1. Probability (Technical note 11.9 v0.6)

1A. Concentration inequalities and tail bounds

Unless otherwise specified, all variables are real $\mathbb{R}$. Inequalities come as one-sided $\Pr(\cdots \leq \cdots)$ and two-sided $\Pr(|\cdots| \leq \cdots)$. Notation: X is a random variable, $\mu := \mathbb{E}[X]$, $\sigma^2 := \mathrm{Var}[X]$, $S_n := X_1 + \cdots + X_n$.

Inequality	Conditions		Common form	Notes / Alternate form
			Single random variable	
Markov[1]	Non-negative	$X \geq 0$	$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$ $\quad\forall\, a > 0$	$\Pr[X \geq kE[X]] \leq \frac{1}{k} \quad k > 1$ [3, Sec. 5.1][6, Thm 1.13]
extension	+ non-negative, strictly increasing func Φ	$X \geq 0$ $\Phi(X) \geq 0$ increasing	$\Pr[X \geq a] = \Pr[\Phi(X) \geq \Phi(a)] \leq \dfrac{\mathrm{E}(\Phi(X))}{\Phi(a)}$	$\forall\, a > 0$ Wiki
Reverse Markov	upper-bounded by U (can be positive)	$\max X = U$	$\Pr[X \leq a] \leq \frac{U - \mathbb{E}[X]}{U-a}$	$\forall\, a > 0$ [1, Sec. 3.1]
Chebyshev[2]	Finite mean and variance	$\mathbb{E}[X], \mathrm{Var}[X]$ finite	$\Pr[\lvert X - \mathbb{E}[X]\rvert \geq a] \leq \frac{\sigma^2}{a^2}$	$\Pr[\lvert X - \mathbb{E}[X]\rvert \geq a \cdot \sigma] \leq \frac{1}{a^2}$ [1, Sec. 3.2] $\forall\, a > 0,\ \sigma^2 = \mathrm{Var}[X]$ [3, Sec. 5.1][2, Thm 18.11]
Cantelli	Improved Chebyshev	(same; but one-sided)	$\Pr[X - \mathbb{E}[X] \geq a]) \leq \dfrac{\sigma^2}{\sigma^2 + a^2}$	$\forall\, a > 0,\ \sigma^2 = \mathrm{Var}[X]$ Wiki
Chernoff[3]	Generic		$\Pr[X \geq a] = \Pr\left[e^{tX} \geq e^{ta}\right]$	$\forall\, t > 0, \quad a \in \mathbb{R}$ [1, Sec. 3.3]
Jensen		$f : \mathbb{R} \to \mathbb{R}$; f convex	$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$	[3, Prob. 5.3][6, Thm 1.14]
Hoeffding's lemma		$\mathbb{E}[X] = \mu$ $a \leq X \leq b$	$\mathbb{E}\left[e^{\lambda X}\right] \leq e^{\lambda\mu} e^{\frac{\lambda^2 (b-a)^2}{8}}$	$\lambda \in \mathbb{R}$ [1, Sec. 3.4]
			Sum of random variables	
Chernoff-Hoeffding (one-sided)	n independent random vars	$X_1, \ldots, X_n$ indep $S_n = X_1 + \cdots + X_n$ $X_i \in [a_i, b_i] \quad \forall i$	$\Pr[S_n - \mathbb{E}[S_n] \geq t] \leq \exp\left(\frac{-2t^2 n^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$	[1, Sec. 3.5]
(two-sided)[4]	(same as above)		$\Pr[\lvert S_n - \mathbb{E}[S_n]\rvert > t] \leq 2\exp\left(\frac{-2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$	$\forall t \in \left(0, \frac{1}{2}\right)$ [5, Thm.1.1]
(two-sided iid)	same plus iid, range, mean μ for each	$X_1, \ldots, X_n \in [0,1]$ $\mathbb{E}[X_i] = \mu$ iid	$\Pr\left[\left\lvert \frac{S_n}{n} - \mu \right\rvert \geq \epsilon\right] \leq 2\exp\left(-2n\epsilon^2\right)$	$\forall\, \epsilon > 0$ [6, Thm 1.16]
Thm 1.3	n independent random vars	$X_1, \ldots, X_n$ indep $S_n = X_1 + \cdots + X_n$	$\Pr[S_n - \mathbb{E}[S_n] > \epsilon] \leq 2\exp\left(\frac{-\epsilon^2}{4\sum_{i=1}^n \mathrm{Var}[X_i]}\right)$	$\epsilon \in (0, 2\,\mathrm{Var}[S_n] / (\max_i \lvert X_i - \mathbb{E}[X_i]\rvert))$ [5, Thm. 1.3]
Azuma				
Weak law of large numbers	n independent iid random vars	$X_1, \ldots, X_n$ indep $\mathbb{E}[X_i] = \mu$ iid	$\lim_{n\to\infty} \Pr\left[\lvert \frac{1}{n} S_n - \mu \rvert \geq \epsilon\right] = 0$	$\forall\, \epsilon > 0$ [3, Sec. 5.2][6, Thm 1.15]
Strong law of large numbers	(same)	(same)	$\Pr\left[\lim_{n\to\infty} \frac{1}{n} S_n = \mu\right] = 1$	[3, Sec. 5.5]
	Advanced			
Bennett	n independent zero-mean	$X_1, \ldots, X_n$ indep $\mathbb{E}[X_i] = 0$ iid	$\Pr[S_n > \epsilon] \leq \exp\left(-n\sigma^2 h\left(\frac{\epsilon}{n\sigma^2}\right)\right)$	$\sigma^2 := \frac{1}{n}\sum_{i=1}^n \mathrm{Var}[X_i]$, $\forall\, \epsilon > 0$, $h(a) := (1+a)\log(1+a) - a$ for $a \geq 0$ [1, 4.1]
Bernstein	(same)	(same)	$\Pr[S_n > \epsilon] \leq \exp\left(\frac{-n\epsilon^2}{2(\sigma^2 + \epsilon/3)}\right)$	(same) [1, 4.2]
Efron-Stein	scalar func of vars $f\colon \chi^n \to \mathbb{R}$	$X_1, \ldots, X_n$ indep w/ values in set χ	$\mathrm{Var}[Z] \leq \sum_{i=1}^n \mathbb{E}\left[(Z - \mathbb{E}_i[Z])^2\right]$	$Z := g(X_1, \ldots, X_n)$ [1, 4.3] $\mathbb{E}_i[Z] := \mathbb{E}[Z \vert X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_n]$
McDiarmid's	scalar func of vars $f\colon \chi^n \to \mathbb{R}$	$X_1, \ldots, X_n$ indep w/ values in set χ	$\Pr[f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)] \geq \epsilon] \leq \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^n c_i^2)}\right)$	condition: c-bounded difference property $\forall\, \epsilon > 0$ [1, 4.4] $\left\lvert f(X_1, \ldots, X_i, \ldots, X_n) - f\left(X_1, \ldots, X_i', \ldots, X_n\right)\right\rvert \leq c_i$

