Notes on Median and Quantile Estimation with Applications

Mark Huber
School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14850
mhuber@orie.cornell.edu

October 22, 1998

Abstract

We compute the rate of convergence for nonparametric selection estimators of the median using Chernoff type bounds, and consider some applications. The accuracy bound derived from Chernoff for Monte Carlo quantile estimators depends upon the quantity we wish to estimate. We give an estimator which runs for a nearly optimal amount of time. Finally, we give a formula for Chernoff bounds which is correct up to a multiplicative constant.

1 Introduction

Sampling to estimate values of medians and quantiles is a tool long familiar to statisticians. For a random variable X with distribution $F(x)$, we will call a point M_p a p-median if $P(X \leq M_p) > p$ and $P(X \geq M_p) > 1 - p$. Given an independent identically distributed sample $X_1, \ldots, X_n$ with order statistics $X_{(1)} \leq \ldots \leq X_{(n)}$. each with distribution F, the classical selection estimator for the p-median is $X_{(\lfloor pn \rfloor)}$. The accuracy of this estimator is often bounded using the central limit theorem. In section 2 we will give more precise bounds using a variant of the Chernoff Bound.

Finding a quantile estimate is in some sense the opposite of the median problem. Given a random variable X, suppose we wish to know $P(X \leq a)$ for some value a. A common estimate of this probability is to take n samples $X_1, \ldots, X_n$ compute $Y_i = 1_{\{X_i \leq a\}}$, and use $\sum_{i=1}^n Y_i/n$ as the estimate. This estimator has the unfortunate property that the relative accuracy is a function of $P(X \leq a)$, the very value we are trying to estimate. We present in section 3 a method which with high probability takes almost exactly as much data as needed to insure that the estimate is relatively accurate.

A tool often used for bounding how close sample approximations are to the correct value is the Central Limit Theorem (CLT). Consider how the CLT applies to binomial distributions. Variance of a single Bernoulli variable is $p(1-p)$ with mean p, so if we let $S_n = X_1 + \ldots + X_n$ the central limit theorem gives

$$\begin{aligned} P(S_n > (1+\delta)np) &\approx 1 - \Phi\left(\delta\sqrt{\frac{np}{1-p}}\right) \\ &= \int_{\sqrt{\frac{np}{1-p}}\delta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \int_1^{\infty} e^{-\left(\frac{np}{1-p}\right)\frac{(y\delta)^2}{2}} \sqrt{\frac{np}{1-p}}\delta\, dy \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\sqrt{2\pi}} \int_1^\infty y e^{-\left(\frac{np}{1-p}\right)\frac{(y\delta)^2}{2}} \sqrt{\frac{np}{1-p}}\delta\, dy \\
&\leq \frac{1}{\delta}\sqrt{\frac{2(1-p)}{\pi np}} \exp\left\{-\frac{\delta^2 np}{2(1-p)}\right\} \\
&= \sqrt{\frac{2}{\pi k}} e^{-k/2}
\end{aligned}
$$

where in the last equality we set $k = \delta^2 \frac{np}{1-p}$. Similarly, $P(S_n < (1-\delta)np) \leq \frac{1}{\sqrt{k}} e^{-k/2}$. In these expressions the exponential factor is dominant, indicating that we must set $n = O(\ln(\epsilon)(1-p)/(p\delta^2))$ to guarantee that $P(|S_n - np| > \delta np) < \epsilon$.

The following version of a Chernoff Bound nearly matches the Central Limit Theorem approximation given above.

Lemma 1 *Let $X_1, \ldots, X_n$ be $[0,1]$ with mean p, $\delta > 0$ and let $S_n = X_1 + \ldots + X_n$. Then*

$$P(S_n > (1+\delta)np) < e^{-np\delta^2/(2(1-p)(1+\delta))}$$

and

$$P(S_n < (1-\delta)np) < e^{-np\delta^2/(2(1-p)(1+\delta))}.$$

The proof of this lemma is in the appendix. The advantage of the Chernoff Bound is that it gives precise bounds on the probability attached to the quality of the estimate. In general, bounds on the accuracy of the Central Limit Theorem are somewhat weak, on the order of $\sqrt{n}$ [8]. In the next section we use this bound for the problem of estimating the median.

