Disclaimer. Though I try to be precise and correct, errors are inevitable. If you spot an error, please mail me. Otherwise, use with caution.

1 Inequalities

1.1 Markov / Chebychev

Let X be a non-negative random variable. For $\varepsilon > 0$,

$$\mathbb{P}\left[X \geq \varepsilon\right] \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

Applying this to $|X - \mathbb{E}[X]|$ one has Chebychev's inequality:

Let X be *any* random variable. For $\varepsilon > 0$,

$$\mathbb{P}\left[|X - \mathbb{E}(X)| \geq \varepsilon\right] \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}$$

[7]

1.2 Paley-Zygmund

Let X be a non-negative random variable. Then for any $0 \leq \lambda < 1$,

$$\mathbb{P}\left[X > \lambda\,\mathbb{E}\left[X\right]\right] \geq (1-\lambda)^2\,\frac{\left(\mathbb{E}\left[X\right]\right)^2}{\mathbb{E}\left[X^2\right]}.$$

Proof. Using the Cauchy-Schwartz inequality (1.7 below),

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[X\mathbf{1}_{\{X\leq\lambda\,\mathbb{E}[X]\}}\right] + \mathbb{E}\left[X\mathbf{1}_{\{X>\lambda\,\mathbb{E}[X]\}}\right] \leq \lambda\,\mathbb{E}\left[X\right] + \sqrt{\mathbb{E}\left[X^2\right]\mathbb{P}\left[X > \lambda\,\mathbb{E}\left[X\right]\right]}.$$

□

1.3 Azuma (Chernoff / Hoeffding)

Let $(M_n)_{n\geq 0}$ be a martingale with $M_0 = 0$. Assume $|M_n - M_{n-1}| \leq 1$ a.s. for all n. Then for $\lambda > 0$,

$$\mathbb{P}[M_n > \lambda] < \exp\left(-\frac{\lambda^2}{2n}\right),$$

$$\mathbb{P}[|M_n| > \lambda] < 2\exp\left(-\frac{\lambda^2}{2n}\right).$$

[1]

1.4 Bernstein

Let $X_1, X_2, \ldots,$ be i.i.d. random variables such that $\mathbb{E}[X_i] = 0$ and $|X_i| \leq 1$ a.s. Let $\sigma^2 = \mathbb{E}[X_i^2]$, and let $S_n = \sum_{i=1}^n X_i$. Then for any $\lambda > 0$

$$\mathbb{P}[S_n > \lambda] < \exp\left(-\frac{\lambda^2}{2\sigma^2 n + \frac{2}{3}\lambda}\right),$$

$$\mathbb{P}[|S_n| > \lambda] < 2\exp\left(-\frac{\lambda^2}{2\sigma^2 n + \frac{2}{3}\lambda}\right).$$

[3]

1.5 Extension of Hoeffding

Let $X_1, \ldots, X_n$ be real-valued random variables, such that

1. For all i, X_i is not independent of at most d other variables; i.e.,

$$\max\left\{|A| \;\middle|\; A \subset [n]\ ,\ X_i \text{ is not independent of } \{X_j\ :\ j \in A\}\right\} \leq d.$$

2. For all i, $|X_i| \leq b$.

Let $S_n = \sum_1^n X_i$. Then for $\lambda > 0$,

$$\mathbb{P}\left[|S_n - \mathbb{E}\left[S_n\right]| \geq \lambda\right] \leq 2\exp\left(-\frac{\lambda^2}{2nb^2(d+1)}\right).$$

[13]

1.6 Jénsen

Let X be a real valued random variable. If $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

1.7 Cauchy-Schwartz and Hölder

Let $p, q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then,

$$\int |fg|\, d\mu \leq \left(\int |f|^p\, d\mu\right)^{\frac{1}{p}} \cdot \left(\int |g|^q\, d\mu\right)^{\frac{1}{q}}.$$

The Cauchy-Schwartz inequality is Hölder with $p = q = 2$.

