Some mathematical formulae

useful in my research

Abbas Mehrabian

University of British Columbia

January 15, 2018

To be added...

(a) Lóvasz Local lemma and variations, e.g. Harvey-Vondrak paper
(b) Tropp's concentration inequalities for sum of i.i.d. matrices, e.g. Tropp's monograph
(c) Suen's inequality, e.g. Janson's paper
(d) CHERNOFF BOUND FOR RANDOM WALKS ON EXPANDER GRAPHS, Gilman's paper

Contents

General remarks.

(a) This document is for personal use. Of course I wouldn't mind if someone else uses it as well, but use at your own risk ;-) Found errors? Kindly email me: *first name followed by last name at gmail dot com.*

(b) The references given for each result is NOT necessarily the first place where the result has been proven. But rather, I try to provide a reference which (i) has a proof, and (ii) is easy to access, e.g. is available on-line, and is published.

(c) Another set of useful results can be found in [11, Part Four].

(d) All log's are in base e.

1 Function Approximations

(a)

$$\begin{aligned}
\exp(x) &\geq 1 + x && \forall x, \\
\exp(x) &\leq 1 + x + x^2 && \forall x \in [0, 1.7932], \\
\exp(-x - x^2) &\leq 1 - x && \forall x \in [0, 0.6838].
\end{aligned}$$

(b)

$$\begin{aligned}
\log(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \ldots \qquad \forall |x| < 1, \\
\log(x) &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \ldots \qquad \forall |x-1| < 1.
\end{aligned}$$

(c) Stirling's formula ([9, equation 9.15 in Chapter II]):

$$\sqrt{2\pi n}(n/e)^n \exp(1/(12n+1)) < n! < \sqrt{2\pi n}(n/e)^n \exp(1/12n) \qquad \forall n \in \mathbb{Z}_+ \,,$$

which gives

$$\log n! = n \log n - n + (\log 2\pi n)/2 + O(1/n) \,.$$

For real n, [12, equation 8.327] or [1, equation 6.1.37] gives

$$\Gamma(x) = x^{x-1/2} e^{-x} \left(1 + \frac{1}{12x} + O(x^{-2})\right) \sqrt{2\pi} \qquad \forall x > 0.$$

(d) Inequalities for the Gamma function: ([19, equation (2.2)])

$$\frac{\Gamma(x+\lambda)}{\Gamma(x+1)} < (x + \lambda/2)^{\lambda - 1} \qquad \forall x \geq 0, \lambda \in (0,1) \cup (2, \infty) \,,$$

and ([19, equation (2.3)])

$$\frac{\Gamma(x+\lambda)}{\Gamma(x+1)} > (x + \lambda/2)^{\lambda - 1} \qquad \forall x \geq 0, \lambda \in (1,2) \,,$$

(e) Harmonic numbers: For every positive integer n we have ([33])

$$\frac{1}{2n+2} < \sum_{i=1}^{n} \frac{1}{i} - \log n - \gamma < \frac{1}{2n} \,,$$

where $\gamma \approx 0.57721$ is Euler's constant.

(f) If G is a connected n-vertex graph with maximum degree $\Delta > 0$ and diameter $D > 0$, then the bound $n < 2\Delta^D$ follows e.g. from Moore bound, see `https://en.wikipedia.org/wiki/Degree_diameter_problem`

(g) This bound for binomial coefficient comes in handy: $\sum_{i=0}^{k} \binom{n}{i} \leq (en/k)^k$ holds for all positive integers $1 \leq k \leq n$, see [4, Exercise 2.14].

(h) [1, 7.1.13] Bounds for the standard Gaussian CDF: Let Z be Gaussian with mean 0 and variance 1. Then, for any $t > 0$ we have

$$\frac{e^{-t^2/2}}{t + \sqrt{t^2 + 4}} \leq \sqrt{\pi/2}\, \Pr(Z > t) \leq \frac{e^{-t^2/2}}{t + \sqrt{t^2 + 8/\pi}}.$$

See `https://arxiv.org/pdf/1012.2063.pdf` for more such bounds.

2 Concentration Inequalities

(a) (**Markov Inequality**) If X is a nonnegative random variable then

$$\mathbf{Pr}[X > t] < \mathbb{E}[X]/t.$$

(b) (**Chebyshev Inequality**) If X is a nonnegative random variable then

$$\mathbf{Pr}[|X - \mathbb{E}[X]| > t] < \mathbf{Var}[X]/t^2.$$

(c) (**Cramér's Theorem** [6, Theorem I.4 and Comments (1), (4), and (5) in Section I.4]) Let $X_1, X_2, \ldots$ be i.i.d. real-valued random variables, and define

$$I(z) = \sup\{zt - \log \mathbb{E}\left[e^{tX_1}\right] : t \in \mathbb{R}\}.$$

For any $a > \mathbb{E}[X_1]$, as n grows we have

$$\Pr(X_1 + \cdots + X_n \geq an) = \exp((-I(a) \pm o(1))n),$$

and for any $a < \mathbb{E}[X_1]$, as n grows we have

$$\Pr(X_1 + \cdots + X_n \leq an) = \exp((-I(a) \pm o(1))n),$$

2.1 Chernoff-Type Inequalities (sums of bounded variables)

Let $X = X_1 + X_2 + \cdots + X_n$ with X_i be independent and bounded in $[0, 1]$ and let $\mu = \mathbb{E}[X]$. The following Chernoff-driven inequalities are true even if $X = X_1 + X_2 + \cdots + X_n$, where $X_1, X_2, \ldots, X_n$ are bounded in $[0, 1]$, and for all subsets S we have

$$\mathbf{Pr}\left[\bigcap_{i \in S} \{X_i = 1\}\right] \leq \prod_{i \in S} \mathbf{Pr}[X_i = 1].$$

(a) (Basic Chernoff Bound) Let $p = \mu/n$.

$$\mathbf{Pr}[X > (p + t)n] < \left[\left(\frac{p}{p+t}\right)^{p+t} \left(\frac{q}{q-t}\right)^{q-t}\right]^n,$$

and

$$\mathbf{Pr}[X < (p - t)n] < \left[\left(\frac{q}{q+t}\right)^{q+t} \left(\frac{p}{p-t}\right)^{p-t}\right]^n.$$

Another, perhaps nicer way to write the above inequalities follows. Define

$$J(x, p) = x \ln\left(\frac{x}{p}\right) + (1 - x) \ln\left(\frac{1-x}{1-p}\right), \qquad \text{for } (x, p) \in (0, 1)^2.$$

Then for all $x \in (0, p)$ we have

$$\Pr(Z_n < nx) < e^{-nJ(x,p)},$$

and for all $x \in (p, 1)$ we have

$$\Pr(Z_n > nx) < e^{-nJ(x,p)}.$$

(b) ([8] Theorem 1.1)

$$\mathbf{Pr}[X > \mu + t], \mathbf{Pr}[X < \mu - t] < \exp(-2t^2/n) \qquad \forall t > 0,$$

and

$$\mathbf{Pr}[X < (1-\epsilon)\mu] < \exp(-\epsilon^2\mu/2) \qquad \forall \epsilon > 0,$$

and

$$\mathbf{Pr}[X > t] < 2^{-t} \qquad \forall 0 < t < 2e\mu.$$

(c) ([24] Theorem 2.3(b)) For all $\epsilon > 0$,

$$\mathbf{Pr}[X \geq (1+\epsilon)\mu] \leq \exp(\mu(\epsilon-(1+\epsilon)\log(1+\epsilon))) \leq \exp\left(-\frac{\epsilon^2\mu}{2+2\epsilon/3}\right) \leq \exp(-\epsilon^2(1-\epsilon)\mu/2),$$

and the leftmost inequality gives

$$\mathbf{Pr}[X \geq (1+\epsilon)\mu] < \exp(-\mu\epsilon^2/3) \qquad \forall\ 0 < \epsilon \leq 1.81,$$

and (see [28, Exercise 4.1])

$$\mathbf{Pr}[X > (1+\epsilon)\mu] < 2^{-(1+\epsilon)\mu} \qquad \forall\ \epsilon > 2e-1\,.$$

Moreover, (see [5, Theorem 2.17])

$$\mathbf{Pr}[X < \epsilon\mu] < \exp(-\mu + 2\mu\epsilon(1-\log\epsilon)) \qquad \forall\ 0 \leq \epsilon \leq 1/e\,.$$

(d) ([29] Lemma 1.1) Define $H : [0,\infty) \to [0,\infty)$ as $H(0) := 0$ and $H(a) := 1 - a + a\log a$. Let $p \in (0,1)$ and $0 < k < n$. If $k \geq \mu$ then

$$\mathbf{Pr}[X \geq k] \leq \exp(-\mu H(k/\mu)),$$

and if $k \leq \mu$ then

$$\mathbf{Pr}[X \leq k] \leq \exp(-\mu H(k/\mu)).$$

Finally, if $k \geq e^2\mu$ then

$$\mathbf{Pr}[X \geq k] \leq \exp\left(-\frac{k}{2}\log(k/\mu)\right).$$

(e) ([8] Theorem 1.2, Bernstein's inequality) Let $X_1, \ldots, X_n$ be independent with $X_i - \mathbb{E}[X_i] \leq b$ for all i. Let $X = \sum X_i$ and let σ^2 be the variance of X. For any $t > 0$,

$$\mathbf{Pr}[X > \mathbb{E}[X] + t] \leq \exp\left(-\frac{t^2}{2\sigma^2(1+bt/3\sigma^2)}\right).$$