2 Median Estimation

In order to describe how a sampling algorithm may approximate the median, we need to define what we mean by an approximate median. Call $N_{(p,\delta)}$ a (p,δ)-median for a random variable X if $N_{(p,\delta)} = M_{p'}$ for some value of p' which satisfies $(1-\delta)p \leq p' \leq (1+\delta)p$. For example, if we wished to find a 0.4-median and $\delta = 0.1$, then any p-median where p ran from 0.36 to 0.44 would be a valid (0.4,0.1)-median. As with the definition of p-median, more than one value may be a valid (p,δ)-median.

Our estimate for the median is the selection estimator. Simply take k samples and use the kp order statistic for the estimate. Given our notation, the following theorem follows almost immediately from Lemma 1

Theorem. *The above algorithm yields a (ρ,δ)-median with probability at least $1-\epsilon$ if k samples are taken, where*

$$k = \frac{2(1+2\delta)}{p\delta^2} \ln\left(\frac{2}{\epsilon}\right).$$

For a $(1/2,\delta)$-median, the above algorithm yeilds a correct answer with prbability at least $1-\epsilon$ if k samples are taken, where

$$k = \frac{2(1+\delta-\delta^2)}{p\delta^2} \ln\left(\frac{2}{\epsilon}\right).$$

Proof: We will use the Chernoff Bound twice, first to show that the probability that the sample median is above $M_{(p(1+\delta))}$ is small, and then a second time to show that the probability that the sample median is below $M_{(p(1-\delta))}$ is also small. Consider the probability that more than kp samples lie too far below the true p-median. The prob that any 1 sample lies below $M_{(p(1-\delta))}$ is $p(1-\delta)$, so if kp samples were to lie below this value, we would have too many by a relative factor of a, where

$$\begin{aligned}(1+a)p(1-\delta) = p \quad &\Leftrightarrow \quad 1+a = \frac{1}{1-\delta} \\ &\Leftrightarrow \quad a = \frac{1-(1-\delta)}{1-\delta} \\ &\Leftrightarrow \quad a = \frac{\delta}{1-\delta}\end{aligned}$$

Now we prove a fact which will be useful in a moment.

$$\frac{a^2}{1+a} = \frac{\left(\frac{\delta}{1-\delta}\right)^2}{1+\frac{\delta}{1-\delta}} = \frac{\delta^2}{1+\delta} \cdot \frac{1}{1-\delta+\delta} = \frac{\delta^2}{1-\delta}.$$

Using Chernoff the chance that kp samples lie below $M_{(p(1-\delta))}$ is

$$\begin{aligned} exp\left\{\frac{-a^2}{1+a} \cdot \frac{p(1-\delta)k}{2(1-p(1-\delta))}\right\} \quad &\leq \quad exp\left\{\frac{-\delta^2}{1-\delta} \cdot \frac{p(1-\delta)k}{2}\right\} \\ &\leq \quad exp\left\{\ln\left(\frac{2}{\epsilon}\right)\right\} \\ &= \quad \epsilon/2 \end{aligned}$$

Note that we did not need the term $1+2\delta$ in our value of k. That term is needed for the second half of the calculation, where we look at the probability that the sample median is above $M_{(p(1-\delta))}$. The prob that any 1 sample lies above $M_{(p(1+\delta))}$ is $p(1+\delta)$, so if kp samples were to lie above this value, we would have too many by a relative factor of b, where

$$\begin{aligned}(1-b)p(1+\delta) = p \quad &\Leftrightarrow \quad 1-b = \frac{1}{1+\delta} \\ &\Leftrightarrow \quad b = \frac{(1+\delta)-1}{1+\delta} \\ &\Leftrightarrow \quad b = \frac{\delta}{1+\delta}.\end{aligned}$$

As before, a fact about b will be useful

$$\frac{b^2}{1+b} = \frac{\left(\frac{\delta}{1+\delta}\right)^2}{1+\frac{\delta}{1+\delta}} = \frac{\delta^2}{1+\delta} \cdot \frac{1}{1+\delta+\delta} = \frac{\delta^2}{(1+\delta)(1+2\delta)}$$

Hence the probability that fewer than kp samples lie below $M_{(p(1+\delta))}$ is

$$\begin{aligned} exp\left\{-\frac{b^2}{1+b} \cdot \frac{p(1+\delta)k}{2(1-p(1+\delta))}\right\} \quad &= \quad exp\left\{\frac{-\delta^2}{(1+\delta)(1+2\delta)} \cdot \frac{p(1+\delta)k}{2}\right\} \\ &= \quad exp\left\{\ln\left(\frac{2}{\epsilon}\right)\right\} \\ &= \quad \epsilon/2. \end{aligned}$$

Adding the probabilities for the two types of failures gives an upper bound on the total failure probability of $\epsilon/2 + \epsilon/2 = \epsilon$ which completes the proof for general p. If $p = 1/2$, then we simply do not neglect the $1 - p(1 - \delta) = 1/2(1 - \delta)$ term in the denominator of the exponential, and the final form for k is as given in the lemma.□

The advantage of finding an approximate median using a selection estimator is that the number of samples required is completely independent of the distribution used. This is a sharp constrast to estimating the expectation of a random variable, which is highly dependent on the size of the variable.