[1] Markov's inequality bounds the first moment of random variable. Use it when a constant probability bound is sufficient [1, Sec. 3.3].

[2] Chebyshev is derived from Markov. It bounds the second moment. It is the appropriate one when the variance σ is known. If σ is unknown, we can use the bounds of $X \in [a, b]$.

[3] Chernoff bound is used to bound the tails of the distribution for a sum of independent random variables. By far the most useful tool in randomized algorithms [1, Sec. 3.3].

[4] This probability can be interpreted as the level of significance ϵ (probability of making an error) for a confidence interval around the mean of size 2ϵ. Therefore, we require at least $\log(2\alpha)/2t^2$ samples to acquire $1 - \alpha$ confidence interval $\mathbb{E}[\bar{X}] \pm t$.

Other expressions

- ➡ Union bound [5, Thm. (4.1)]
- ➡ Schwarz inequality $(\mathbb{E}[XY])^2 \leq \mathbb{E}[X]^2\, \mathbb{E}[Y]^2$[3, Ch. 5]
- ➡ Exponential inequalities. $\frac{1}{2}\left(e^x + e^{-x}\right) \leq e^{x^2/2}$ [5, Eq. (4.1)]
- ➡ Core references: [3, Ch. 5] (core classical theory), [6, Ch. 1] (quantum info essentials, formal), [4, App A], [5], [1], [2, Ch. 18]; see summary on wiki.

Acknowledgments: Many thanks to John Watrous, who was kind to provide me with some useful references on the subject. I have included these in the bibliography.

Bibliography

[1] Kumar Abhishek, Sneha Maheshwari, and Sujit Gujar. Introduction to Concentration Inequalities. oct 2019.

[2] Boaz Barak. Introduction to Theoretical Computer Science.

[3] D Bertsekas and J N Tsitsiklis. *Introduction to Probability*. Athena Scientific optimization and computation series. Athena Scientific, 2008.

[4] P Kaye, P.K.R.L.M. Mosca, I.Q.C.P. Kaye, R Laflamme, and M Mosca. *An Introduction to Quantum Computing*. OUP Oxford, 2007.

[5] Jeff M. Phillips. Chernoff-Hoeffding Inequality and Applications. sep 2012.

[6] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, apr 2018.

Caveat emptor These pages are a work in progress, inevitably imperfect, incomplete, and surely enriched with typos and unannounced inaccuracies. Sources credited in Bibliography to the best of my ability, though certain omissions certainly remain.