Data Structures and Algorithms
Lecture Outline
March 31, 2016

Example. For any $\epsilon > 0$, if n balls are thrown independently and uniformly at random into $n^{2+\epsilon}$ bins, then with high probability, no bin contains more than one ball.

Solution. Let X_{ij} be an indicator random variable that is 1, iff ball i and ball j ($i \neq j$) land in the same bin. Let X be the random variable denoting the total number of collisions. Clearly, $X = \sum_{i \neq j} X_{ij}$. By the linearity of expectation, we have $\mathbf{E}[X] = \frac{\binom{n}{2}}{m}$. Note that if $m = n^2$ then $\mathbf{E}[X] < 1/2$, i.e, the expected number of pairwise collisions is less than $1/2$.

Let X_j denote the number of balls that land in bin j. We want to bound the probability that $X_j \geq 2$. To calculate $\Pr[X_j \geq 2]$, we first calculate $\Pr[X_j = 0]$ and $\Pr[X_j = 1]$.

$$\begin{aligned}
\Pr[X_j = 0] &= \left(1 - \frac{1}{m}\right)^n \\
&= \sum_{i=0}^{n} \binom{n}{i} \left(-\frac{1}{m}\right)^i \\
&= 1 - \frac{n}{m} + \binom{n}{2}\frac{1}{m^2} - \binom{n}{3}\frac{1}{m^3} + \cdots \\
&> 1 - \frac{n}{m}
\end{aligned}$$

Similarly, we lower bound the probability that bin j has exactly one ball.

$$\begin{aligned}
\Pr[X_j = 1] &= \left(\frac{n}{m}\right)\left(1 - \frac{1}{m}\right)^{n-1} \\
&= \frac{n}{m}\left(1 - \frac{n-1}{m} + \binom{n-1}{2} \cdot \frac{1}{m^2} - \binom{n-1}{3}\frac{1}{m^3} + \cdots\right) \\
&> \frac{n}{m} - \frac{n(n-1)}{m^2}
\end{aligned}$$

Thus, we have

$$\Pr[X_j \geq 2] < 1 - \left(1 - \frac{n}{m}\right) - \left(\frac{n}{m} - \frac{n(n-1)}{m^2}\right) = \frac{n(n-1)}{m^2}$$

Using the union bound we get

$$\Pr[X_1 \geq 2 \vee X_2 \geq 2 \vee \cdots \vee X_m \geq 2] < m \cdot \frac{n(n-1)}{m^2} = \frac{n(n-1)}{m}$$

By setting $m = n^{2+\epsilon}$, for any constant $\epsilon > 0$, we get that the probability of some bin having more than one ball is less than $1/n^\epsilon$ and this the probability of no bin having more than one ball is greater than $1 - \frac{1}{n^\epsilon}$.