoff Mean)** Let $X_1, \dots, X_n$ be iid, $\epsilon > 0$, $\bar{X_n} = \sum_{i=1}^n X_i$, $M_U(t)$, $-h_U \le t \le h_U$ be the moment-generating function of $U = X_1 - \mathbb{E}X_1 - \epsilon$, and $M_V(t)$, $-h_V \le t \le h_V$ be the moment-generating function of $V = -X_1 + \mathbb{E}X_1 - \epsilon$. Then there exist for some $0 < t_U \le h_U$ and $-h_V \le t_V < 0$ [1] such that

$$\mathbf{Prob}[|\bar{X_n} - \mathbb{E}X_1| > \epsilon] \le 2c^n, \text{ where } c = \max\{M_U(t_U), M_V(t_V)\} \in (0,1). \text{ 5 p.262}$$

[1] Such a t_U and t_V exist since $\mathbb{E}U < 0$ and $\mathbb{E}V < 0$, guaranteeing that M_U and M_V are decreasing in a neighborhood of zero.

- **(Chernoff Poisson Trials I)** Let X_i be n independent Poisson trials [2]. Let $X = \sum_{i=1}^n X_i$. Then for $\delta > 0$,

$$\mathbf{Prob}[X \geq (1+\delta)\mathbb{E}X] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mathbb{E}X}. \text{ 16 p.64}$$

- **(Chernoff Poisson Trials II)** Let X_i be n independent Poisson trials. Let $X = \sum_{i=1}^n X_i$. Then for $0 < \delta \leq 1$,

$$\mathbf{Prob}[X \geq (1+\delta)\mathbb{E}X] < e^{-(\mathbb{E}X)\delta^2/3}. \text{ 16 p.64}$$

- **(Chernoff Poisson Trials III)** Let X_i be n independent Poisson trials. Let $X = \sum_{i=1}^n X_i$. Then for $R \geq gEX$,

$$\mathbf{Prob}[X \geq R] < 2^{-R}. \text{ 16 p.64}$$

- **(Chernoff Poisson Trials IV)** Let X_i be n independent Poisson trials. Let $X = \sum_{i=1}^n X_i$. Then for $0 < \delta < 1$,

$$\mathbf{Prob}[X \leq (1-\delta)\mathbb{E}X] < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mathbb{E}X}. \text{ 16 p.65}$$

- **(Chernoff Poisson Trials V)** Let X_i be n independent Poisson trials. Let $X = \sum_{i=1}^n X_i$. Then for $0 < \delta < 1$,

$$\mathbf{Prob}[X \leq (1-\delta)\mathbb{E}X] < e^{-\delta^2\mathbb{E}X/2}. \text{ 16 p.65}$$

- **(Chernoff Rademacher I)** Suppose $X_1, \ldots, X_n$ be iid such that $\mathbf{Prob}[X_i = 1] = \mathbf{Prob}[X_i = -1] = \frac{1}{2}$. If $X = \sum_{i=1}^n X_i$ and $a > 0$, then $\mathbf{Prob}[X \geq a] \leq e^{\frac{-a^2}{2n}}$. 16 p.69
- **(Chernoff Rademacher II)** Suppose $X_1, \ldots, X_n$ be iid such that $\mathbf{Prob}[X_i = 1] = \mathbf{Prob}[X_i = -1] = \frac{1}{2}$. If $X = \sum_{i=1}^n X_i$ and $a > 0$, then $\mathbf{Prob}[|X| \geq a] \leq 2e^{\frac{-a^2}{2n}}$. 16 p.70
- **(Chernoff Bernoulli I)** Suppose $X_1, \ldots, X_n$ be iid Bernoulli$\left(\frac{1}{2}\right)$. If $X = \sum_{i=1}^n X_i$ and $0 < a < \frac{n}{2}$, then $\mathbf{Prob}\left[X \leq \frac{n}{2} - a\right] \leq 2e^{\frac{-2a^2}{n}}$. 16 p.71
- **(Chernoff Bernoulli II)** Suppose $X_1, \ldots, X_n$ be iid Bernoulli$\left(\frac{1}{2}\right)$. If $X = \sum_{i=1}^n X_i$ and $0 < \delta < 1$, then $\mathbf{Prob}\left[X \leq \frac{n}{2}(1-\delta)\right] \leq 2e^{\frac{-n\delta^2}{2}}$. 16 p.71
- **(Chernoff Bernoulli III)** Suppose $X_1, \ldots, X_n$ be iid Bernoulli$\left(\frac{1}{2}\right)$. If $X = \sum_{i=1}^n X_i$ and $a > 0$, then $\mathbf{Prob}\left[X \geq \frac{n}{2} + a\right] \leq 2e^{\frac{-2a^2}{n}}$. 16 p.70
- **(Chernoff Bernoulli IV)** Suppose $X_1, \ldots, X_n$ be iid Bernoulli$\left(\frac{1}{2}\right)$. If $X = \sum_{i=1}^n X_i$ and $\delta > 0$, then $\mathbf{Prob}\left[X \leq \frac{n}{2}(1+\delta)\right] \leq 2e^{\frac{-n\delta^2}{2}}$. 16 p.70
- **(Markov)** If $X \geq 0$ and $\mathbf{Prob}[X = 0] < 1$, then for $r > 0$, $\mathbf{Prob}[X \geq r] \leq \frac{\mathbb{E}X}{r}$. 5 p.136
- **(Gauss)** Suppose X follows a unimodal distribution with mode ν, and define $\tau^2 = \mathbb{E}(X-\nu)^2$. Then

