MS-E1603 - Useful formulas

Inequalities

Markov	$\mathbb{P}(\lvert X\rvert \geq a) \leq \frac{\mathbb{E}(\lvert X\rvert)}{a}$, $a \geq 0$
Chebyshev	$\mathbb{P}(\lvert X-\mathbb{E}(X)\rvert \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}$
Cauchy-Schwarz	$\mathbb{E}(XY) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$
	$\left(\sum_{i=1}^n x_i y_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)$
Hoeffding	$\mathbb{P}(\lvert S_n-\mathbb{E}(S_n)\rvert \geq a) \leq 2e^{-\frac{2a^2}{\sum_i (b_i-a_i)^2}}$,
	where $S_n=\sum_{i=1}^n X_i$ with $a_i \leq X_i \leq b_i$, $\perp\!\!\!\perp$
Jensen	$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))$ for convex ϕ
Chernoff bound	$\mathbb{P}(X \geq a) \leq \inf_{t>0} e^{-ta}\mathbb{E}(e^{tX})$,
Union bound	$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i)$,
	where $n \in \mathbb{N}$ or $n=\infty$

Statistics and combinatorics

Binomial coefficient	$\binom{n}{k} = \frac{n!}{(n-k)!k!}$
	$\frac{n^k}{k^k} \leq \binom{n}{k} \leq \frac{n^k}{k!}$
Binomial theorem	$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k}y^k$
Inclusion-exclusion	$\mathbb{P}(\bigcup_i^n A_i) = \sum_i^n \mathbb{P}(A_i) - \sum_{i<j}\mathbb{P}(A_i \cap A_j) + \ldots$
Order statistics	$\mathbb{P}(X_{(k)} \leq a) = \sum_{i=0}^{n-k}\binom{n}{i}F(a)^{n-i}(1-F(a))^i$
	where $X_1,\ldots,X_n$ i.i.d. with cdf $F(x)$
Law of total probability	$\mathbb{P}(A) = \mathbb{E}[\mathbb{P}(A\vert X)]$
	esp. $\mathbb{P}(A) = \sum_i \mathbb{P}(A\vert X=x_i)\mathbb{P}(X=x_i)$
Law of total expectation	$\mathbb{E}(X) = \mathbb{E}[\mathbb{E}[X\vert Y]]$
Law of total variance	$\mathrm{Var}(X) = \mathrm{Var}(\mathbb{E}[X\vert Y]) + \mathbb{E}[\mathrm{Var}(X\vert Y)]$

Convergence of random variables

In probability	$X_n \xrightarrow{p} X \iff \mathbb{P}(\lvert X_n - X\rvert \geq a) \to 0$,
	for all $a>0$
Weakly/in distribution	$X_n \xrightarrow{d} X \iff \mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$
	for all bounded and continuous f
Almost surely*	$X_n \xrightarrow[n\to\infty]{a.s.} X \iff \mathbb{P}(X_n \to X) = 1$
Weak/strong laws of large numbers	(weak) $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{p} \mathbb{E}(X)$
	(strong) $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{a.s.} \mathbb{E}(X)$
	where $X, X_1, X_2, \ldots$ i.i.d.
Central limit theorem	$\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n (X_i - \mathbb{E}(X)) \xrightarrow{d} \mathcal{N}(0,1)$
	where $X, X_1, X_2, \ldots$ i.i.d. and $\sigma^2 = \mathrm{Var}(X) < \infty$

Miscellaneous

"Little-o" notation	$f(x) = o(g(x)) \iff f(x)/g(x) \to 0$
"Big-o" notation	$f(x) = O(g(x)) \iff \limsup \lvert f(x)/g(x)\rvert < \infty$
"Big-theta" notation	$f(x) = \Theta(g(x)) \iff f(x) = O(g(x))$
	and $g(x) = O(f(x))$
"A with high probability/w.h.p"	$\Pr(A) \to 1$
Taylor's approximation	$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(k)}(a)}{k!}(x-a)^k$
	$+ o(\lvert x-a\rvert^k)$
Stirling's approximation	$n! = \Gamma(n+1) = \sqrt{2\pi n}e^{-n}n^n(1+O(n^{-1}))$
	$\ln n! = n\ln n - n + O(\ln n)$
Dominated convergence*	$X_n \to X$, $\lvert X_n\rvert \leq Y$ and $\mathbb{E}(Y) < \infty$
	$\Rightarrow \mathbb{E}(X_n) \to \mathbb{E}(X)$
Monotone convergence*	$X_n \to X$ and $0 \leq X_1 \leq X_2 \leq \ldots$
	$\Rightarrow \mathbb{E}(X_n) \to \mathbb{E}(X)$

* see e.g. MS-E1600 Probability theory for details