(f) ([30] Theorem 5) If X is the sum of k-wise independent random variables taking values in $[0,1]$, and $\mu = \mathbb{E}[X]$, then

$$\begin{aligned}
&\mathbf{Pr}(|X-\mu| > \epsilon\mu) < \exp(-\lfloor k/2 \rfloor) \qquad \forall \epsilon \le 1, k \le \lfloor \epsilon^2 \mu e^{-1/3} \rfloor \\
&\mathbf{Pr}(|X-\mu| > \epsilon\mu) < \exp(-\lfloor \epsilon^2 \mu/3 \rfloor) \qquad \forall \epsilon \le 1, k \ge \lfloor \epsilon^2 \mu e^{-1/3} \rfloor \\
&\mathbf{Pr}(|X-\mu| > \epsilon\mu) < \exp(-\lfloor k/2 \rfloor) \qquad \forall \epsilon \ge 1, k \le \lfloor \epsilon \mu e^{-1/3} \rfloor \\
&\mathbf{Pr}(|X-\mu| > \epsilon\mu) < \exp(-\lfloor \epsilon \mu/3 \rfloor) \qquad \forall \epsilon \ge 1, k \ge \lfloor \epsilon \mu e^{-1/3} \rfloor \\
&\mathbf{Pr}(|X-\mu| > \epsilon\mu) < \exp(-\epsilon \ln(1+\epsilon)\mu/2) < \exp(-\epsilon\mu/3) \qquad \forall \epsilon \ge 1, k \ge \lceil \epsilon\mu \rceil
\end{aligned}$$

(g) ([3] Lemmas 2.2 and 2.3) Let k be an even integer, and let X be the sum of n k-wise independent random variables taking values in $[0,1]$. Let $\mu = \mathbb{E}[X]$ and $a > 0$. Then we have

$$\begin{aligned}
\mathbf{Pr}[|X-\mu| > a] &< 1.0004 \left(\frac{nk}{a^2}\right)^{k/2} \\
\mathbf{Pr}[|X-\mu| > a] &< 8 \left(\frac{k\mu + k^2}{a^2}\right)^{k/2}.
\end{aligned}$$

2.2 Martingale-Based Inequalities

(a) (**[24] Theorem 3.1**) Let $\overrightarrow{X} = (X_1, X_2, \ldots, X_n)$, where X_i's are independent random variables, with $X_i \in A_i$. Suppose that the real-valued function f defined on $\prod A_i$ satisfies

$$|f(\overrightarrow{x}) - f(\overrightarrow{y})| \le c_i,$$

whenever the vectors $\overrightarrow{x}$ and $\overrightarrow{y}$ differ only in the i-th coordinate. Then for any $t \ge 0$,

$$\mathbf{Pr}\Big(f\big(\overrightarrow{X}\big) - \mathbb{E}\Big[f\big(\overrightarrow{X}\big)\Big] < -t\Big), \mathbf{Pr}\Big(f\big(\overrightarrow{X}\big) - \mathbb{E}\Big[f\big(\overrightarrow{X}\big)\Big] > t\Big) < \exp\Big(-2t^2/\sum c_i^2\Big).$$

(b) (**[24] Theorem 3.7**) Let $\overrightarrow{X} = (X_1, X_2, \ldots, X_n)$, where X_i's are random variables, with $X_i \in A_i$. Suppose that the real-valued function f defined on $\prod A_i$ satisfies

$$|\mathbb{E}[f|X_1 = a_1, \ldots, X_{i-1} = a_{i-1}, X_i = x_i] - \mathbb{E}[f|X_1 = a_1, \ldots, X_{i-1} = a_{i-1}, X_i = y_i]| \le c_i,$$

for all $a_1, a_2, \ldots, a_{i-1}, x_i, y_i$ for which the LHS is well-defined. Then for any $t \ge 0$,

$$\mathbf{Pr}\Big(f\big(\overrightarrow{X}\big) - \mathbb{E}\Big[f\big(\overrightarrow{X}\big)\Big] < -t\Big), \mathbf{Pr}\Big(f\big(\overrightarrow{X}\big) - \mathbb{E}\Big[f\big(\overrightarrow{X}\big)\Big] > t\Big) < \exp\Big(-2t^2/\sum c_i^2\Big).$$

(c) (**[8] Theorem 5.2, Azuma-Hoeffding inequality for supermartingales)** Let $X_0, \ldots, X_n$ be random variables, and

$$Y_i = g_i(X_0, X_1, \ldots, X_i) \qquad i = 0, 1, \ldots, n$$

be such that

$$\mathbb{E}[Y_i|X_0, \ldots, X_i] \le Y_{i-1} \qquad \forall 1 \le i \le n.$$

Suppose further that

$$a_i \leq Y_i - Y_{i-1} \leq b_i \qquad \forall 1 \leq i \leq n.$$

Then for any $t \geq 0$,

$$\mathbf{Pr}(Y_n > Y_0 + t] < \exp\Big(-2t^2 / \sum (b_i - a_i)^2\Big).$$

(d) (**[25] Azuma-Hoeffding inequality for centering sequences)** Let $0 = X_0, X_1, X_2, \ldots, X_n$ be a sequence and let $Y_k = X_k - X_{k-1}$ for $1 \leq k \leq n$. Assume that $\mathbf{E}[Y_k | X_{k-1} = x]$ is a non-increasing function of x. (If this condition is satisfied then (X_i) is called a *centering* sequence.)
(a) (Theorem 2.2 in [25]) If $0 \leq Y_k \leq 1$ for each k, then

$$\begin{aligned}
\mathbf{Pr}[X_n - \mathbf{E}[X_n] > t] &< \exp(-2t^2/n) \qquad \forall t > 0, \\
\mathbf{Pr}[X_n - \mathbf{E}[X_n] < -t] &< \exp(-2t^2/n) \qquad \forall t > 0, \\
\mathbf{Pr}[X_n > (1+\epsilon)\mathbf{E}[X_n]] &< \exp(-\epsilon^2 \mathbf{E}[X_n]/3) \qquad \forall 0 < \epsilon \leq 1, \\
\mathbf{Pr}[X_n < (1-\epsilon)\mathbf{E}[X_n]] &< \exp(-\epsilon^2 \mathbf{E}[X_n]/2) \qquad \forall 0 < \epsilon \leq 1.
\end{aligned}$$

(b) (Theorem 2.3 in [25]) If $a_k \leq Y_k \leq b_k$ for all k, then for any $t > 0$,

$$\begin{aligned}
\mathbf{Pr}[X_n - \mathbf{E}[X_n] > t] &< \exp\Big(-2t^2 / \sum (b_k - a_k)^2\Big), \\
\mathbf{Pr}[X_n - \mathbf{E}[X_n] < -t] &< \exp\Big(-2t^2 / \sum (b_k - a_k)^2\Big).
\end{aligned}$$

(c) (Concluding remarks of [25]) If $a_k \leq Y_k \leq b_k$ for all k, then for any $t > 0$,

$$\mathbf{Pr}[|X_n - \mathbf{E}[X_n]| > t] < \left(\frac{\sum (b_k - a_k)^2}{2t}\right)^2.$$

(This may be better than (b) only for very small $t > 0$.)

More inequalities of the type given in Sections 2.1 and 2.2 can be found in [5, Chapter 2].

2.3 Sums of Poisson variables

Let $X \sim \mathbf{Po}(\lambda)$.

(a) ([2] Theorem A.1.15)

$$\begin{aligned}
&\mathbf{Pr}[X < (1-\epsilon)\lambda] < \exp(\epsilon^2 \lambda / 2), \\
&\mathbf{Pr}[X > (1+\epsilon)\lambda] < \exp(\lambda(\epsilon - (1+\epsilon)\log(1+\epsilon))).
\end{aligned}$$

(b) ([29] Lemma 1.2) Let $H(a) := 1 - a + a\log a$, $k, \lambda > 0$. If $k \geq \lambda$ then

$$\mathbf{Pr}[X \geq k] \leq \exp(\lambda H(k/\lambda)),$$

and if $k \leq k$ then

$$\mathbf{Pr}[X \leq k] \leq \exp(\lambda H(k/\lambda)),$$

and if $k \geq e^2\lambda$ then

$$\mathbf{Pr}[X \geq k] \leq \exp\left(-\frac{k}{2}\log(k/\lambda)\right).$$

(c) (**[32, Exercise 2.7]**) Let $X_1, X_2, \ldots$ be independent Poisson variables with mean λ, and let $I(a) = a\log(a/\lambda) - a + \lambda$. If $a > \lambda$ then

$$\Pr(X_1 + \cdots + X_n \geq na) \leq e^{-nI(a)},$$

and if $a < \lambda$ then

$$\Pr(X_1 + \cdots + X_n \leq na) \leq e^{-nI(a)}.$$

Moreover, $I(a) > 0$ for all $a \neq \lambda$.