$$\mathbf{Prob}[|X-\nu| > \epsilon] \leq \begin{cases} \frac{4\tau^2}{9\epsilon^2}, & \epsilon \geq \sqrt{\frac{4}{3}}\tau \\ 1 - \frac{\epsilon}{\tau\sqrt{3}}, & \epsilon \leq \sqrt{\frac{4}{3}}\tau \end{cases} \text{ 5 p.137}$$

- **(Vysochanskiĭ-Petunin)** Suppose X follows a unimodal distribution, and define $\xi^2 = \mathbb{E}(X-\alpha)^2$ for arbitrary α. Then

$$\mathbf{Prob}\left[|X-\alpha| > \epsilon\right] \leq \begin{cases} \frac{4\xi^2}{9\epsilon^2}, & \epsilon \geq \sqrt{\frac{8}{3}}\xi \\ \frac{4\xi^2}{9\epsilon^2} - \frac{1}{3}, & \epsilon \leq \sqrt{\frac{8}{3}}\xi \end{cases} \text{ 5 p.137}$$

[2] Each X_i is a Bernoulli(p_i).

- **(Hoeffding I)** Let $Y_1, \ldots, Y_n$ be independent observations such that $\mathbb{E}Y_i = 0$ and $a_i \leq Y_i \leq b_i$ for all i. If $\epsilon > 0$ and $t > 0$, then
$$\textbf{Prob}\left[\sum_{i=1}^{n} Y_i \geq \epsilon\right] \leq e^{-t\epsilon} \prod_{i=1}^{n} e^{t^2(b_i-a_i)^2/8} \text{ 24 p.64}$$

- **(Hoeffding II)** Let $X_1, \ldots, X_n$ be independent Bernoulli(p). If $\epsilon > 0$, then
$$\textbf{Prob}\left[\left|\sum_{i=1}^{n} X_i - np\right| \geq \epsilon\right] \leq 2e^{-2n\epsilon^2} \text{ 24 p.65}$$

- **(Saw)** Suppose $X_1, \ldots, X_n$ are iid with finite first and second order moments. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Let $k > 0$, $\nu(t) = \max\left\{m \in \mathbb{N} | m < \frac{n+1}{t}\right\}$, $\alpha(t) = \frac{(n+1)(n+1-\nu(t))}{1+\nu(t)(n+1-\nu(t))}$, and $\beta = \frac{n(n+1)k^2}{n-1+(n+1)k^2}$. Then
$$\textbf{Prob}[|X - \bar{X}| \geq kS] \leq \begin{cases} \frac{1}{n+1}(\nu(\beta) - 1) & \text{if } \nu \text{ is odd and } \beta > \alpha(\beta) \\ \frac{1}{n+1}\nu(\beta) & \text{otherwise. 5 p.268} \end{cases}$$