1.8 Doob's Maximal L^p-inequality

Let $0 = M_0, M_1, \ldots, M_n$ be a martingale (or positive sub-martingale). Let $M^* = \max_{0\le k\le n} |M_k|$. Then, for $p \ge 1$ and any $\lambda > 0$,

$$\mathbb{P}[M^* \ge \lambda] \le \frac{\|M_n\|_p^p}{\lambda^p},$$

and for any $p > 1$,

$$\|M_n\|_p \le \|M^*\|_p \le \frac{p}{p-1} \|M_n\|_p .$$

[14]

1.9 Harmonic-Geometric-Arithmetic Means

Let $a_1, a_2, \ldots, a_n$ be positive real numbers. Then,

$$\frac{n}{\sum_{j=1}^n a_j^{-1}} \le \prod_{j=1}^n a_j^{1/n} \le \frac{1}{n}\sum_{j=1}^n a_j.$$

[10]

1.10 Prékopa-Leindler

Let $f, g, h : \mathbb{R}^n \to \mathbb{R}$ be non-negative integrable functions and let $0 < \lambda < 1$. If $h((1-\lambda)x + \lambda y) \ge f^{1-\lambda}(x) g^{\lambda}(y)$ for all $x, y \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} h(x)dx \ge \left(\int_{\mathbb{R}^n} f(x)dx\right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g(x)dx\right)^{\lambda}.$$

[8]

1.11 Young and inverse Young

Let $p, q, r > 0$ be such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. Let $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ be non-negative functions. Then,

$$\text{if } p,q,r \ge 1 \text{ then } \|f * g\|_r \le C^n \|f\|_p \|g\|_q ,$$

$$\text{if } p,q,r \le 1 \text{ then } \|f * g\|_r \ge C^n \|f\|_p \|g\|_q ,$$

where $C = \frac{C_p C_q}{C_r}$, and

$$C_s = \sqrt{\frac{|s|^{1/s}}{|s'|^{1/s'}}},$$

for $\frac{1}{s} + \frac{1}{s'} = 1$.

[8]

1.12 Brunn-Minkowski

Let X, Y be non-empty bounded measurable sets in R^n. Let $s, t > 0$ and assume that

$$sX + tY \triangleq \left\{ sx + ty \;\middle|\; x \in X, y \in Y \right\}$$

is also measurable. Then,

$$\text{Vol}(sX + tY)^{1/n} \geq s\text{Vol}(X)^{1/n} + t\text{Vol}(Y)^{1/n}.$$

Equivalently, for any $0 < \lambda < 1$,

$$\text{Vol}(\lambda X + (1 - \lambda)Y) \geq \min \{\text{Vol}(X), \text{Vol}(Y)\}.$$

1.13 Newton

Let $a_1, \ldots, a_n$ be non-zero real numbers (positive or negative). For $0 \leq j \leq n$ let

$$p_j = \frac{1}{\binom{n}{j}} \sum_{\substack{S \subset \{1,\ldots,n\} \\ |S| = j}} \prod_{i \in S} a_i.$$

Then, for all $1 \leq j \leq n - 1$,

$$p_{j-1} p_{j+1} \leq p_j^2.$$

Corollary. Write

$$\prod_{i=1}^{n} (x + a_i) = \sum_{j=0}^{n} \binom{n}{j} p_j x^j.$$

Then, $p_{j-1} p_{j+1} \leq p_j^2$ and if the a_i's are all positive then

$$p_1 \geq p_2^{1/2} \geq \cdots \geq p_n^{1/n}.$$

[10]

1.14 Four Functions Theorem and FKG

Let L be a finite distributive lattice; i.e. L is a partially ordered set, such that for all $x, y, z \in L$,

- There exists a unique element of L that is the maximal lower bound of x, y, denoted by $x \wedge y$ (and called the *meet* of x, y).
- There exists a unique element of L that is the minimal upper bound of x, y, denoted by $x \vee y$ (and called the *join* of x, y).
- (Distributivity)
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

For $X, Y \subset L$ define:

$$X \vee Y = \left\{ x \vee y \;\middle|\; x \in X, y \in Y \right\} \quad X \wedge Y = \left\{ x \wedge y \;\middle|\; x \in X, y \in Y \right\}.$$