2.4 Sums of exponential variables

Note: for some clean lower and upper bounds for sums of exponentials and sums of geometrics, see Janson's paper, TAIL BOUNDS FOR SUMS OF GEOMETRIC AND EXPONENTIAL VARIABLES, available at `http://www2.math.uu.se/~svante/papers/sjN14.pdf`

(a) (**[27, Lemma 6]**) Let $\Upsilon(x) = x - 1 - \log(x)$ and let $E_1, E_2, \ldots, E_m$ be independent exponential random variables with mean 1. For any fixed $0 < x < 1$, as $m \to \infty$ we have

$$\exp(-\Upsilon(x)m - o(m)) \leq \Pr(E_1 + E_2 + \cdots + E_m \leq xm) \leq \exp(-\Upsilon(x)m)$$

(this is what Cramér's Theorem gives, so is almost tight).

2.5 Sums of geometric variables

Let $p \in (0, 1)$ and let $Z_1, Z_2, \ldots, Z_m$ be independent geometric random variables with parameter p and mean $1/p$, namely for every positive integer s, $\Pr(Z_1 = s) = (1-p)^{s-1}p$.

(a) (**[27, Lemma 7]**) If $r \geq 1/p$, then $\Pr(Z_1 + Z_2 + \cdots + Z_m \geq rm) \leq \left(r^r p(1-p)^{r-1}(r-1)^{1-r}\right)^m$ (this is what Cramér's Theorem gives, so is almost tight).

(b) (**[10, Lemma 21]**) For any $\varepsilon > 0$,

$$\Pr\left(Z_1 + \cdots + Z_m \geq (1+\varepsilon)\frac{m}{p}\right) \leq \exp\left(-\frac{\varepsilon^2}{2(1+\varepsilon)}m\right).$$

3 Correlation Inequalities

The treatment here is from [15, Section 5], which is essentially the same as that in [14, Section 4] and [13, Section 2]. Let E be a finite non-empty set, and let $\Omega = \Omega_E = \{0,1\}^E$. A probability distribution μ on Ω_E is called *positive* if $\mu(\omega) > 0$ for all $\omega \in \Omega_E$. For $a, b \in \Omega_E$, $\max\{a,b\}$ and $\min\{a,b\}$ denote the component-wise maximum and minimum (i.e. bit-wise OR and bit-wise AND). A random variable $X : \Omega_E \to \mathbb{R}$ is *increasing* if flipping a bit from 0 to 1 does not decrease the value of X. An event $A \subseteq \Omega_E$ is increasing if its indicator function is increasing.

(a) (**FKG inequality, Theorem 5.1 in [15]**) Let μ be a positive probability distribution on Ω_E such that for all $a, b \in \Omega_E$,

$$\mu(\max\{a,b\})\mu(\min\{a,b\}) \geq \mu(a)\mu(b) . \tag{1}$$

(For example, the product measure is positive and satisfies this condition.) Then for any increasing random variables X and Y,

$$\mathbb{E}[[]\mu]XY \geq \mathbb{E}[[]\mu]X\mathbb{E}[[]\mu]Y .$$

For example, if A and B are increasing events, FKG inequality gives $\mu(A \cap B) \geq \mu(A)\mu(B)$.

(b) (**Holley's inequality, Theorem 5.5 in [15]**) Let μ_1 and μ_2 be positive probability distributions on Ω_E such that for all $a, b \in \Omega_E$,

$$\mu_1(\max\{a,b\})\mu_2(\min\{a,b\}) \geq \mu_1(a)\mu_2(b) .$$

Then for any increasing random variable X,

$$\mathbb{E}[[]\mu_1]X \geq \mathbb{E}[[]\mu_2]X .$$

For the following two inequalities, we consider the product measure on Ω_E: suppose $\{p_e\}_{e\in E}$ are given, and define

$$\mathbb{P}[\omega] := \prod_{e:\omega(e)=1} p_e \prod_{e:\omega(e)=0} (1-p_e) .$$

(c) (**BK inequality, [15, Theorem 5.11]**) For $F \subseteq E$ and $\omega \in \Omega_E$ define $\omega_F \in \Omega_E$ as

$$\omega_F(e) = \begin{cases} \omega(e) & \text{if } e \in F \\ 0 & \text{if } e \notin F , \end{cases}$$

and for increasing events A and B define

$$A \circ B := \Big\{\omega : \text{there exists } F \subseteq E \text{ such that } \omega_F \in A \text{ and } \omega_{E\setminus F} \in B\Big\} .$$

(The canonical example in percolation theory is the existence of edge-disjoint paths.)
Then, for increasing events A and B we have

$$\Pr(A \circ B) \leq \Pr(A)\Pr(B) .$$

(d) (**Reimer's inequality, [15, Theorem 5.12]**) For $\omega \in \Omega_E$ and $F \subseteq E$ define the cylinder event

$$C(\omega, F) = \{\omega' : \omega'(e) = \omega(e) \text{ for } e \in F\} ,$$

and for events A and B define

$$A \Box B = \{\omega : \text{ there exists } F \subseteq E \text{ such that } C(\omega, F) \subseteq A \text{ and } C(\omega, E \setminus F) \subseteq B\} .$$

In words, this is the set of ω for which there exists $F \subseteq E$ such that agreeing with ω on F guarantees A happens, and agreeing on $E \setminus F$ guarantees B happens. Then for any two events A and B we have

$$\Pr(A \Box B) \leq \Pr(A) \Pr(B) .$$

4 Other Probability Bounds

(a) (**[2] Theorems 8.1.1 and 8.1.2, The (Extended) Janson Inequality**) Let Ω be a finite universal set, and let R be a random subset of Ω given by $\mathbf{Pr}[r \in R] = p_r$, these events mutually independent. Let $A_1, \dots, A_n$ be subsets of Ω, and B_i be the event $A_i \subseteq R$. Write $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Let

$$\Delta = \sum_{i<j} \mathbf{Pr}[B_i \text{ AND } B_j], M = \prod (1 - \mathbf{Pr}[B_i]), \mu = \sum \mathbf{Pr}[B_i],$$

and assume that $\mathbf{Pr}[B_i] \leq \epsilon$ for all i. Then

$$M \leq \mathbf{Pr}[\text{no } B_i \text{ occurs}] \leq M \exp\left(\frac{\Delta}{1-\epsilon}\right),$$

and

$$\mathbf{Pr}[\text{no } B_i \text{ occurs}] \leq \exp(\Delta - \mu).$$

If also $\Delta \geq \mu$ then

$$\mathbf{Pr}[\text{no } B_i \text{ occurs}] \leq \exp\left(-\frac{\mu^2}{4\Delta}\right).$$

5 Eigenvalues of graphs, random walks and graph expansion

It is known that for a given graph, there are connections between combinatorial expansion, mixing rate of random walks, and eigenvalues. Here are some relevant results. For other results and references, see [22, Section 3] (a 1995 survey, perhaps not up to date!).

Let P denote the transition probability matrix of an irreducible reversible Markov chain with finite state space X, and suppose the spectrum of P is

$$1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{|X|} \geq -1.$$

The facts that all eigenvalues are real and lie in $[-1, 1]$ are well known (see, e.g., [7, first page]). (Also, $\lambda_{|X|} > -1$ if and only if the chain is aperiodic.) Let π denote the stationary distribution, and for $S \subseteq X$ define $\pi(S) := \sum_{x\in S} \pi(x)$, and let $\pi_{\min} := \min\{\pi(x) : x \in X\}$. The *total variation distance* between two distributions μ and π is

$$\|\mu - \pi\| = \max\{|\mu(A) - \pi(A)| : A \subseteq X\} = \frac{1}{2}\sum_{x\in X} |\mu(x) - \pi(x)| \,.$$

We define the *conductance* of the chain as

$$\Phi := \min\left\{ \frac{\sum_{(x,y)\in S\times S^c} \pi(x)P(x,y)}{\pi(S)} : S \subseteq X, 0 < \pi(S) \le 1/2 \right\} \tag{2}$$

A reversible Markov chain is equivalent to a random walk on a weighted undirected graph (with all weights positive, and with possible self-loops, but no parallel edges). The chain is irreducible/aperiodic if and only if the graph is connected/non-bipartite. The transition probability matrix is also called the random walk matrix of the underlying, possibly weighted, graph.

Now consider a simple random walk on an unweighted graph. This corresponds to choosing all weights to be 1. The stationary distribution is $\pi(x) = \deg(x)/(2|E(G)|)$. The difference $1 - \lambda_2$ is called the *spectral gap* of the graph. Graphs with larger spectral gaps expand better. The formula for Φ simplifies into

$$\Phi := \min\left\{ \frac{e(S, S^c)}{\sum_{x\in S} \deg(x)} : S \subseteq X, 0 < \sum_{x\in S} \deg(x) \le |E(G)| \right\}, \tag{3}$$

where $e(S, S^c)$ denotes the number of edges between S and S^c.