- **(Talagrand)** Let $\mathbf{X}$ be chosen randomly uniformly from $\{-1, 1\}^n$, let A be a convex subset of $\mathbb{R}^n$, $A_t = \{\mathbf{p} \in \mathbb{R}^n | dist(\mathbf{p}, A) \leq t\}$. Then there exists $c > 0$ such that $\textbf{Prob}[\mathbf{X} \in A]\textbf{Prob}[\mathbf{X} \notin A_t] \leq e^{-ct^2}$ for all $t > 0$. 22

- **(Talagrand Large Deviation)** Let $\mathbf{X}$ be chosen randomly uniformly from $\{-1, 1\}^n$, V be a d-dimensional subspace of $\mathbb{R}^n$. Then there exist constants $c, C > 0$ such that $\textbf{Prob}[|dist(X, V) - \sqrt{n-d}| \geq t] \leq Ce^{-ct^2}$ for all $t > 0$. 22

- **(Gaussian for Lipschitz)** Let $\mathbf{X}$ be an n-dimensional random vector such that each X_i is an independent $n(0, 1)$ variable. If $f : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz function with scale constant 1[3], then there exists a constant $c > 0$ such that $\textbf{Prob}[|f(\mathbf{X}) - \mathbb{E}f(\mathbf{X})| \geq t] \leq e^{-ct^2}$ for all $t > 0$. 22

- **(Azuma)** Suppose $X_0, \ldots, X_n$ is a *martingale* ($\mathbb{E}(X_i | X_1, \ldots, X_{i-1}) = X_{i-1}$ for $i = 1, \ldots, n$); suppose further that X is $\mathbf{c}$-Lipschitz ($|X_i - X_{i-1}| \leq c_i$ for $i = 1, \ldots, n$, $c \in \mathbb{R}^n$ positive); then
$$\textbf{Prob}[X_n - X_0 \geq \lambda] \leq 2e^{\frac{-\lambda^2}{2\sum_{i=1}^{n} c_i^2}}. \text{ 7 p. 37-38}$$

- **(Bennett)** Let $X_1, \ldots, X_n$ be independent random variables of zero mean such that $\textbf{Prob}[X_i \leq 1] = 1$. Let $h(u) = (1 + u)\log(1 + u) - u$ for $u \geq 0$ and $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} \text{Var}(X_i)$. Then for $t > 0$,
$$\textbf{Prob}\left[\sum_{i=1}^{n} X_i > t\right] \leq e^{-n\sigma^2 h\left(\frac{t}{n\sigma^2}\right)}. \text{ 4 p.218}$$

- **(Bernstein)** Let $X_1, \ldots, X_n$ be independent random variables of zero mean such that $\textbf{Prob}[X_i \leq 1] = 1$. Let $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} \text{Var}(X_i)$. Then for $\epsilon > 0$,
$$\textbf{Prob}\left[\frac{1}{n}\sum_{i=1}^{n} X_i > \epsilon\right] \leq e^{\frac{-n\epsilon^2}{2(\sigma^2+\epsilon/3)}}. \text{ 4 p.219}$$

- **(McDiarmid Bounded Differences I)** Let $X_1, \ldots, X_n$ be independent random variables each whose domain is χ. If $f : \chi^n \to \mathbb{R}^n$ is a function such that for all $\mathbf{x} \in \chi^n$, $y \in \chi$, and $i \in \{1, \ldots, n\}$, there exists a constant $c_i > 0$ such that $|f(\mathbf{x}) - f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)| \leq c_i$, then
$$\textbf{Prob}[f(\mathbf{X}) - \mathbb{E}f(\mathbf{X}) \geq t] \geq e^{\frac{-t^2}{\sum_{i=1}^{n} c_i^2}} \text{ for all } t > 0. \text{ 1}$$

[3] A Lipschitz function f satisfies $|f(x) - f(y)| \leq M||x - y||$ for all $x, y \in domain(f)$.