For a real valued function $\varphi : L \to \mathbb{R}$ and $X \subset L$ define

$$\varphi(X) = \sum_{x \in X} \varphi(x).$$

Four Functions Theorem (FFT). If $\alpha, \beta, \gamma, \delta : L \to \mathbb{R}^+$ are four non-negative real valued functions of L such that for any $x, y \in L$

$$\alpha(x)\beta(y) \leq \gamma(x \vee y)\delta(x \wedge y),$$

then for any $X, Y \subset L$,

$$\alpha(X)\beta(Y) \leq \gamma(X \vee Y)\delta(X \wedge Y).$$

FKG. Let μ be a probability measure on L such that

$$\mu(x)\mu(y) \leq \mu(x \vee y)\mu(x \wedge y).$$

Let $f, g, h : L \to \mathbb{R}^+$ be non-negative real valued function of L such that f, g are increasing and h is decreasing. Then,

$$\mathbb{E}[fg] \geq \mathbb{E}[f]\,\mathbb{E}[g], \quad \text{and} \quad \mathbb{E}[fh] \leq \mathbb{E}[f]\,\mathbb{E}[h].$$

In other words, f, g are positively correlated, and f, h are negatively correlated.

[1]

1.15 Shearer

For a random vector $X = (X_1, \ldots, X_m)$ and $A \subset \{1, 2, \ldots, m\}$, set $X_A = (X_i : i \in A)$. $H(\cdot)$ is the entropy function.

Let $X = (X_1, \ldots, X_m)$ be a random vector, and let $\mathcal{A}$ be a collection of subsets of $\{1, 2, \ldots, m\}$, possibly with repeats, such that every element of $\{1, 2, \ldots, m\}$ is contained at least t sets in $\mathcal{A}$. Then for any partial order $\prec$ on $\{1, 2, \ldots, m\}$,

$$H(X) \leq \frac{1}{t} \sum_{A \in \mathcal{A}} H\left(X_A \mid (X_i : i \prec A)\right).$$

[5]

1.16 Berry-Esséen

Let $X_1, \ldots, X_n$ be i.i.d. with $\mathbb{E}[X_i] = 0$. Assume that $\sigma \triangleq \mathbb{E}[X_i^2] < \infty$ and $\rho \triangleq \mathbb{E}[|X_i|^3] < \infty$. Let N be a normal random variable with mean 0 and variance 1. Set $X = \sum_{i=1}^n X_i$. Then,

$$\left|\mathbb{P}\left[X > \lambda\sigma\sqrt{n}\right] - \mathbb{P}\left[N > \lambda\right]\right| = \left|\mathbb{P}\left[X \leq \lambda\sigma\sqrt{n}\right] - \mathbb{P}\left[N \leq \lambda\right]\right| \leq \frac{\rho}{\sigma^3\sqrt{n}}.$$

[7]

1.17 Chen-Stein Method (simplified)

Let $X_1, \ldots, X_N$ be N Bernoulli random variable with $p_i = \mathbb{E}\left[X_i\right] > 0$, and let $S_N = \sum_{i=1}^N X_i$. Let $G = (V, E)$ be a graph on the vertex set $V = \{1, \ldots, N\}$, such that $\{i, i\} \in E$ is an edge (self loop) for all $i \in V$. Assume that for all $i \in V$,

$$\mathbb{E}\left[X_i \mid \sum_{j \not\sim i} X_j\right] = p_i.$$

(e.g. this holds if X_i is independent of the set $\{X_j \ : \ j \not\sim i\}$.)

Let Z be a Poisson random variable with mean

$$\mathbb{E}\left[Z\right] = \lambda = \mathbb{E}\left[S_N\right] = \sum_{i=1}^N p_i.$$

Define

$$B_1 = \sum_{i=1}^N \sum_{j \sim i} \mathbb{E}\left[X_i\right] \mathbb{E}\left[X_j\right],$$

$$B_2 = \sum_{i=1}^{N} \sum_{\substack{j \sim i \\ j \neq i}} \mathbb{E}\left[X_i X_j\right].$$

Then,

$$\sup_{A \subseteq \mathbb{N}} \left|\mathbb{P}\left[S_N \in A\right] - \mathbb{P}\left[Z \in A\right]\right| \leq \frac{1 - e^{-\lambda}}{\lambda} \cdot (B_1 + B_2).$$

[2]

1.18 Suen's Inequality

For a graph $G = (V, E)$ write $i \sim j$ if $\{i, j\}$ is an edge. For subsets $A, B \subseteq V$, write $A \sim B$ if there exists an edge between A and B. Thus, $A \not\sim B$ means that there is no edge between A and B. $i \sim A$ means there is an edge between i and some element of A.

Let $X_1, \ldots, X_N$ be N Bernoulli random variables, and let $S_N = \sum_{i=1}^{N} X_i$.