If G is also d-regular, everything simplifies further. The stationary distribution is simply $\pi(x) = 1/n$. The random walk matrix is simply $\frac{1}{d}A$, where A is the adjacency matrix. The formula for Φ simplifies into

$$\Phi := \min\left\{ \frac{e(S, S^c)}{d|S|} : S \subseteq X, 0 < |S| \le n/2 \right\}. \tag{4}$$

5.1 Eigenvalues and mixing of random walks

(a) [7, Proposition 3] For all $x \in X$ and all positive integer m,

$$\|P^m(x,\cdot) - \pi\| \le \sqrt{\frac{1-\pi(x)}{4\pi(x)}} \max\{|\lambda_2|, |\lambda_{|X|}|\}^m \le \sqrt{\frac{1}{4\pi_{\min}}} \max\{|\lambda_2|, |\lambda_{|X|}|\}^m$$

[21, Theorem 2.12] For all $x \in X$, $A \subseteq X$ and positive integer m,

$$|\Pr(P^m(x,A)) - \pi(A)| \le \sqrt{\frac{\pi(A)}{\pi(x)}} \max\{|\lambda_2|, |\lambda_{|X|}|\}^m$$

In applications, the appearance of the smallest eigenvalue $\lambda_{|X|}$ is usually not important, and what we need to work on is bounding the eigenvalue gap $1 - \lambda_2$. The trick is the following: If the smallest eigenvalue is too small, then we can modify the walk as follows. At each step, we flip a coin and move with probability 1/2 and stay where we are with probability 1/2. The stationary distribution of this modified walk is the same, and the transition matrix is replaced with $\frac{1}{2}(P+I)$. For this modified walk, all eigenvalues are nonnegative, and the eigenvalue gap is half of the original. So applying the theorem to this, we only lose a factor of 2.

(b) For $\alpha > 0$, a *continuous-time Markov chain with rate* α is a Markov chain combined with an exponential clock with parameter α: whenever the clock rings, the walk moves to a random location using the transition matrix. Formally, for a starting vertex x and $t > 0$, the probability that the walk is at vertex y at time t equals

$$P_t(x,y) = e^{-\alpha t}\sum_{n=0}^{\infty}\frac{(\alpha t)^n}{n!}P^n(x,y).$$

[7, Proposition 3] For all $x \in X$ and all positive t,

$$\|P_t(x,\cdot)-\pi\| \leq \sqrt{\frac{1-\pi(x)}{4\pi(x)}}e^{-(1-\lambda_2)\alpha t} \leq \sqrt{\frac{1}{4\pi_{\min}}}e^{-(1-\lambda_2)\alpha t}$$

(c) (**[16] Theorems 3.6, 3.9, and 3.10**) Let G be a d-regular graph on n vertices, and suppose $|\lambda_2|, |\lambda_n| \leq \alpha$, and let $B \subseteq V$ with $|B| = \beta n$. Let $X_0, X_1, \ldots, X_t$ be a random walk on G, where X_0 is chosen uniformly at random. Then we have

$$\mathbb{P}[\forall 0 \leq i \leq t \quad X_i \in B] \leq (\alpha+\beta)^t \,,$$

and for every subset $K \subseteq \{0, \ldots, t\}$,

$$\mathbb{P}[\forall i \in K \quad X_i \in B] \leq (\alpha+\beta)^{|K|-1} \,,$$

and if $\beta > 6\alpha$ then

$$\beta(\beta-2\alpha)^t \leq \mathbb{P}[\forall 0 \leq i \leq t \quad X_i \in B] \leq \beta(\beta+2\alpha)^t \,.$$

5.2 Eigenvalues and expansion

(a) [7, Proposition 6] If the Markov chain is aperiodic then

$$1-2\Phi \leq \lambda_2 \leq 1-\Phi^2 \,.$$

If Markov chain is not aperiodic, one can consider its lazy version, hence [31, Theorem 2]

$$1-2\Phi \leq \lambda_2 \leq 1-\Phi^2/2 \,.$$

(b) Let G be a d-regular graph on n vertices (many of the results below can be extended to general weighted graphs, but assuming regularity makes the formulae cleaner) For $S \subseteq V(G)$ define

$$\phi(S) = \frac{e(S,S^c)}{d|S|} \,.$$

For $k \geq 2$ let

$$\phi_k(G) = \min\{\max\{\phi(S_i) : i = 1, 2, \ldots, k\} : S_1, S_2, \ldots, S_k \text{ partition } V(G)\} .$$

Then, [20, Theorem 3.8] gives

$$(1 - \lambda_k)/2 \leq \phi_k(G) = O\Big(k^4 \sqrt{1 - \lambda_k}\Big) .$$

Also, [18, Theorem 1] gives

$$\phi_2(G) = O\Big(k(1 - \lambda_2)/\sqrt{\lambda_k}\Big) .$$

Results in this item are algorithmic: i.e., the authors also give an algorithm for finding partitions with "small" conductance.

6 Urn theory

6.1 Pólya-Eggenberger urns

Start with W_0 white and B_0 blue balls in an urn. In every step a ball is picked from the urn uniformly at random, the ball is returned to the urn, and s balls of the same colour are added to the urn. Let W_n denote the number of white balls after n draws, and let $\tau_0 = W_0 + B_0$.

(a) (**[26, Proposition 1]**) For $c \geq (W_0 + B_0)/s$ we have

$$\mathbb{P}[W_n = W_0] \leq \left(\frac{c}{c+n}\right)^{W_0/s} .$$

(b) (**[23, Corollary 3.1]**)

$$\mathbb{E}[W_n] = W_0 + \frac{W_0}{\tau_0}\, sn ,$$
$$\mathrm{Var}[W_n] = \frac{W_0 B_0 s^2 n(sn + \tau_0)}{\tau_0^2(\tau_0 + s)} .$$

(c) (**[23, Theorem 3.2]**) For any fixed $x \in [0, 1]$ we have

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{W_n - W_0}{sn} \leq x\right] = \mathbb{P}[\beta(W_0/s, B_0/s) \leq x] = \frac{\Gamma((W_0 + B_0)/s)}{\Gamma(W_0/s)\Gamma(B_0/s)} \int_0^x u^{-1+W_0/s}(1-u)^{-1+B_0/s} \mathrm{d}u .$$

(d) (**[17, page 181]**) Let Z be a beta random variable with parameters W_0/s and B_0/s. Then $(W_n - W_0)/s$, the number of white draws, is distributed as a binomial with parameters n and Z (so it has a mixture distribution). This follows from above result and de Finetti's theorem, since the draws are exchangeable.

7 Singular value decomposition

Let A be an $n \times d$ real matrix of rank r.

(a) There exist $u_1, \dots, u_r \in \mathbf{R}^n$, called the left-singular vectors, and $v_1, \dots, v_r \in \mathbf{R}^d$, called the right-singular vectors, and $\sigma_1, \dots, \sigma_r > 0$ such that

$$A = \sum \sigma_i u_i v_i^T = UDV^T \ ,$$

where $U = [u_1 \dots u_r]$, $V = [v_1 \dots v_r]$, and $D = diag(\sigma_1, \dots, \sigma_r)$. Moreover, U and V are orthogonal: $U^T U = I = V^T V$, and hence $A^{-1} = VD^{-1}U^T$.
For each i, v_i is an eigenvector of $A^T A$ with an eigenvalue of σ_i^2, and u_i is an eigenvector of AA^T with an eigenvalue of σ_i^2. The matrices AA^T and $A^T A$ have eigenvalues $\sigma_1^2, \dots, \sigma_r^2$, plus possibly some zero eigenvalues. Finally, we have

$$\|A\|_F^2 = \sum A_{i,j}^2 = \sum \sigma_i^2 \ .$$

(b) Suppose we arrange the σ_i such that

$$\sigma_1 \geq \dots \geq \sigma_r \ .$$

Then, v_i maximizes $\|Av\|_2$ subject to v having norm 1 and being orthogonal to $v_1, \dots, v_{i-1}$. Similarly, u_i maximizes $\|A^T u\|_2$ subject to u having norm 1 and being orthogonal to $u_1, \dots, u_{i-1}$. Also, we have $Av_i = \sigma_i u_i$ and $A^T u_i = \sigma_i v_i$.
Let $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ for some $k \leq r$. Then, for any $n \times d$ matrix B of rank k, we have

$$\sqrt{\sum_{i=k+1}^{r} \sigma_i^2} = \|A - A_k\|_F \leq \|A - B\|_F,$$
$$\sigma_{k+1} = \|A - A_k\|_2 \leq \|A - B\|_2,$$

where $\|X\|_2$ denotes the operator norm (or spectral norm) of X, e.g. $\|A\|_2 = \sigma_1$.

References

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York; National Bureau of Standards, Washington, DC, 1984. Reprint of the 1972 edition, Selected Government Publications,available at `http://people.math.sfu.ca/~cbm/aands/`.

