Assignment 7

15-359 Probability and Computing
Name: Shashank Singh
Email: sss1@andrew.cmu.edu
Section: B
Due: Friday, March 30, 2012

Problem 1: Upper tail (10 pts.)

Clearly, examining the Taylor series expansion of $\ln(1+\delta)$, $\forall \delta \in (0, 1]$,

$$\delta^2/3 + \delta \leq (1+\delta)(\delta - \delta^2/2 + \delta^3/3 - \delta^4/4 \ldots) = (1+\delta)\ln(1+\delta).$$

Thus, since $x \mapsto e^x$ is an increasing function, $\frac{e^\delta}{(1+\delta)^{1+\delta}} < e^{-\delta^2/3}$ Therefore,

$$\left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu <= e^{-\mu\delta^2/3},$$

so that the strong Chernoff bound implies the simplified bound

$$P(X \geq (1+\delta)\mu) < e^{-\mu\delta^2/3} \quad \blacksquare.$$

Problem 2: The $\frac{1}{2}$-th moment (10 pts.)

Suppose $X \sim \text{Exp}(1)$. Then, substituting $u = \sqrt{x}$ gives

$$E\left[\sqrt{X}\right] = \int_0^\infty \sqrt{X} f_X(x)\, dx = \int_0^\infty \sqrt{X} e^{-x}\, dx = \int_0^\infty 2u^2 e^{-u^2}\, du.$$

Integration by parts then gives

$$E\left[\sqrt{X}\right] = -0 * e^{-0} - \lim_{x\to\infty}\left(-xe^{-x}\right) - \int_0^\infty -e^{-u^2}\, dx = \int_0^\infty e^{-u^2}\, dx = \frac{1}{2}\int_{-\infty}^\infty e^{-u^2}\, dx,$$

(since e^{-x^2} is even in x). This integral is well-known to be $\sqrt{\pi}$, so that

$$E\left[\sqrt{X}\right] = \boxed{\frac{\sqrt{\pi}}{2}.}$$

Problem 3: Sorting our loose ends (20 pts.)

A. $\sum_{i=1}^{t} X_i$ is the number of times, in the first t steps, that a ended up in the smaller of the two pieces of the array after partitioning around the pivot. Each time an array is partitioned into two pieces, the size of the smaller piece is at most half the size of the initial array. Thus, if a ended up in the smaller of the two pieces of the array partitioned $X = \sum_{i=1}^{t} X_i$ times, then it is in a part that is of size at most $\frac{n}{2^X}$, so that, when X exceeds $\log_2 n$, a is in a part of size 1, so that it must be in its correct position in the sorted array. ■

B. Since each X_i is 1 with probability $\frac{1}{2}$ and 0 otherwise, $X \sim \text{Binomial}(\frac{1}{2}, t)$, so that $\mu := E[X] = \frac{1}{2} C \log_2 n$. For $\delta = \frac{C-2}{2}$, then, by the simplified Chernoff bound for the Lower Tail,

$$P(X < \log_2 n) = P(X < (1-\delta)\mu) < e^{-\mu\delta^2/2}. \quad ■$$

C. If A is the event that the array is not sorted, and A_i is the event that a_i is not sorted, then, after $t = C \log_2 n$ iterations of quicksort, by the Union Bound,

$$P(A) \leq \sum_{i=1}^{t} P(A_i) \leq t e^{-\mu\delta^2/2}.$$

For $C \leq 4.6$, $c(c-2)^2/16 \geq 2$, so that $p(A_i) \leq \frac{1}{n^2}$. ■

Problem 5: Predictions are guesses anyway (8 pts.)

Suppose $f(h) \in \{0, 1\}$. Then, $F(h)$ is distributed according to $\text{Binomial}(P(f(h) = 1), n)$, so that, since $F(\Lambda) = E[F(h)]$ we can apply Hoeffding's Inequality to give that, $\forall \epsilon > 0$,

$$P(F(h) - F(\Lambda) > \epsilon) < e^{-2\epsilon^2 n}. \quad ■$$

Problem 6: PB wrap (15 pts.)

Let $s = F_m(h)$ and $d = F(h)$.

$$\begin{aligned}
E_s(Z) &= E_s\left(E_{h\sim\Pi}\left(\exp\left(m\left(s\ln\left(\frac{s}{d}\right) + (1-s)\ln\left(\frac{1-s}{1-d}\right)\right)\right)\right)\right) \\
&= E_s\left(E_{h\sim\Pi}\left(\left(\frac{s}{d}\right)^{m\cdot s}\left(\frac{1-s}{1-d}\right)^{m(1-s)}\right)\right).
\end{aligned}$$

s can take $m+1$ values: $\{0/m, 1/m, \ldots, m/m\}$ and has a binomial distribution with parameter(s) d and m. Thus,

$$\begin{aligned}
E_s\left(\left(\frac{s}{d}\right)^{m\cdot s}\left(\frac{1-s}{1-d}\right)^{m(1-s)}\right) &= \sum_{k=0}^{m}\binom{m}{k}d^k(1-d)^{m-k}\left(\frac{k/m}{d}\right)^k\left(\frac{1-k/m}{1-d}\right)^{m-k} \\
&= \sum_{k=0}^{m}\binom{m}{k}\left(\frac{k}{m}\right)^k\left(1-\frac{k}{m}\right)^{m-k}.
\end{aligned}$$

Since each term of this summation is a probability (e.g., of getting k heads out of m flips of a biased coin with k/m chance of getting heads), each of the $m+1$ terms is at most 1. Thus, $E_s(Z) \le m+1$.
■