Let G be a *dependency graph* of $\{X_i\}_{i=1}^{N}$; i.e. $G = (V, E)$ is a graph on the vertex set $V = \{1, \ldots, N\}$ such that for any two disjoint subsets $A, B \subset V$, if $A \not\sim B$ then the two families $\{X_i\}_{i \in A}$ and $\{X_i\}_{i \in B}$ are independent of each other. (In some texts G is called a superdependency digraph.)

Define

$$\Delta = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \sim i} \mathbb{E}\left[X_i X_j\right] \prod_{k \sim \{i,j\}} \left(1 - \mathbb{E}\left[X_k\right]\right)^{-1}.$$

$$\Delta^* = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \sim i} \mathbb{E}\left[X_i\right] \mathbb{E}\left[X_j\right] \prod_{k \sim \{i,j\}} \left(1 - \mathbb{E}\left[X_k\right]\right)^{-1}.$$

Then,

$$\mathbb{P}\left[S_N = 0\right] \leq e^{\Delta} \prod_{i=1}^{N} \left(1 - \mathbb{E}\left[X_i\right]\right),$$

$$\mathbb{P}\left[S_N = 0\right] \geq \left(1 - \Delta^* e^{\Delta}\right) \prod_{i=1}^{N} \left(1 - \mathbb{E}\left[X_i\right]\right).$$

Note that if $\{X_i\}_{i=1}^{N}$ are all independent, then equality holds in both inequalities (choosing the empty graph).

[12, 16]

2 Convergence Theorems

2.1 The Law of the Iterated Logarithm

Let $X_1, X_2, \dots$ be i.i.d. with $\mathbb{E}[X_i] = 0$ and $0 < \sigma^2 \triangleq \mathbb{E}[X_i^2] < \infty$. Set $S_n = \sum_1^n X_i$. Then,

$$\mathbb{P}\left[\limsup \frac{S_n}{\sqrt{2\sigma^2 n \log\log n}} = 1\right] = 1.$$

[15]

2.2 Monotone Convergence

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $0 \le f_1 \le f_2 \le \cdots$ be a monotone sequence of non-negative measurable functions. Then,

$$f = \lim_{n\to\infty} f_n$$

exists a.e., and is a measurable function. Further,

$$\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu.$$

[6]

2.3 Fatou's Lemma

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of non-negative measurable functions. Then,

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \le \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

[6]

2.4 Dominated Convergence

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $\{f_n\}, f, g$ be measurable functions such that:

$$\begin{aligned} &(1) \qquad f = \lim_{n\to\infty} f_n \quad \text{a.e.,} \\ &(2) \qquad \forall\ n \quad |f_n| \le g \quad \text{a.e.,} \\ &(3) \qquad \int_\Omega g d\mu < \infty. \end{aligned}$$

Then,

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu.$$

Remark: The condition of a.e. convergence in (1), can be replaced by convergence in measure; i.e. it suffices to require that

$$(1) \qquad \forall\, \varepsilon > 0 \quad \lim_{n\to\infty} \mu\left\{|f_n - f| > \varepsilon\right\} = 0.$$

[6]

3 Formulas

3.1 Stirling's Formula

$$n! \sim \sqrt{2\pi} n^{(n+1/2)} e^{-n}$$

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n}}$$

For $0 < \alpha < 1$,

$$\binom{n}{\alpha n} = \left(1 \pm O\left(n^{-1}\right)\right) \cdot C \cdot \frac{1}{\sqrt{n}} \cdot 2^{nH(\alpha)},$$

where $H(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$, and the constant is $C = \frac{1}{\sqrt{2\pi\alpha(1-\alpha)}}$.

Less accurate, but easier to use:

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

3.2 $\left(1 - \frac{1}{x}\right)$ and e:

For all $x \geq 1$,

$$\left(1 - \frac{1}{x}\right)^x \leq e^{-1} \leq \left(1 - \frac{1}{x+1}\right)^x.$$

3.3 Finite sums of powers

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{k=1}^{n} k^3 = \left[\frac{n(n+1)}{2}\right]^2$$

$$\sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$$

[9]

3.4 Permutation fixed points

Let $D(n,m)$ denote the number of permutations of $\{1,\ldots,n\}$ that have exactly m fixed points. Then,

$$D(n,m) = \frac{n!}{m!}\sum_{k=0}^{n-m}\frac{(-1)^k}{k!}.$$

The proof uses the inclusion-exclusion principle, see e.g.
`Wikipedia: Random permutation statistics` .

