1 Concentration Inequalities

Suppose $X$ is a random variable and $m$ is a constant (usually the mean or median of $X$). We seek bounds like

$$P(X - m > t) \leq \exp(-f(t)), \quad P(X - m < -t) \leq \exp(-g(t)).$$

Typically, $f(0) = 0$ and $\lim_{t \to \infty} f(t) = \infty$, and similarly for $g$. For example, if $X_i$ are iid with $P(X_i = -1) = P(X_i = 1) = 0.5$, then

$$P\left(\frac{1}{n} \sum_{i=1}^{n} X_i > t\right) \leq \exp\left(-\frac{nt^2}{2}\right),$$

and similarly for the lower bound, so

$$P\left(\left|\frac{1}{n} \sum_{i=1}^{n} X_i\right| > t\right) \leq 2 \exp\left(-\frac{nt^2}{2}\right).$$

Obviously this provides some justification for the statement that $\frac{1}{n} \sum_{i=1}^{n} X_i$ is concentrated near 0.

Now the issue is to read off the “typical deviation” of $X$ from $m$. The method is to find the range of $t$ for which $f$ is “like a constant,” say, e.g., equal to 1. In the example above, $t = n^{-1/2}$ is the typical deviation. If $t \ll n^{-1/2}$ then $nt^2/2$ is near 0, so the bound provides no information. If $t \gg n^{-1/2}$, then $nt^2/2$ is quite large, so the upper bound on $P(X - m > t)$ is near 0, and we are left wondering whether there is a smaller neighborhood of $m$ around which $X$ concentrates.

Caution: Don’t assume that upper bounds are sharp.

The simplest concentration inequalities come from variance bounds: the typical deviation of $X$ from $EX$ is $\sqrt{\text{Var}(X)}$. The following is a useful bound on the variance of a function of several random variables.

**Theorem 1** Efron Stein-Inequality (or Influence Inequality, or MG bound on Variance). Suppose that $X_1, ..., X_n$, $X'_1, ..., X'_n$ are independent with $X'_i \overset{d}{=} X_i$ for all $i$. Let $X = \frac{1}{n} \sum_{i=1}^{n} X'_i$.
Let $X' = (X'_1, \ldots, X'_n)$, $X^{[i]} = (X'_1, \ldots, X'_i, X_{i+1}, \ldots, X_n)$; note $X^{[0]} = X$ and $X^{[n]} = X'$.

\[
Var(f(X)) = Ef(X)^2 - (Ef(X))^2
= Ef(X)^2 - Ef(X)Ef(X')
= Ef(X)(f(X) - f(X'))
= \sum_{i=1}^{n} Ef(X) \left( f(X^{[i-1]}) - f(X^{[i]}) \right)
\]

Fix $i$ and note that $f(X) \left( f(X^{[i-1]}) - f(X^{[i]}) \right)$ is a function of $(X_1, \ldots, X_n, X'_1, \ldots, X'_n) =: X^*$. The distribution of $X^*$ remains unchanged if we switch $X_i$ and $X'_i$. Under this switching operation,

\[
f(X) \left( f(X^{[i-1]}) - f(X^{[i]}) \right) \rightarrow f(X^{(i)}) \left( f(X^{[i]}) - f(X^{[i-1]}) \right),
\]

so these two quantities are equal in law. It follows that

\[
a = Ef(X) \left( f(X^{[i-1]}) - f(X^{[i]}) \right) = Ef(X^{(i)}) \left( f(X^{[i]}) - f(X^{[i-1]}) \right) = b.
\]

Observing that $a = b$ implies $a = b = (a + b)/2$, we obtain by Cauchy-Schwarz

\[
Ef(X) \left( f(X^{[i-1]}) - f(X^{[i]}) \right) = \frac{1}{2} Ef(X) Ef(X^{(i)}) \left( f(X^{[i-1]}) - f(X^{[i]}) \right)
\leq \frac{1}{2} \left( Ef(X)^2 \right) \left( Ef(X^{[i]})^2 \right) \left( Ef(X^{[i-1]})^2 \right)
= \frac{1}{2} Ef(X)^2,
\]

where the second step follows by noticing that

\[
Ef(X)^2 = Ef(X^{[i-1]})^2 = Ef(X^{[i]})^2
\]

(by $i - 1$ applications of the switching operation). Sum over $i$ to complete the proof.

2 Application: First Passage Percolation

We will apply the Efron-Stein inequality to study first-passage percolation. This section gives definitions; the application will be finished next class.
Consider the lattice $\mathbb{Z}^2$, with iid nonnegative random edge weights $(w_e)_{e \in E}$, $E$ being the set of edges of the lattice. Let $t_n = \min\{\sum_{e \in p} w_e : p \text{ is a path from } (0,0) \text{ to } (n,0)\}$. In words, $t_n$ is the first time that the vertex $(n,0)$ is reached by a liquid that is spilled onto the origin and which takes a random amount of time to flow over each edge in the graph.

Theorem\(^1\): $t_n/n \to \mu$ in probability where $\mu$ depends on the distribution of edge weights.

Note that the weight of a minimal path from $(0,0)$ to $(n+m,0)$ is less than or equal to the sum of the weights of the minimal paths from $(0,0)$ to $(0,n)$ and from $(0,n)$ to $(0,n+m)$. Therefore

$$E(t_{n+m}) \leq E(t_n) + E(t_m).$$

Next time we’ll begin with the subadditive lemma: if $\{a_n\}$ is a sequence of real random numbers satisfying $a_{n+m} \leq a_n + a_m$ (i.e., $\{a_n\}$ is subadditive), then $\lim_n a_n/n$ exists in $[-\infty, \infty)$ and equals $\inf_n a_n/n$.

---