- **(McDiarmid Bounded Differences II)** Let $X_1, \ldots, X_n$ be independent random variables each whose domain is χ. If $f : \chi^n \to \mathbb{R}^n$ is a function such that for all $\mathbf{x} \in \chi^n$, $y \in \chi$, and $i \in \{1, \ldots, n\}$, there exists a constant $c_i > 0$ such that $|f(\mathbf{x}) - f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)| \leq c_i$, then

$$\mathbf{Prob}[f(\mathbf{X}) - \mathbb{E}f(\mathbf{X}) \leq -t] \geq e^{\frac{-t^2}{\sum_{i=1}^n c_i^2}} \text{ for all } t > 0. \ 1$$

- **(Kannan Chromatic Number)** Let $G(\{1, \ldots, n\}, P)$ be a random graph with edge probabilities $P = (p_{ij})$. The *chromatic number* $\chi = \chi(G)$ is the least number of colors necessary to color G such that no two vertices sharing an edge receive the same color. Let $p = \frac{\sum_{i,j} p_{ij}}{\binom{n}{2}}$. Then there exists a constant $c > 0$ such that for $t \in (0, n\sqrt{p})$,

$$\mathbf{Prob}[|\chi(G) - \mathbb{E}\chi(G)| \geq t] \leq e^{\frac{-ct^2}{n\sqrt{p}\log n}}. \ 13 \text{ p.5}$$

- **(Johnson-Lindenstrauss Random Projection)** Suppose $k \leq n$, and we pick $V_1, \ldots, V_k$ uniformly randomly from the surface of the unit ball in $\mathbb{R}^n$. Then for $\epsilon \in (0, 1)$, there exist constants $c_1, c_2 > 0$ such that

$$\mathbf{Prob}\left[\left|\sum_{i=1}^k v_i^2 - \frac{k}{n}\right| \geq \frac{\epsilon k}{n}\right] \leq c_1 e^{-c_2\epsilon^2 k} \ 13 \text{ p.5}$$

- **(Kannan Random Projection)** Suppose m is an even positive integer and $X_1, \ldots, X_n$ are real-valued random observations satisfying the *strong negative correlation principle*. That is, for all i, $\mathbb{E}X_i(X_1 + \cdots + X_{i-1})^l < 0$ when $l < m$ is odd and $\mathbb{E}(X_i^l | X_1 + \cdots + X_{i-1}) \leq \left(\frac{n}{m}\right)^{\frac{l-2}{2}} l!$ for $l \leq m$ even. Define constants $\{M_{i,l}\}$, $\{K_{i,l}\}$, and $\{L_{i,l}\}$ such that $\mathbb{E}(X_i^l | X_1 + \cdots + X_{i-1}) \leq M_{i,l}$, each $K_{i,l}$ is an indicator variable on the *typical* case of the conditional expectation where $\mathbf{Prob}[K_{i,l}] = 1 - \delta_{i,l}$, and $\mathbb{E}(X_i^l | X_1 + \cdots + X_{i-1}, K_{i,l}) \leq L_{i,l}$ for $l = 2, 4, \ldots, m$ and $i = 1, \ldots, n$. Finally, let $X = \sum_{i=1}^n X_i$. Then

$$\mathbb{E}X^m \leq (cm)^{\frac{m}{2}+1}\left(\sum_{l=1}^{m/2} \frac{m^{1-\frac{1}{l}}}{l^2}\left(\sum_{i=1}^n L_{i,2l}\right)^{\frac{1}{l}}\right)^{\frac{m}{2}} + (cm)^{m+2}\sum_{l=1}^{m/2}\frac{1}{nl^2}\sum_{i=1}^n\left(nM_{i,2l}\delta_{i,2l}^{2/(m-2l+2)}\right)^{\frac{2}{ml}}. \ 13 \text{ p.5}$$

- **(Kannan Bin Packing)** Suppose $Y_1, \ldots, Y_n$ are iid from a discrete distribution of r atoms each with probability at least $\frac{1}{\log n}$ and $\mathbb{E}Y_1 \leq \frac{1}{r^2 \log n}$. Let $f(Y_1, \ldots, Y_n)$ be the minimum number of unit capacity bins necessary to pack the $Y_1, \ldots, Y_n$ items. Then there exist constants $c_1, c_2 > 0$ such that if $t \in (0, n[(\mathbb{E}Y_i)^3 + \mathrm{Var}(Y_i)])$, then

$$\mathbf{Prob}[|f - \mathbb{E}f| \geq t + r] \leq c_1 e^{\frac{-c_2 t^2}{n[(\mathbb{E}Y_i)^3 + \mathrm{Var}(Y_i)]}}. \ 13 \text{ p.8}$$