4 Complex Analysis

Denote the unit disc in the complex plane by $\mathbb{U}$;

$$\mathbb{U} = \{z \in \mathbb{C} \ : \ |z| < 1\}.$$

4.1 Schwartz Lemma

Let $f : \mathbb{U} \to \mathbb{C}$ be a bounded analytic function with $f(0) = 0$. Then,

$$|f(z)| \leq |z| \cdot \|f\|_\infty \qquad \text{and} \qquad |f'(0)| \leq \|f\|_\infty.$$

If there exists $z \neq 0$ such that equality holds for one of the above, then there exists $\theta \in [0, 2\pi)$ such that $f(z) = e^{i\theta}\|f\|_\infty z$ for all z.

[4]

4.2 Bieberbach Conjecture (De Branges Theorem)

Let $f : \mathbb{U} \to \mathbb{C}$ be a conformal map of the unit disc (i.e. f is injective and analytic in the unit disc). Then for all $n \geq 1$,

$$|f^{(n)}(0)| \leq n \cdot n! \cdot |f'(0)|.$$

[4]

4.3 Koebe $\frac{1}{4}$ Theorem

Let $f : \mathbb{U} \to \mathbb{C}$ be a conformal map of the unit disc (i.e. f is injective and analytic in the unit disc). Then,

$$\left\{ z \in \mathbb{C} \ : \ |z - f(0)| < \tfrac{|f'(0)|}{4} \right\} = \tfrac{1}{4}|f'(0)|\mathbb{U} + f(0) \subseteq f(\mathbb{U}).$$

Furthermore, let $d_f(z)$ be the distance of $f(z)$ from $\partial f(\mathbb{U})$. Then,

$$\tfrac{1}{4} \cdot |f'(z)| \cdot (1 - |z|^2) \leq d_f(z) \leq |f'(z)| \cdot (1 - |z|^2).$$

[4]

4.4 Koebe Distortion Theorem

Let $f : \mathbb{U} \to \mathbb{C}$ be a conformal map of the unit disc (i.e. f is injective and analytic in the unit disc). Then,

$$\frac{|z|}{(1 + |z|)^2} \cdot |f'(0)| \leq |f(z) - f(0)| \leq \frac{|z|}{(1 - |z|)^2} \cdot |f'(0)|,$$

$$\frac{1 - |z|}{(1 + |z|)^3} \cdot |f'(0)| \leq |f'(z)| \leq \frac{1 + |z|}{(1 - |z|)^3} \cdot |f'(0)|.$$

[4]

References

[1] N. Alon, J. H. Spencer, *The Probabilistic Method* (2000), John Wiley & Sons, Inc.

[2] R. Arratia, L. Goldstein, L. Gordon, Two moments suffice for Poisson approximations: The Chen-Stein method. *Ann. Probab.* **17** (1989), 9–25.

[3] N. Berestycki, R. Nickl, *Concentration of Measure* (2009), Section 1.5, lecture notes. `available here`

[4] J. B. Conway, *Functions of One Complex Variable II* (1995), Springer-Verlag.

[5] T. M. Cover, J. A. Thomas, *Elements of Information Theory* (1991), John Wiley & Sons, Inc.

[6] J. L. Doob, *Measure Theory* (1994), Springer.

[7] W. Feller, *Introduction to Probability Theorey and its Applications* (1966), John Wiley & Sons, Inc.

[8] R. J. Gardner, The Brunn-Minkowski Inequality, *Bulletin of the American Mathematical Society* **39** (2002), no. 3, 355–405. `available here`

[9] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products* (1965), New York : Academic Press.

[10] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities* (1952), Cambridge University Press.

[11] F. den Hollander, *Large Deviations* (2000), AMS.

[12] S. Janson, New versions of Suen's correlation inequality. *Random Structures and Algorithms* **13** (1998), 467–483.

[13] S. Janson, Large deviations for sums of partly dependent random variables. *Random Structures and Algorithms* **24** (2004), no. 3, 234–248. `available here`

[14] D. Revuz, M. Yor, *Continuous martingales and Brownian motion* (1991), Springer-Verlag.

[15] A. N. Shiryaev, *Probability*, Translated by R. P. Boas (1996), Springer.

[16] W. C. S. Suen, A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph. *Random Structures and Algorithms*, **1** (1990), 231–242.