[2] Noga Alon and Joel H. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[3] M. Bellare and J. Rompel. Randomness-efficient oblivious sampling. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 276–287, Washington, DC, USA, 1994. IEEE Computer Society.

[4] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence.* Oxford University Press, Oxford, 2013.

[5] Fan Chung and Linyuan Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. first 3 chapters are available at `http://www.math.ucsd.edu/~fan/complex/`.

[6] Frank den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs.* American Mathematical Society, Providence, RI, 2000.

[7] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

[8] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press, Cambridge, 2009.

[9] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[10] Tobias Friedrich, Thomas Sauerwald, and Alexandre Stauffer. Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica*, 67(1):65–88, 2013.

[11] Alan Frieze and Michał Karoński. *Introduction to random graphs.* Cambridge University Press, 2015. Book in preparation, available at `http://www.math.cmu.edu/~af1p/Book.html`.

[12] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products.* Academic Press, Inc., San Diego, CA, sixth edition, 2000. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger.

[13] Geoffrey Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 1999.

[14] Geoffrey Grimmett. *Probability on graphs*, volume 1 of *Institute of Mathematical Statistics Textbooks.* Cambridge University Press, Cambridge, 2010. Random processes on graphs and lattices, available at `http://www.statslab.cam.ac.uk/~grg/books/pgs.html`.

[15] Geoffrey Grimmett. Percolation and disordered systems. In *Percolation theory at Saint-Flour*, Probab. St.-Flour, pages 141–303. Springer, Heidelberg, 2012.

[16] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006. available at `http://www.cs.huji.ac.il/~nati/PAPERS/expander_survey.pdf`.

[17] Norman L. Johnson and Samuel Kotz. *Urn models and their application.* John Wiley & Sons, New York-London-Sydney, 1977. An approach to modern discrete probability theory, Wiley Series in Probability and Mathematical Statistics.

[18] Tsz Chiu Kwok, Lap Chi Lau, Yin Tat Lee, Shayan Oveis Gharan, and Luca Trevisan. Improved cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*, STOC '13, pages 11–20, New York, NY, USA, 2013. ACM.

[19] A. Laforgia. Further inequalities for the gamma function. *Mathematics of Computation*, 42(166):597–600, 1984.

[20] James R. Lee, Shayan Oveis Gharan, and Luca Trevisan. Multi-way spectral partitioning and higher-order cheeger inequalities. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 1117–1130, New York, NY, USA, 2012. ACM.

[21] László Lovász. Eigenvalues of graphs. available in `www.cs.elte.hu/~lovasz/eigenvals-x.pdf`, November 2007.

[22] László Lovász and Peter Winkler. Mixing of random walks and other diffusions on a graph. In Peter Rowlinson, editor, *Surveys in Combinatorics, 1995*, pages 119–154. Cambridge University Press, New York, NY, USA, 1995. available in `http://www.cs.elte.hu/~lovasz/morepapers/stirl.pdf`.

[23] Hosam M. Mahmoud. *Pólya urn models.* Texts in Statistical Science Series. CRC Press, Boca Raton, FL, 2009.

[24] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998. available at `http://www.stats.ox.ac.uk/people/academic_staff/colin_mcdiarmid/?a=4139`.

[25] Colin McDiarmid. Centering sequences with bounded differences. *Comb. Probab. Comput.*, 6:79–86, March 1997.

[26] A. Mehrabian and A. Pourmiri. Randomized rumor spreading in poorly connected small-world networks. In F. Kuhn, editor, *Distributed Computing (DISC '14)*, volume 8784 of *Lecture Notes in Computer Science*, pages 346–360. Springer Berlin Heidelberg, 2014.

[27] Abbas Mehrabian and Nick Wormald. It's a small world for random surfers. *arXiv*, 1404.7189v1 [cs.DM], 2014.

[28] Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms*. Cambridge University Press, New York, NY, USA, 1995.

[29] Mathew Penrose. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2003.

[30] Jeanette P. Schmidt, Alan Siegel, and Aravind Srinivasan. Chernoff-Hoeffding bounds for applications with limited independence. *SIAM J. Discrete Math.*, 8(2):223–250, 1995.

[31] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992.

[32] Remco van der Hofstad. Random graphs and complex networks. Lecture notes in preparation, available at `http://www.win.tue.nl/~rhofstad/`, June 2014.

[33] Robert M. Young. 75.9 euler's constant. *The Mathematical Gazette*, 75(472):pp. 187–190, 1991.

Pages below are taken from the homepage of László Kozma.

Useful Inequalities $\{x^2 \geqslant 0\}$ `v0.27b · August 19, 2015`

Cauchy-Schwarz $\left(\sum\limits_{i=1}^{n} x_i y_i\right)^2 \le \left(\sum\limits_{i=1}^{n} x_i^2\right)\left(\sum\limits_{i=1}^{n} y_i^2\right)$

Minkowski $\left(\sum\limits_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \le \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum\limits_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$ for $p \ge 1$.

Hölder $\sum\limits_{i=1}^{n} |x_i y_i| \le \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum\limits_{i=1}^{n} |y_i|^q\right)^{1/q}$ for $p, q > 1,\ \frac{1}{p} + \frac{1}{q} = 1$.

Bernoulli

$(1+x)^r \ge 1 + rx$ for $x \ge -1,\ r \in \mathbb{R} \setminus (0,1)$. Reverse for $r \in [0,1]$.

$(1+x)^r \le 1 + (2^r - 1)x$ for $x \in [0,1],\ r \in \mathbb{R} \setminus (0,1)$.

$(1+x)^n \le \frac{1}{1-nx}$ for $x \in [-1, 0],\ n \in \mathbb{N}$.

$(1+x)^r \le 1 + \frac{rx}{1-(r-1)x}$ for $x \in [-1, \frac{1}{r-1}),\ r > 1$.

$(1+nx)^{n+1} \ge (1+(n+1)x)^n$ for $x \in \mathbb{R},\ n \in \mathbb{N}$.

$(a+b)^n \le a^n + nb(a+b)^{n-1}$ for $a, b \ge 0,\ n \in \mathbb{N}$.

$\left(1 + \frac{x}{p}\right)^p \ge \left(1 + \frac{x}{q}\right)^q$ for (i) $x > 0,\ p > q > 0$,
(ii) $-p < -q < x < 0$, (iii) $-q > -p > x > 0$. Reverse for:
(iv) $q < 0 < p$, $-q > x > 0$, (v) $q < 0 < p$, $-p < x < 0$.

exponential

$e^x \ge \left(1 + \frac{x}{n}\right)^n \ge 1 + x$, $\left(1 + \frac{x}{n}\right)^n \ge e^x\left(1 - \frac{x^2}{n}\right)$ for $n > 1,\ |x| \le n$.

$e^x \ge x^e$ for $x \in \mathbb{R}$, and $\frac{x^n}{n!} + 1 \le e^x \le \left(1 + \frac{x}{n}\right)^{n + x/2}$ for $x, n > 0$.

$e^x \ge 1 + x + \frac{x^2}{2}$ for $x \ge 0$, reverse for $x \le 0$.

$e^{-x} \le 1 - \frac{x}{2}$ for $x \in [0, \sim 1.59]$ and $2^{-x} \le 1 - \frac{x}{2}$ for $x \in [0,1]$.

$\frac{1}{2-x} < x^x < x^2 - x + 1$ for $x \in (0,1)$.

$x^{1/r}(x-1) \le rx(x^{1/r} - 1)$ for $x, r \ge 1$.

$x^y + y^x > 1$ and $e^x > \left(1 + \frac{x}{y}\right)^y > e^{\frac{xy}{x+y}}$ for $x, y > 0$.

$2 - y - e^{-x-y} \le 1 + x \le y + e^{x-y}$, and $e^x \le x + e^{x^2}$ for $x, y \in \mathbb{R}$.

logarithm

$\frac{x-1}{x} \le \ln(x) \le \frac{x^2-1}{2x} \le x - 1$, $\ln(x) \le n(x^{\frac{1}{n}} - 1)$ for $x, n > 0$.

$\frac{2x}{2+x} \le \ln(1+x) \le \frac{x}{\sqrt{x+1}}$ for $x \ge 0$, reverse for $x \in (-1, 0]$.

$\ln(n+1) < \ln(n) + \frac{1}{n} \le \sum_{i=1}^{n} \frac{1}{i} \le \ln(n) + 1$

$\ln(1+x) \ge \frac{x}{2}$ for $x \in [0, \sim 2.51]$, reverse elsewhere.

$\ln(1+x) \ge x - \frac{x^2}{2} + \frac{x^3}{4}$ for $x \in [0, \sim 0.45]$, reverse elsewhere.