- **(Dvoretzky Kiefer Wolfowitz I)** Suppose $X_1, \ldots, X_n$ are iid univariate random variables following cdf F. Let $F_n(x) = \frac{1}{n}\sum_{i=1}^n 1_{X_i \leq x}$ be the empirical distribution. Then for $\epsilon > \sqrt{\frac{1}{2n}\log 2}$,

$$\mathbf{Prob}\left[\sup_{x\in\mathbb{R}}(F_n(x) - F(x)) > \epsilon\right] \leq e^{-2n\epsilon^2}. \ 10 \text{ p.642–669}$$

- **(Dvoretzky Kiefer Wolfowitz II)** Suppose $X_1, \ldots, X_n$ are iid univariate random variables following cdf F. Let $F_n(x) = \frac{1}{n}\sum_{i=1}^n 1_{X_i \leq x}$ be the empirical distribution. Then for $\epsilon > 0$,

$$\mathbf{Prob}\left[\sup_{x\in\mathbb{R}}|F_n(x) - F(x)| > \epsilon\right] \leq 2e^{-2n\epsilon^2}. \ 10 \text{ p.642–669}$$

- **(Etemadi Differing Means)** Let $X_1, \ldots, X_n$ be random variables with common support. Let $S_k = \sum_{i=1}^k X_k$ be the kth partial sum. Then for $\epsilon > 0$,

$$\mathbf{Prob}\left[\max_{1\leq k\leq n}|S_k| \geq 3\epsilon\right] \leq 3\max_{1\leq k\leq n}\mathbf{Prob}\left[|S_k| \geq \epsilon\right]. \ 11 \text{ p.215–221}$$

- **(Etemadi Shared Means)** Let $X_1, \ldots, X_n$ be random variables with common support and equal means. Let $S_k = \sum_{i=1}^k X_k$ be the kth partial sum. Then for $\epsilon > 0$,

$$\mathbf{Prob}\left[\max_{1 \le k \le n} |S_k| \ge \epsilon\right] \le \frac{27}{\epsilon^2}\mathrm{Var}(S_n).$$ 11 p.215–221

- **(Kolmogorov)** Let $X_1, \ldots, X_n$ be independent random variables with common support such that $\mathbb{E}X_i = 0$ and $\mathrm{Var}(X_i) < \infty$ for $i = 1, \ldots, n$. Let $S_k = \sum_{i=1}^k X_k$ be the kth partial sum. Then for $\epsilon > 0$,

$$\mathbf{Prob}\left[\max_{1 \le k \le n} |S_k| \ge \epsilon\right] \le \frac{1}{\epsilon^2}\sum_{i=1}^n \mathrm{Var}(X_i).$$ 3 Theorem 22.4

- **(Chebychev Multidimensional)** Let $\mathbf{X} \in \mathbb{R}^n$ be a random vector with covariance matrix $V = \mathbb{E}\left[(\mathbf{X} - \mathbb{E}\mathbf{X})(\mathbf{X} - \mathbb{E}\mathbf{X})^T\right]$. Then for $t > 0$,

$$\mathbf{Prob}\left[\sqrt{(\mathbf{X} - \mathbb{E}\mathbf{X})^T V^{-1}(\mathbf{X} - \mathbb{E}\mathbf{X})}\right] \le \frac{n}{t^2}.$$ 27

- **(Leguerre Samuelson)** Let $X_1, \ldots, X_n$ be random variables with common support, and define $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$. Then for $i = 1, \ldots, n$ with probability one,

$$\bar{X} - S\sqrt{n-1} \le X_i \le \bar{X} + S\sqrt{n-1}.$$ 21 p.1522–1525

- **(LeCam)** Suppose $X_1, \ldots, X_n$ are independent binomial random variables with respective success parameters $p_1, \ldots, p_n$. Letting $\lambda_n = \sum_{i=1}^n p_i$, we have

$$\sum_{k=0}^{\infty}\left|\mathbf{Prob}\left[\sum_{i=1}^n X_i = k\right] - \frac{\lambda_n^k e^{-\lambda_n}}{k!}\right| \le 2\sum_{i=1}^n p_i^2.$$ 14 p.1181–1197