Moreover, in many cases (such as the exponential and uniform distributions) the median has a simple relationship to the parameter of the distribution. An exponential distribution with parameter λ will have a p-median at $-\log(p)/\lambda$ so from the value of the median will come the value of p. Cohen [1] uses this technique to give one estimator for the transitive neighborhood of a node in a graph.

3 Quantile Estimation

Given a random variable X with distribution F, the quantile estimation problem is to find an estimate for $P(X \leq a)$, that is, to approximate $F(a)$ for some a. A common estimate for this value comes from taking n samples $X_1, \ldots, X_n$, letting $Y_i = 1_{\{X_i \leq a\}}$, and using $\frac{1}{n}\sum_{i=1}^{n} Y_i$ as the estimate. (This is equivalent to using the empirical distribution function to estimate the actual distribution function at the point a.) The Y_i variables are 0-1 random variables with expectation $F(a)$, so we may use Lemma 1 to bound how quickly this estimate converges. Unfortunately, as was pointed out in Section 1, to achieve a desired accuracy requires that we know beforehand the value of p which we are trying to find. A bootstrapping approach could be employed to slowly narrow in on this value, but such an approach is inefficient. Usually lower bounds are computed for p using external information about the problem. However, the true value of p may be much larger than the lower bound, resulting in wasted iterations.

Fortunately in simulation and randomized algorithms, unlike many statistics applications, it is possible to generate as much data as is needed to make the estimate. We can therefore assume that we have an infinite stream of independent identically distributed random variables $X_1, X_2, \ldots$ of which we wish to use as few as possible to construct our estimate. We now introduce a sampling procedure first introduced by Dagum, Karp, Luby, and Ross [3].

Procedure: *Optimal Sampling*

Input k
Let $S_0 = 0$, and $m = 0$
Repeat until $S_m \geq k$
  Let $m \leftarrow m + 1$
  Sample X_m
  Let $S_m \leftarrow S_{m-1} + X_m$
Output $\bar{\mu} = \frac{k}{m}$

Figure 1: Optimal Sampling Procedure

In the usual methodology, the variable k is random and m is fixed, but in this algorithm k is fixed and m (which is also the number of samples needed by the procedure) is random. Since the random variables are $[0,1]$, Wald's identity [9] can be used to say that $p(1-\frac{1}{k+1}) < E[\bar{\mu}] \leq p$, so while this is not quite an unbiased estimation procedure, it is close.

The following lemma uses the Chernoff bound to show that this procedure produces a good estimate with high probability. Simultaneously, this shows that the probability of the running time being too long is also small. (Dagum, et. al. showed that in fact this method takes the fewest number of samples possible to obtain this level of accuracy.)

Lemma 2 *Let $\delta > 0$, $\epsilon > 0$, and $X_1, \ldots$ be a sequence of $[0,1]$ random variables with mean p. Then if*

$$k = 2\ln(2/\epsilon)(1+\delta)/\delta^2,$$

then

$$P(|\bar{\mu} - p| > \delta p) < \epsilon,$$

and

$$P(m > \frac{kp}{1-\delta}) < \epsilon/2.$$

Consider our estimate $\bar{\mu} = \frac{k}{m}$.

$$\begin{aligned} \frac{k}{m} > (1+\delta)p \quad &\Leftrightarrow \quad m < \left\lceil \frac{k}{p(1+\delta)} \right\rceil \\ &\Leftrightarrow \quad \sum_{i=1}^{\left\lceil \frac{k}{p(1+\delta)} \right\rceil} X_i > k \end{aligned}$$