$\ln(1-x) \ge -x - \frac{x^2}{2} - \frac{x^3}{2}$ for $x \in [0, \sim 0.43]$, reverse elsewhere.

trigonometric $x - \frac{x^3}{2} \le x \cos x \le \frac{x \cos x}{1 - x^2/3} \le x \sqrt[3]{\cos x} \le x - x^3/6 \le x \cos \frac{x}{\sqrt{3}} \le \sin x$,

hyperbolic $x \cos x \le \frac{x^3}{\sinh^2 x} \le x \cos^2(x/2) \le \sin x \le (x \cos x + 2x)/3 \le \frac{x^2}{\sinh x}$,

$\frac{2}{\pi} x \le \sin x \le x \cos(x/2) \le x \le x + \frac{x^3}{3} \le \tan x$ all for $x \in \left[0, \frac{\pi}{2}\right]$.

$\cosh(x) + \alpha \sinh(x) \le e^{x(\alpha + x/2)}$ for $x \in \mathbb{R},\ \alpha \in [-1, 1]$.

binomial

$\max\left\{\frac{n^k}{k^k}, \frac{(n-k+1)^k}{k!}\right\} \le \binom{n}{k} \le \frac{n^k}{k!} \le \frac{(en)^k}{k^k}$ and $\binom{n}{k} \le \frac{n^n}{k^k (n-k)^{n-k}} \le 2^n$.

$\frac{n^k}{4k!} \le \binom{n}{k}$ for $\sqrt{n} \ge k \ge 0$ and $\frac{4^n}{\sqrt{\pi n}}\left(1 - \frac{1}{8n}\right) \le \binom{2n}{n} \le \frac{4^n}{\sqrt{\pi n}}\left(1 - \frac{1}{9n}\right)$.

$\binom{n_1}{k_1}\binom{n_2}{k_2} \le \binom{n_1 + n_2}{k_1 + k_2}$ for $n_1 \ge k_1 \ge 0,\ n_2 \ge k_2 \ge 0$.

$\frac{\sqrt{\pi}}{2} G \le \binom{n}{\alpha n} \le G$ for $G = \frac{2^{nH(\alpha)}}{\sqrt{2\pi n \alpha(1-\alpha)}}$, $H(x) = -\log_2(x^x (1-x)^{1-x})$.

$\sum_{i=0}^{d} \binom{n}{i} \le n^d + 1$ and $\sum_{i=0}^{d} \binom{n}{i} \le 2^n$ for $n \ge d \ge 0$.

$\sum_{i=0}^{d} \binom{n}{i} \le \left(\frac{en}{d}\right)^d$ for $n \ge d \ge 1$.

$\sum_{i=0}^{d} \binom{n}{i} \le \binom{n}{d}\left(1 + \frac{d}{n - 2d + 1}\right)$ for $\frac{n}{2} \ge d \ge 0$.

$\binom{n}{\alpha n} \le \sum_{i=0}^{\alpha n} \binom{n}{i} \le \frac{1-\alpha}{1-2\alpha}\binom{n}{\alpha n}$ for $\alpha \in (0, \frac{1}{2})$.

square root $2\sqrt{x+1} - 2\sqrt{x} < \frac{1}{\sqrt{x}} < \sqrt{x+1} - \sqrt{x-1} < 2\sqrt{x} - 2\sqrt{x-1}$ for $x \ge 1$.

Stirling $e\left(\frac{n}{e}\right)^n \le \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n+1)} \le n! \le \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n} \le en\left(\frac{n}{e}\right)^n$

means $\min\{x_i\} \le \frac{n}{\sum_i x_i^{-1}} \le \left(\prod_i x_i\right)^{1/n} \le \frac{1}{n}\sum_i x_i \le \sqrt{\frac{1}{n}\sum_i {x_i}^2} \le \max\{x_i\}$

power means $M_p \le M_q$ for $p \le q$, where $M_p = \left(\sum_i w_i |x_i|^p\right)^{1/p}$, $w_i \ge 0$, $\sum_i w_i = 1$.
In the limit $M_0 = \prod_i |x_i|^{w_i}$, $M_{-\infty} = \min_i\{x_i\}$, $M_\infty = \max_i\{x_i\}$.

Lehmer $\frac{\sum_i w_i |x_i|^p}{\sum_i w_i |x_i|^{p-1}} \le \frac{\sum_i w_i |x_i|^q}{\sum_i w_i |x_i|^{q-1}}$ for $p \le q,\ w_i \ge 0$.

log mean $\sqrt{xy} \le \left(\frac{\sqrt{x} + \sqrt{y}}{2}\right)(xy)^{\frac{1}{4}} \le \frac{x-y}{\ln(x) - \ln(y)} \le \left(\frac{\sqrt{x} + \sqrt{y}}{2}\right)^2 \le \frac{x+y}{2}$ for $x, y > 0$.

Heinz $\sqrt{xy} \le \frac{x^{1-\alpha} y^\alpha + x^\alpha y^{1-\alpha}}{2} \le \frac{x+y}{2}$ for $x, y > 0,\ \alpha \in [0,1]$.

Maclaurin-Newton ${S_k}^2 \ge S_{k-1} S_{k+1}$ and $\sqrt[k]{S_k} \ge \sqrt[(k+1)]{S_{k+1}}$ for $1 \le k < n$,
$S_k = \frac{1}{\binom{n}{k}} \sum\limits_{1 \le i_1 < \cdots < i_k \le n} a_{i_1} a_{i_2} \cdots a_{i_k}$, and $a_i \ge 0$.

Jensen $\varphi\left(\sum_i p_i x_i\right) \le \sum_i p_i \varphi(x_i)$ where $p_i \ge 0$, $\sum p_i = 1$, and φ convex.
Alternatively: $\varphi(\mathrm{E}[X]) \le \mathrm{E}[\varphi(X)]$. For concave φ the reverse holds.

Chebyshev $\sum\limits_{i=1}^{n} f(a_i) g(b_i) p_i \ge \left(\sum\limits_{i=1}^{n} f(a_i) p_i\right)\left(\sum\limits_{i=1}^{n} g(b_i) p_i\right) \ge \sum\limits_{i=1}^{n} f(a_i) g(b_{n-i+1}) p_i$
for $a_1 \le \cdots \le a_n$, $b_1 \le \cdots \le b_n$ and f, g nondecreasing, $p_i \ge 0$, $\sum p_i = 1$.
Alternatively: $\mathrm{E}\big[f(X) g(X)\big] \ge \mathrm{E}\big[f(X)\big]\mathrm{E}\big[g(X)\big]$.

rearrangement $\sum\limits_{i=1}^{n} a_i b_i \ge \sum\limits_{i=1}^{n} a_i b_{\pi(i)} \ge \sum\limits_{i=1}^{n} a_i b_{n-i+1}$ for $a_1 \le \cdots \le a_n$,
$b_1 \le \cdots \le b_n$ and π a permutation of $[n]$. More generally:
$\sum\limits_{i=1}^{n} f_i(b_i) \ge \sum\limits_{i=1}^{n} f_i(b_{\pi(i)}) \ge \sum\limits_{i=1}^{n} f_i(b_{n-i+1})$
with $\big(f_{i+1}(x) - f_i(x)\big)$ nondecreasing for all $1 \le i < n$.

Weierstrass $\prod_i (1-x_i)^{w_i} \geq 1 - \sum_i w_i x_i$ where $x_i \leq 1$, and
either $w_i \geq 1$ (for all i) or $w_i \leq 0$ (for all i).
If $w_i \in [0,1]$, $\sum w_i \leq 1$ and $x_i \leq 1$, the reverse holds.

Young $(\frac{1}{px^p} + \frac{1}{qx^q})^{-1} \leq xy \leq \frac{x^p}{p} + \frac{y^q}{q}$ for $x, y \geq 0$, $p, q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$.

Kantorovich $\left(\sum_i {x_i}^2\right)\left(\sum_i {y_i}^2\right) \leq \left(\frac{A}{G}\right)^2 \left(\sum_i x_i y_i\right)^2$ for $x_i, y_i > 0$,
$0 < m \leq \frac{x_i}{y_i} \leq M < \infty$, $A = (m+M)/2$, $G = \sqrt{mM}$.

sum-integral $\int_{L-1}^{U} f(x)\,dx \leq \sum_{i=L}^{U} f(i) \leq \int_{L}^{U+1} f(x)\,dx$ for f nondecreasing.

Cauchy $\varphi'(a) \leq \frac{f(b)-f(a)}{b-a} \leq \varphi'(b)$ where $a < b$, and φ convex.

Hermite $\varphi\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b \varphi(x)\,dx \leq \frac{\varphi(a)+\varphi(b)}{2}$ for φ convex.

Chong $\sum_{i=1}^{n} \frac{a_i}{a_{\pi(i)}} \geq n$ and $\prod_{i=1}^{n} {a_i}^{a_i} \geq \prod_{i=1}^{n} {a_i}^{a_{\pi(i)}}$ for $a_i > 0$.

Gibbs $\sum_i a_i \log \frac{a_i}{b_i} \geq a \log \frac{a}{b}$ for $a_i, b_i \geq 0$, or more generally:
$\sum_i a_i \varphi\left(\frac{b_i}{a_i}\right) \leq a\,\varphi\left(\frac{b}{a}\right)$ for φ concave, and $a := \sum a_i$, $b := \sum b_i$.