- **(Doob Martingale)** Let $\mathbf{X} \in \mathbb{R}^n$ be a *martingale* ($\mathbb{E}(X_i | X_1, \ldots, X_{i-1}) = X_{i-1}$ for $i = 2, \ldots, n$). Then for $C > 0$, $p \ge 1$,

$$\mathbf{Prob}\left[\sup_{1 \le i \le n} X_i \ge C\right] \le \frac{\mathbb{E}X_n^p}{C^p}.$$ 18 (Theorem II.1.7)

5 References

1. P. Bartlett. Lecture on Concentration Inequalities. URL: www.cs.berkeley.edu/ bartlett/courses/281b-sp08/13.pdf, 2008.
2. R. Bhatia and C. Davis. A Better Bound on the Variance. *American Mathematical Monthly (Mathematical Association of America)* 107 (4), 2000.
3. P. Billingsley. *Probability and Measure.* New York: John Wiley, 1995.
4. S. Boucheron, O. Bousquet, G. Lugosi. Concentration inequalities, *Advanced Lectures in Machine Learning.* Springer, 2004.
5. G. Casella, R.L. Berger. *Statistical Inference*, Duxbury, 2002.
6. D.G. Chapman and H. Robbins. Minimum variance estimation without regularity assumptions, *Annals of Mathematical Statistics* 22 (4), 1951.
7. F. Chung, L. Lu. *Complex Graphs and Networks*, AMS, 2006.
8. A.Dembo, T.M. Cover, and J.A. Cover. Information-theoretic inequalities, *IEEE Trans. Inform. Theory* 37 (6), 1991.
9. J. Duchi. Probability Bounds. URL: http://www.cs.berkeley.edu/ jduchi/projects/probability_bounds.pdf.
10. A. Dvoretzky, J. Kiefer, and J. Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator, *Annals of Mathematical Statistics* 27 (3), 1956.
11. N. Etemadi. On some classical results in probability theory, *Sankhy Ser* A 47 (2), 1985.

12. W. Hürlimann. Generalized Algebraic Bounds on Order Statistics Functions, with Application to Reinsurance and Catastrophe. URL: www.actuaries.org/ASTIN/Colloquia/Porto_Cervo/Huerlimann.pdf, 1970?.

13. R. Kannan. A New Probability Inequality Using Typical Moments and Concentration Results. URL: www.crm.umontreal.ca/CARP09/pdf/kannan.pdf, 2009.

14. L. LeCam. An Approximation Theorem for the Poisson Binomial Distribution. *Pacific Journal of Mathematics* 10 (4), 1960.

15. J. Marcinkiewicz and A. Zygmund. Sur les foncions independantes. *Fund. Math.*, 1937. Reprinted in Jzef Marcinkiewicz, *Collected papers*, edited by Antoni Zygmund, Panstwowe Wydawnictwo Naukowe, Warsaw, 1964.

16. M. Mitzenmacher, E. Upfal. *Probability and Computing*, Cambridge, 2005.

17. N. Papadatos. Maximum Variance of Order Statistics. URL: www.ism.ac.jp/editsec/aism/pdf/047_1_0185.pdf, 1994.

18. D. Revuz and M. Yor. *Continuous martingales and Brownian motion (Third ed.)* Berlin: Springer, 1999.

19. H.P. Rosenthal. On the subspaces of L_p 2 spanned by sequences of independent random variables. *Israel J. Math*, 1970.

20. B.W. Silverman. *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC, 1986.

21. P. Samuelson. How Deviant Can You Be?, *Journal of the American Statistical Association*, volume 63, number 324, 1968.

22. T. Tao. Talagrand's Concentration Inequality. URL: http://terrytao.wordpress.com/2009/06/09/talagrands-concentration-inequality, 2009.

23. C. Vasile. Two Generalizations of Popovicius Inequality, *Crux Mathematicorum* 5, 2001.

24. L. Wasserman. *All of Statistics*, Springer, 2004.

25. L. Wasserman. Lecture on Probability Inequalities, www.stat.cmu.edu/ larry/=stat705/Lecture2.pdf, 2008.

26. T. Wolff. Lectures on Harmonic Analysis, *American Mathematical Society*, University Lecture Series, 2003.

27. Wikipedia. Multidimensional Chebychev's Inequality, URL: http://en.wikipedia.org/wiki/Multidimensional_Chebyshev%27s_inequality, 2012.