

## 10 Tail Bounds (October 18)

### 10.1 Markov's Inequality

**Markov's Inequality.** If  $X$  is a non-negative integer random variable, then  $\Pr[X \geq t] \leq E[X]/t$  for any  $t > 0$ .

**Proof:**

$$\begin{aligned}
 E[X] &= \sum_{k=0}^{\infty} k \cdot \Pr[X = k] && \text{[def. expectation]} \\
 &= \sum_{k=0}^{\infty} \Pr[X \geq k] && \text{[algebra]} \\
 &\geq \sum_{k=0}^{t-1} \Pr[X \geq k] && \text{[since } k < \infty\text{]} \\
 &\geq \sum_{k=0}^{t-1} \Pr[X \geq t] && \text{[since } k < t\text{]} \\
 &= t \cdot \Pr[X \geq t] && \text{[algebra]}
 \end{aligned}$$

Since  $t > 0$ , we're done. □

### 10.2 Chernoff's Inequalities

Recall that random variables  $X$  and  $Y$  are *independent* if  $\Pr[X = x \wedge Y = y] = \Pr[X = x] \cdot \Pr[Y = y]$  for all  $x$  and  $y$ . If  $X$  and  $Y$  are independent, then  $E[XY] = E[X] \cdot E[Y]$ .

Let  $X = \sum_i X_i$  be a sum of independent random indicator variables  $X_i$ . For each  $i$ , let  $p_i = \Pr[X_i = 1]$ , and let  $\mu = E[X] = \sum_i E[X_i] = \sum_i p_i$ .

**Chernoff Bound (Upper Tail).**  $\Pr[X > (1 + \delta)\mu] < \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu$  for any  $\delta > 0$ .

**Proof:** The proof is fairly long, but it relies on just a few basic components: a clever substitution, Markov's inequality, the independence of the  $X_i$ 's, The World's Most Useful Inequality  $e^x > 1 + x$ , a tiny bit of calculus, and lots of high-school algebra.

We start by introducing a variable  $t$ , whose role will become clear shortly.

$$\Pr[X > (1 + \delta)\mu] = \Pr[e^{tX} > e^{t(1+\delta)\mu}]$$

To cut down on the superscripts, I'll usually write  $\exp(x)$  instead of  $e^x$  in the rest of the proof. Now apply Markov's inequality to the right side of this equation:

$$\Pr[X > (1 + \delta)\mu] < \frac{E[\exp(tX)]}{\exp(t(1 + \delta)\mu)}.$$

We can simplify the expectation on the right using the fact that the terms  $X_i$  are independent.

$$E[\exp(tX)] = E\left[\exp\left(t \sum_i X_i\right)\right] = E\left[\prod_i \exp(tX_i)\right] = \prod_i E[\exp(tX_i)]$$

We can bound the individual expectations  $E[e^{tX_i}]$  using The World's Most Useful Inequality:

$$E[\exp(tX_i)] = p_i e^t + (1 - p_i) = 1 + (e^t - 1)p_i < \exp((e^t - 1)p_i)$$

This inequality gives us a simple upper bound for  $E[e^{tX}]$ :

$$E[\exp(tX)] < \prod_i \exp((e^t - 1)p_i) < \exp\left(\sum_i (e^t - 1)p_i\right) = \exp((e^t - 1)\mu)$$

Substituting this back into our original fraction from Markov's inequality, we obtain

$$\Pr[X > (1 + \delta)\mu] < \frac{E[\exp(tX)]}{\exp(t(1 + \delta)\mu)} < \frac{\exp((e^t - 1)\mu)}{\exp(t(1 + \delta)\mu)} = (\exp(e^t - 1 - t(1 + \delta)))^\mu$$

Notice that this last inequality holds for *all* possible values of  $t$ . To obtain the final tail bound, we will choose  $t$  to make this bound as tight as possible. To minimize  $e^t - 1 - t - t\delta$ , we take its derivative with respect to  $t$  and set it to zero:

$$\frac{d}{dt}(e^t - 1 - t(1 + \delta)) = e^t - 1 - \delta = 0.$$

(And you thought calculus would never be useful!) This equation has just one solution  $t = \ln(1 + \delta)$ . Plugging this back into our bound gives us

$$\Pr[X > (1 + \delta)\mu] < (\exp(\delta - (1 + \delta)\ln(1 + \delta)))^\mu = \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}}\right)^\mu$$

And we're done! □

This form of the Chernoff bound can be a bit clumsy to use. A more complicated argument gives us the bound

$$\Pr[X > (1 + \delta)\mu] < e^{-\mu\delta^2/3}$$

for any  $0 < \delta < 1$ .

A similar inequality bounds the probability that  $X$  is much smaller than its expected value:

**Chernoff Bound (Lower Tail).**  $\Pr[X < (1 - \delta)\mu] < \left(\frac{e^\delta}{(1 - \delta)^{1 - \delta}}\right)^\mu < e^{-\mu\delta^2/2}$  for any  $\delta > 0$ .

### 10.3 Treaps

In our analysis of randomized treaps, we defined the indicator variable  $A_k^i$  to have the value 1 if and only if the node with the  $i$ th smallest key ('node  $i$ ') was a proper ancestor of the node with the  $k$ th smallest key ('node  $k$ '). We argued that

$$\Pr[A_k^i = 1] = \frac{[i \neq k]}{|k - i| + 1},$$

and from this we concluded that the expected depth of node  $k$  is

$$E[\text{depth}(k)] = \sum_{i=1}^n \Pr[A_k^i = 1] = H_k + H_{n-k} - 2 < 2 \ln n.$$

To prove a worst-case expected bound on the depth of the tree, we need to argue that the *maximum* depth of any node is small. Chernoff bounds make this argument easy.

**Lemma:** *The depth of a randomized treap with  $n$  nodes is  $O(\log n)$  with high probability.*

**Proof:** First let's bound the probability that the depth of node  $k$  is at most  $6 \ln n$ . (There's nothing special about the constant 6 here; I'm being somewhat generous to make the analysis easier.) The depth is a sum of independent indicator variables  $A_k^i$ , so we can apply Chernoff's inequality with  $\mu = E[\text{depth}(k)] < 2 \ln n$  and  $\delta = 2$ .

$$\begin{aligned} \Pr[\text{depth}(k) > 6 \ln n] &< \Pr[\text{depth}(k) > 3\mu] \\ &< \left(\frac{e^2}{3^3}\right)^\mu \\ &< \left(\frac{e^2}{3^3}\right)^{2 \ln n} \\ &= n^{2 \ln(e^2/3^3)} = n^{4-6 \ln 3} < \frac{1}{n^2}. \end{aligned}$$

(The last step just uses the fact that  $e < 3$ .)

Now consider the probability that the treap has depth greater than  $6 \ln n$ . Even though the distributions of different nodes' depths are not independent, we can conservatively bound the probability of failure as follows:

$$\Pr \left[ \max_k \text{depth}(k) > 6 \ln n \right] < \sum_k \Pr[\text{depth}(k) > 6 \ln n] < \frac{1}{n}.$$

More generally, this same argument implies that for any constant  $\Delta$ , the depth of the treap is less than  $2\Delta \ln n$  with probability at most  $1/n^{1+2\Delta(\ln \Delta - 1)}$ .  $\square$