Shapiro $\sum_{i=1}^{n} \frac{x_i}{x_{i+1}+x_{i+2}} \geq \frac{n}{2}$ where $x_i > 0$, $(x_{n+1}, x_{n+2}) := (x_1, x_2)$,
and $n \leq 12$ if even, $n \leq 23$ if odd.

Schur $x^t(x-y)(x-z) + y^t(y-z)(y-x) + z^t(z-x)(z-y) \geq 0$
where $x, y, z \geq 0$, $t > 0$

Hadamard $(\det A)^2 \leq \prod_{i=1}^{n} \sum_{j=1}^{n} A_{ij}^2$ where A is an $n \times n$ matrix.

Schur $\sum_{i=1}^{n} \lambda_i^2 \leq \sum_{i,j=1}^{n} A_{ij}^2$ and $\sum_{i=1}^{k} d_i \leq \sum_{i=1}^{k} \lambda_i$ for $1 \leq k \leq n$.
A is an $n \times n$ matrix. For the second inequality A is symmetric.
$\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues, $d_1 \geq \cdots \geq d_n$ the diagonal elements.

Ky Fan $\frac{\prod_{i=1}^{n} x_i^{a_i}}{\prod_{i=1}^{n} (1-x_i)^{a_i}} \leq \frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i (1-x_i)}$ for $x_i \in [0, \frac{1}{2}]$, $a_i \in [0,1]$, $\sum a_i = 1$.

Aczél $\left(a_1 b_1 - \sum_{i=2}^{n} a_i b_i\right)^2 \geq \left(a_1^2 - \sum_{i=2}^{n} a_i^2\right)\left(b_1^2 - \sum_{i=2}^{n} b_i^2\right)$
given that $a_1^2 > \sum_{i=2}^{n} a_i^2$ or $b_1^2 > \sum_{i=2}^{n} b_i^2$.

Mahler $\prod_{i=1}^{n} (x_i + y_i)^{1/n} \geq \prod_{i=1}^{n} x_i^{1/n} + \prod_{i=1}^{n} y_i^{1/n}$ where $x_i, y_i > 0$.

Abel $b_1 \min_k \sum_{i=1}^{k} a_i \leq \sum_{i=1}^{n} a_i b_i \leq b_1 \max_k \sum_{i=1}^{k} a_i$ for $b_1 \geq \cdots \geq b_n \geq 0$.

Milne $\left(\sum_{i=1}^{n} (a_i + b_i)\right)\left(\sum_{i=1}^{n} \frac{a_i b_i}{a_i + b_i}\right) \leq \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right)$

Carleman $\sum_{k=1}^{n} \left(\prod_{i=1}^{k} |a_i|\right)^{1/k} \leq e \sum_{k=1}^{n} |a_k|$

sum & product $\sum_{j=1}^{m} \prod_{i=1}^{n} a_{ij} \geq \sum_{j=1}^{m} \prod_{i=1}^{n} a_{i\pi(j)}$ and $\prod_{j=1}^{m} \sum_{i=1}^{n} a_{ij} \leq \prod_{j=1}^{m} \sum_{i=1}^{n} a_{i\pi(j)}$
where $0 \leq a_{i1} \leq \cdots \leq a_{im}$ for $i = 1, \ldots, n$ and π is a permutation of $[n]$.

$\left|\prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i\right| \leq \sum_{i=1}^{n} |a_i - b_i|$ for $|a_i|, |b_i| \leq 1$.

$\prod_{i=1}^{n} (\alpha + a_i) \geq (1+\alpha)^n$, where $\prod_{i=1}^{n} a_i \geq 1$, $a_i > 0$, $\alpha > 0$.

Callebaut $\left(\sum_i a_i^{1+x} b_i^{1-x}\right)\left(\sum_i a_i^{1-x} b_i^{1+x}\right) \geq \left(\sum_i a_i^{1+y} b_i^{1-y}\right)\left(\sum_i a_i^{1-y} b_i^{1+y}\right)$
for $1 \geq x \geq y \geq 0$, and $i = 1, \ldots, n$.

Karamata $\sum_{i=1}^{n} \varphi(a_i) \geq \sum_{i=1}^{n} \varphi(b_i)$ for $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$,
and $\{a_i\} \succeq \{b_i\}$ (majorization), i.e. $\sum_{i=1}^{t} a_i \geq \sum_{i=1}^{t} b_i$ for all $1 \leq t \leq n$,
with equality for $t = n$ and φ is convex (for concave φ the reverse holds).

Muirhead $\frac{1}{n!} \sum_\pi x_{\pi(1)}^{a_1} \cdots x_{\pi(n)}^{a_n} \geq \frac{1}{n!} \sum_\pi x_{\pi(1)}^{b_1} \cdots x_{\pi(n)}^{b_n}$
where $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ and $\{a_k\} \succeq \{b_k\}$,
$x_i \geq 0$ and the sums extend over all permutations π of $[n]$.

Hilbert $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{a_m b_n}{m+n} \leq \pi \left(\sum_{m=1}^{\infty} a_m^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} b_n^2\right)^{\frac{1}{2}}$ for $a_m, b_n \in \mathbb{R}$.
With $\max\{m, n\}$ instead of $m + n$, we have 4 instead of π.

Hardy $\sum_{n=1}^{\infty} \left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^p \leq \left(\frac{p}{p-1}\right)^p \sum_{n=1}^{\infty} a_n^p$ for $a_n \geq 0$, $p > 1$.

Carlson $\left(\sum_{n=1}^{\infty} a_n\right)^4 \leq \pi^2 \sum_{n=1}^{\infty} a_n^2 \sum_{n=1}^{\infty} n^2 a_n^2$ for $a_n \in \mathbb{R}$.

Mathieu $\frac{1}{c^2 + 1/2} < \sum_{n=1}^{\infty} \frac{2n}{(n^2 + c^2)^2} < \frac{1}{c^2}$ for $c \neq 0$.

Copson $\sum_{n=1}^{\infty} \left(\sum_{k \geq n} \frac{a_k}{k}\right)^p \leq p^p \sum_{n=1}^{\infty} a_n{}^p$ for $a_n \geq 0$, $p > 1$, reverse if $p \in (0, 1)$.

Kraft $\sum 2^{-c(i)} \leq 1$ for $c(i)$ depth of leaf i of binary tree, sum over all leaves.

LYM $\sum_{X \in \mathcal{A}} \binom{n}{|X|}^{-1} \leq 1$, $\mathcal{A} \subset 2^{[n]}$, no set in $\mathcal{A}$ is subset of another set in $\mathcal{A}$.

Sauer-Shelah $|\mathcal{A}| \leq |\mathrm{str}(\mathcal{A})| \leq \sum_{i=0}^{\mathrm{vc}(\mathcal{A})} \binom{n}{i}$ for $\mathcal{A} \subseteq 2^{[n]}$, and
$\mathrm{str}(\mathcal{A}) = \{X \subseteq [n] : X \text{ shattered by } \mathcal{A}\}$, $\mathrm{vc}(\mathcal{A}) = \max\{|X| : X \in \mathrm{str}(\mathcal{A})\}$.

Bonferroni $\Pr\left[\bigvee_{i=1}^{n} A_i\right] \leq \sum_{j=1}^{k} (-1)^{j-1} S_j$ for $1 \leq k \leq n$, k odd,
$\Pr\left[\bigvee_{i=1}^{n} A_i\right] \geq \sum_{j=1}^{k} (-1)^{j-1} S_j$ for $2 \leq k \leq n$, k even.
$S_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \Pr\left[A_{i_1} \wedge \cdots \wedge A_{i_k}\right]$ where A_i are events.

Bhatia-Davis $\mathrm{Var}[X] \leq \big(M - \mathrm{E}[X]\big)\big(\mathrm{E}[X] - m\big)$ where $X \in [m, M]$.

Samuelson $\mu - \sigma\sqrt{n-1} \leq x_i \leq \mu + \sigma\sqrt{n-1}$ for $i = 1, \dots, n$.
Where $\mu = \sum x_i / n$, $\sigma^2 = \sum (x_i - \mu)^2 / n$.

Markov $\Pr\big[|X| \geq a\big] \leq \mathrm{E}\big[|X|\big]/a$ where X is a random variable (r.v.), $a > 0$.
$\Pr\big[X \leq c\big] \leq (1 - \mathrm{E}[X])/(1-c)$ for $X \in [0,1]$ and $c \in \big[0, \mathrm{E}[X]\big]$.
$\Pr\big[X \in S\big] \leq \mathrm{E}[f(X)]/s$ for $f \geq 0$, and $f(x) \geq s > 0$ for all $x \in S$.

Chebyshev $\Pr\big[\big|X - \mathrm{E}[X]\big| \geq t\big] \leq \mathrm{Var}[X]/t^2$ where $t > 0$.
$\Pr\big[X - \mathrm{E}[X] \geq t\big] \leq \mathrm{Var}[X]/(\mathrm{Var}[X] + t^2)$ where $t > 0$.

