Daniel Golovin Fall 2008
Notes: Large Deviation Inequalities

Large Deviation Bounds

- **Markov:** For a positive random variable X, and any $a \geq 1$, $\mathbf{Pr}[X > a\mathbf{E}[X]] \leq 1/a$
- **Chebyshev:** For a positive random variable X with standard deviation σ, and any $k \geq 1$, $\mathbf{Pr}[|X - \mathbf{E}[X]| \geq k\sigma] \leq 1/k^2$.
- **Simple Version of Chernoff:** $\mathbf{Pr}[X \geq t\mathbf{E}[X]] = \mathbf{Pr}\left[e^X \geq e^{t\mathbf{E}[X]}\right] \leq \mathbf{E}\left[e^X\right]/e^{t\mathbf{E}[X]}$
- **Chernoff:** (Theorem 5 from Chung & Lu's survey on concentration inequalities [1])
  Let $\{X_i \mid 1 \leq i \leq n\}$ be independent *binary* random variables with $\mathbf{Pr}[X_i = 1] = p_i$ and $\mathbf{Pr}[X_i = 0] = 1 - p_i$. Fix any positive $a_1, \ldots, a_n$ and let $Y := \sum_i a_i X_i$, let $\mu := \mathbf{E}[Y] = \sum_i a_i p_i$, let $\beta := \sum_i a_i^2 p_i$, and let $a := \max_i\{a_i\}$. Then
  $$\mathbf{Pr}[Y \leq \mu - \lambda] \leq \exp\left(-\lambda^2/2\beta\right) \quad (1)$$
  $$\mathbf{Pr}[Y \geq \mu + \lambda] \leq \exp\left(-\lambda^2/2(\beta + a\lambda/3)\right) \quad (2)$$
- **Azuma-Hoeffding:** Let X be a random variable determined by n trials $\{T_i | i \in [n]\}$, and satisfying for each i
  $$\max_v |\mathbf{E}[X \mid \forall j \in [1, i+1] : T_j = v_j] - \mathbf{E}[X \mid \forall j \in [1, i] : T_j = v_j]| \leq c_i$$
  then $\mathbf{Pr}[|X - \mathbf{E}[X]| > t] \leq 2\exp\{-t^2/2\sum_i c_i^2\}$.
- **Talagrand:** This is actually a useful corollary of Talagrand's inequality, from "The Probabilistic Method" by Noga Alon and Joel Spencer (2nd edition).
  Let $\Omega = \prod_{i=1}^n \Omega_i$ where Ω_i is a probability space and Ω has the product measure. Fix $h : \Omega \to \mathbb{R}$, such that h is 1-*Lipschitz* (that is, $|h(x) - h(y)| \leq 1$ if x and y differ in at most one coordinate) and f-*certifiable* for some $f : \mathbb{N} \to \mathbb{N}$ (that is, if $h(x) \geq s$ then there exists $I \subset [n]$ with $|I| \leq f(s)$ so that all $y \in \Omega$ that agree on x on coordinates of I have $h(y) \geq s$). Then for all b and t
  $$\mathbf{Pr}\left[h \leq b - t\sqrt{f(b)}\right] \cdot \mathbf{Pr}[X \geq b] \leq \exp\{-t^2/4\}$$
- **Efron-Stein:** Let $X_1, \ldots, X_n$ be independent random variables taking values in S (they may be distributed differently), and fix any $g : S^n \to \mathbb{R}$. Define random variable $Z = g(X_1, X_2, \ldots, X_n)$, and let $\mu_i := \mathbf{E}[Z \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n]$ for all $i \in [n]$. Then
  $$\mathbf{Var}[Z] \leq \sum_i \mathbf{E}\left[\left(Z - \mu_i\right)^2\right]$$
  As a corollary, if g is $(c_1, \ldots, c_n)$-Lipschitz, then $\mathbf{Var}[Z] \leq \frac{1}{2}\sum_i c_i^2$.

References

[1] Fan Chung and Linyuan Lu. Concentration inequalities and martingale inequalities – a survey. Internet Math., to appear.