Bennett vs. Bernstein

Behrad Moniri
University of Pennsylvania
bemoniri@seas.upenn.edu

Abstract

In this note, we will compare the Bennett's and Bernstein's tail bound for bounded variables.

Let X_i be a mean-zero random variables such that $|X_i| < b$. First, we bound the moment generating function from above. For every $\lambda \in \mathbb{R}$ we can write

$$\begin{aligned}
\mathbb{E}(e^{\lambda X_i}) &= 1 + \mathbb{E}\Big(\sum_{k=2}^{\infty} \frac{\lambda^k X_i^k}{k!}\Big) \\
&= 1 + \sum_{k=2}^{\infty} \frac{\lambda^k \mathbb{E}(X_i^2 X_i^{k-2})}{k!} \\
&\leq 1 + \sum_{k=2}^{\infty} \frac{\lambda^k \sigma_i^2 b^{k-2}}{k!} \\
&= 1 + \frac{\sigma_i^2}{b^2} \sum_{k=2}^{\infty} \frac{\lambda^k b^k}{k!} \\
&= 1 + \frac{\sigma_i^2}{b^2}(e^{\lambda b} - \lambda b - 1) \\
&\leq \exp\Big(\frac{\sigma_i^2}{b^2}(e^{\lambda b} - \lambda b - 1)\Big).
\end{aligned}$$

Hence

$$\log \mathbb{E}(e^{\lambda X_i}) \leq \frac{\sigma_i^2}{b^2}(e^{\lambda b} - \lambda b - 1). \tag{1}$$

Since X_is are independent random variables, we can write

$$\mathbb{P}\Big(\sum_{i=1}^{n} X_i \geq t\Big) \leq \frac{\mathbb{E}\Big(\exp(\sum_{i=1}^{n} \lambda X_i)\Big)}{e^{\lambda t}} = \frac{\prod_{i=1}^{n} \mathbb{E}\Big(\exp(\lambda X_i)\Big)}{e^{\lambda t}} \tag{2}$$

Substituting (1) in (2) we arrive at the following bound:

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=1}^{n} X_i \geq t\Big) &\leq e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}(\exp(\lambda X_i)) \\
&\leq e^{-\lambda t} \prod_{i=1}^{n} \exp\Big(\frac{\sigma_i^2}{b^2}(e^{\lambda b} - 1 - \lambda b)\Big).
\end{aligned}$$

Optimizing λ yield the Bennett's inequality

$$\mathbb{P}\Big(\sum_{i=1}^{n} X_i \geq t\Big) \leq \exp\Big(-\frac{n\sigma^2}{b^2} h\Big(\frac{bt}{n\sigma^2}\Big)\Big), \tag{3}$$

where $h(t) = (1+t)\log(1+t)$ for $t > 0$. Bernstein's inequality is a weakening of Bennett. Let $\Phi(x) = \frac{x^2}{2+\frac{2}{3}x}$. It is easy to see that

$$\forall x \geq 0 \quad (1+x)\log(1+x) - \frac{x^2}{2+2/3x} \geq 0. \tag{4}$$

Substituting h with Φ in Bennett yield the Bernstein's inequality.