2^{nd} *moment* $\Pr\big[X > 0\big] \geq (\mathrm{E}[X])^2/(\mathrm{E}[X^2])$ where $\mathrm{E}[X] \geq 0$.
$\Pr\big[X = 0\big] \leq \mathrm{Var}[X]/(\mathrm{E}[X^2])$ where $\mathrm{E}[X^2] \neq 0$.

k^{th} *moment* $\Pr\big[|X - \mu| \geq t\big] \leq \dfrac{\mathrm{E}\left[(X-\mu)^k\right]}{t^k}$ and
$\Pr\big[|X - \mu| \geq t\big] \leq C_k \left(\dfrac{nk}{t^2}\right)^{k/2}$ for $X_i \in [0,1]$ k-wise indep. r.v.,
$X = \sum X_i$, $i = 1, \dots, n$, $\mu = \mathrm{E}[X]$, $C_k = 2\sqrt{\pi k} e^{1/6k} \leq 1.0004$, k even.

Chernoff $\Pr[X \geq t] \leq F(a)/a^t$ for X r.v., $\Pr[X = k] = p_k$,
$F(z) = \sum_k p_k z^k$ probability gen. func., and $a \geq 1$.

$$\Pr\big[X \geq (1+\delta)\mu\big] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \leq \exp\left(\frac{-\mu\delta^2}{3}\right)$$

for X_i i.r.v. from [0,1], $X = \sum X_i$, $\mu = \mathrm{E}[X]$, $\delta \geq 0$ resp. $\delta \in [0,1)$.

$$\Pr\big[X \leq (1-\delta)\mu\big] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \leq \exp\left(\frac{-\mu\delta^2}{2}\right) \text{ for } \delta \in [0,1).$$

Further from the mean: $\Pr\big[X \geq R\big] \leq 2^{-R}$ for $R \geq 2e\mu$ ($\approx 5.44\mu$).

$$\Pr\big[X \geq t\big] \leq \frac{\binom{n}{k} p^k}{\binom{t}{k}} \quad \text{for } X_i \in \{0,1\} \text{ } k\text{-wise i.r.v., } \mathrm{E}[X_i] = p,\ X = \textstyle\sum X_i.$$

$$\Pr\big[X \geq (1+\delta)\mu\big] \leq \binom{n}{\hat{k}} p^{\hat{k}} \Big/ \binom{(1+\delta)\mu}{\hat{k}} \quad \text{for } X_i \in [0,1] \text{ } k\text{-wise i.r.v.,}$$

$k \geq \hat{k} = \lceil \mu\delta/(1-p) \rceil$, $\mathrm{E}[X_i] = p_i$, $X = \sum X_i$, $\mu = \mathrm{E}[X]$, $p = \frac{\mu}{n}$, $\delta > 0$.

Hoeffding $\Pr\big[\big|X - \mathrm{E}[X]\big| \geq \delta\big] \leq 2\exp\left(\dfrac{-2\delta^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$ for X_i i.r.v.,
$X_i \in [a_i, b_i]$ (w. prob. 1), $X = \sum X_i$, $\delta \geq 0$.
A related lemma, assuming $\mathrm{E}[X] = 0$, $X \in [a,b]$ (w. prob. 1) and $\lambda \in \mathbb{R}$:

$$\mathrm{E}\big[e^{\lambda X}\big] \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

Kolmogorov $\Pr\big[\max_k |S_k| \geq \varepsilon\big] \leq \frac{1}{\varepsilon^2}\mathrm{Var}[S_n] = \frac{1}{\varepsilon^2}\sum_i \mathrm{Var}[X_i]$
where $X_1, \dots, X_n$ are i.r.v., $\mathrm{E}[X_i] = 0$,
$\mathrm{Var}[X_i] < \infty$ for all i, $S_k = \sum_{i=1}^k X_i$ and $\varepsilon > 0$.

Paley-Zygmund $\Pr\big[X \geq \mu\,\mathrm{E}[X]\,\big] \geq 1 - \dfrac{\mathrm{Var}[X]}{(1-\mu)^2\,(\mathrm{E}[X])^2 + \mathrm{Var}[X]}$ for $X \geq 0$,
$\mathrm{Var}[X] < \infty$, and $\mu \in (0,1)$.

Vysochanskij-Petunin-Gauss $\Pr\big[\big|X - \mathrm{E}[X]\big| \geq \lambda\sigma\big] \leq \frac{4}{9\lambda^2}$ if $\lambda \geq \sqrt{\frac{8}{3}}$,
$\Pr\big[|X - m| \geq \varepsilon\big] \leq \frac{4\tau^2}{9\varepsilon^2}$ if $\varepsilon \geq \frac{2\tau}{\sqrt{3}}$,
$\Pr\big[|X - m| \geq \varepsilon\big] \leq 1 - \frac{\varepsilon}{\sqrt{3}\tau}$ if $\varepsilon \leq \frac{2\tau}{\sqrt{3}}$.
Where X is a unimodal r.v. with mode m,
$\sigma^2 = \mathrm{Var}[X] < \infty$, $\tau^2 = \mathrm{Var}[X] + (\mathrm{E}[X] - m)^2 = \mathrm{E}[(X-m)^2]$.

Etemadi $\Pr\big[\max_{1 \leq k \leq n} |S_k| \geq 3\alpha\big] \leq 3 \max_{1 \leq k \leq n}\big(\Pr\big[\,|S_k| \geq \alpha\big]\big)$
where X_i are i.r.v., $S_k = \sum_{i=1}^k X_i$, $\alpha \geq 0$.

Doob $\Pr\big[\max_{1 \leq k \leq n} |X_k| \geq \varepsilon\big] \leq \mathrm{E}\big[|X_n|\big]/\varepsilon$ for martingale (X_k) and $\varepsilon > 0$.

Bennett $\Pr\Big[\sum_{i=1}^n X_i \geq \varepsilon\Big] \leq \exp\left(-\dfrac{n\sigma^2}{M^2}\,\theta\Big(\dfrac{M\varepsilon}{n\sigma^2}\Big)\right)$ where X_i i.r.v.,
$\mathrm{E}[X_i] = 0$, $\sigma^2 = \frac{1}{n}\sum \mathrm{Var}[X_i]$, $|X_i| \leq M$ (w. prob. 1), $\varepsilon \geq 0$,
$\theta(u) = (1+u)\log(1+u) - u$.

Bernstein $\Pr\Big[\sum_{i=1}^n X_i \geq \varepsilon\Big] \leq \exp\left(\dfrac{-\varepsilon^2}{2(n\sigma^2 + M\varepsilon/3)}\right)$ for X_i i.r.v.,
$\mathrm{E}[X_i] = 0$, $|X_i| < M$ (w. prob. 1) for all i, $\sigma^2 = \frac{1}{n}\sum \mathrm{Var}[X_i]$, $\varepsilon \geq 0$.

Azuma $\Pr\big[|X_n - X_0| \geq \delta\big] \leq 2\exp\left(\dfrac{-\delta^2}{2\sum_{i=1}^n {c_i}^2}\right)$ for martingale (X_k) s.t.
$\big|X_i - X_{i-1}\big| < c_i$ (w. prob. 1), for $i = 1, \dots, n$, $\delta \geq 0$.

Efron-Stein $\mathrm{Var}[Z] \leq \frac{1}{2}\,\mathrm{E}\left[\sum_{i=1}^n \big(Z - Z^{(i)}\big)^2\right]$ for $X_i, X_i{}' \in \mathcal{X}$ i.r.v.,
$f : \mathcal{X}^n \to \mathbb{R}$, $Z = f(X_1, \dots, X_n)$, $Z^{(i)} = f(X_1, \dots, X_i{}', \dots, X_n)$.

McDiarmid $\Pr\big[\big|Z - \mathrm{E}[Z]\big| \geq \delta\big] \leq 2\exp\left(\dfrac{-2\delta^2}{\sum_{i=1}^n c_i^2}\right)$ for $X_i, X_i{}' \in \mathcal{X}$ i.r.v.,
Z, $Z^{(i)}$ as before, s.t. $\big|Z - Z^{(i)}\big| \leq c_i$ for all i, and $\delta \geq 0$.

Janson $M \leq \Pr\big[\bigwedge \overline{B}_i\big] \leq M\exp\left(\dfrac{\Delta}{2 - 2\varepsilon}\right)$ where $\Pr[B_i] \leq \varepsilon$ for all i,
$M = \prod(1 - \Pr[B_i])$, $\Delta = \sum_{i \neq j, B_i \sim B_j} \Pr[B_i \wedge B_j]$.

Lovász $\Pr\big[\bigwedge \overline{B}_i\big] \geq \prod(1 - x_i) > 0$ where $\Pr[B_i] \leq x_i \cdot \prod_{(i,j) \in D}(1 - x_j)$,
for $x_i \in [0,1)$ for all $i = 1, \dots, n$ and D the dependency graph.
If each B_i mutually indep. of the set of all other events, exc. at most d,
$\Pr[B_i] \leq p$ for all $i = 1, \dots, n$, then if $ep(d+1) \leq 1$ then $\Pr\big[\bigwedge \overline{B}_i\big] > 0$.