Chernoff Bound

Kirill Levchenko

We will prove a fairly general form of the Chernoff bound. This proof is due to Van Vu, who gave it at a lecture at the University of California, San Diego.

Theorem 1. *Let* $X_1, \ldots, X_n$ *be discrete, independent random variables such that* $\mathrm{E}[X_i] = 0$ *and* $|X_i| \le 1$ *for all* i*. Let* $X = \sum_{i=1}^{n} X_i$ *and* σ^2 *be the variance of* X*. Then*

$$\Pr[|X| \ge \lambda\sigma] \le 2e^{-\lambda^2/4}$$

for any $0 \le \lambda \le 2\sigma$*.*

Proof. We will prove

$$\Pr[X \ge \lambda\sigma] \le e^{-\lambda^2/4},$$

which will suffice, by the symmetry of argument. Let $0 \le t \le 1$, to be determined later. Note that

$$\begin{aligned} \Pr[X \ge \lambda\sigma] &= \Pr[tX \ge t\lambda\sigma] \\ &= \Pr[e^{tX} \ge e^{t\lambda\sigma}] \\ &\le \frac{\mathrm{E}[e^{tX}]}{e^{t\lambda\sigma}}, \end{aligned}$$

the last by the Markov Inequality.

Before going any further, we establish a bound on $\mathrm{E}[e^{tZ}]$, where $-1 \le Z \le 1$ and $\mathrm{E}[Z] = 0$. Additionally, let $t \le 1$. By the definition of expectation,

$$\begin{aligned} \mathrm{E}[e^{tZ}] &= \sum_{j=1}^{m} p_j e^{tz_j} \\ &= \sum_{j=1}^{m} p_j \big(1 + tz_j + \frac{1}{2!}(tz_j)^2 + \frac{1}{3!}(tz_j)^3 + \cdots\big) \\ &= \underbrace{\sum_{j=1}^{m} p_j}_{A} + t\underbrace{\sum_{j=1}^{m} p_j z_j}_{B} + \underbrace{\sum_{j=1}^{m} p_j \big(\frac{1}{2!}(tz_j)^2 + \frac{1}{3!}(tz_j)^3 + \cdots\big)}_{C}. \end{aligned}$$

Summation A is the sum of all probabilities, so $A = 1$. Summation B is the expectation of Z, so $B = 0$. Since $|ta_j| \le 1$, we can upper-bound C by

$$\sum_{j=1}^{m} p_j (ta_j)^2 \big(\frac{1}{2!} + \frac{1}{3!} + \cdots\big) \le t^2 \sum_{j=1}^{m} p_j a_j^2.$$

But the above summation is just $\mathrm{Var}[Z]$, giving

$$\mathrm{E}[e^{tZ}] \le 1 + t^2 \mathrm{Var}[Z].$$

Returning to our main computation,

$$
\begin{aligned}
\mathrm{E}[e^{tX}] &= \mathrm{E}[e^{t(X_1+X_2+\cdots+X_n)}] \\
&= \mathrm{E}\big[\textstyle\prod_{i=1}^{n} e^{tX_i}\big] \\
&= \textstyle\prod_{i=1}^{n} \mathrm{E}[e^{tX_i}] && \text{by the independence of } X_i \\
&\leq \textstyle\prod_{i=1}^{n} \big(1 + t^2\mathrm{Var}[X_i]\big) \\
&\leq \textstyle\prod_{i=1}^{n} e^{t^2\mathrm{Var}[X_i]} && \text{since } 1+\alpha \leq e^{\alpha} \text{ for } \alpha \geq 0 \\
&= e^{t^2\sigma^2}. && \text{by the independence of } X_i
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\Pr[X \geq \lambda\sigma] &\leq \frac{e^{t^2\sigma^2}}{e^{t\lambda\sigma}} \\
&= e^{t\sigma(t\sigma-\lambda)}.
\end{aligned}
$$

Optimizing t we get $t = \lambda/2\sigma$, which gives

$$\Pr[X \geq \lambda\sigma] \leq e^{-\lambda^2/4}.$$

$\square$