So we may apply Chernoff with $n = \left\lceil \frac{k}{p(1+\delta)} \right\rceil$ so that $np(1+\delta) = k$ and we have

$$\begin{aligned} P\left(\frac{k}{m} > (1+\delta)p\right) &= P\left(\sum_{i=1}^{\left\lceil \frac{k}{p(1+\delta)} \right\rceil} X_i > k\right) \\ &\leq \exp(-k\delta^2/(2(1+\delta))) \\ &= \exp(-\ln(2/\epsilon) \\ &= \epsilon/2 \end{aligned}$$

Similarly,

$$\frac{k}{m} < (1-\delta)p \quad \Leftrightarrow \quad \sum_{i=1}^{\left\lfloor \frac{k}{p(1-\delta)} \right\rfloor} X_i < k$$

so applying Chernoff with $n = \left\lfloor \frac{k}{p(1-\delta)} \right\rfloor$ gives:

$$\begin{aligned} P\left(\frac{k}{m} < (1-\delta)p\right) &< \exp(-\left\lfloor \frac{k}{p(1-\delta)} \right\rfloor p\delta^2/2) \\ &= \exp\left(-\left\lfloor \frac{(1+\delta)\ln(2/\epsilon)}{p(1-\delta)\delta^2} \right\rfloor p\delta^2\right) \\ &= \exp\left(-\frac{(1+\delta)\ln(2/\epsilon)p\delta^2}{(1-\delta)p\delta^2} - p\delta^2\right) \\ &= \epsilon/2\exp\left(p\delta^2 - \frac{1+\delta}{1-\delta}\right) \\ &< \epsilon/2 \end{aligned}$$

Combining the two probabilities gives the desired result.

The property that makes this procedure useful is that the value of k does not depend at all upon p, the value that we wish to estimate. No complicated bootstrapping procedure is needed since the algorithm automatically runs for as long as is needed to ensure the desired accuracy.

This theorem is applicable to any situation in simulation where the results are bounded random variable, but the mean is not known in advance. Consider the problem of estimating the average shortest path length between pairs of distinct nodes in an unweighted graph. This value can range from 1 to D, the diameter of the graph. Then $X_i' = X_i/D$ are $[0,1]$ random variables, and we may apply this new sampling procedure. Since $X_i \geq 1$, we know that $p = E(X_i')$ satisfies $p \geq 1/D$. However, this is a quite weak lower bound and will in general take a longer period of time to run. The procedure above will run for a nearly optimal amount of time.

Another application of the continuous theorem occurs in estimating the price of an American option for a stock. Here the price of the option has a strict upper bound (the strike price of the option). Thus if the expected price is estimated using simulation this technique is applicable.

In many cases, the random variables that we are sampling from are not $[0,1]$, but are $\{0,1\}$. The fact that these variables only take on two values allows us to analyze them in a different fashion.

First we restate the sampling procedure for the discrete case.

Procedure: *Optimal Sampling–Bernoulli Random Variables*

Input k
Let $m = 0$
Repeat until at least k of $\{X_1, \ldots, X_m\}$ equal 1
  Let $m \leftarrow m + 1$
  Sample X_m
Output $\bar{\mu} = \frac{k}{m}$

Figure 1: Optimal Sampling for Bernoulli Random Variables

This way of looking at the algorithm for Bernoulli random variables gives an insight into a different analysis. We know that the difference between times of successive ones in a sequence of Bernoulli random variables is distributed geometrically with parameter p. We introduce the following variables. Let $W_n = \inf_{j \geq 0}\{\sum_{i=1}^{j} X_i \geq n\}$, and let $Z_i = W_i - W_{i-1}$. The Z_i's are just the distance between successive 1's in the sequence $X_1, X_2, \ldots$, and so are geometrically distributed. Also, $\sum_{i=1}^{k} Z_i = \sum_{i=1}^{k}(W_i - W_{i-1}) = W_k - W_0 = m - 0 = m$. Thus we wish to use a Chernoff type bound for the sum of geometrically distributed random variables. The following lemma is proved in the appendix.

Lemma 3 *Let $Z_1, \ldots, Z_k$ be geometric random variables with mean $1/p$, and let $S_k = Z_1 + \ldots + Z_k$. Then*

$$P(S_k > (1+\delta)k/p) < e^{-k\delta^2/(2(1+\delta))}$$

and

$$P(S_k < (1-\delta)k/p) < e^{-k\delta^2/2}.$$

As with the other analysis, the essential property of this bound for geometric random variables is that the probability of error does not depend on the mean in any way. Since m is the sum of k exponential random variables with mean $1/p$, this theorem immediately gives a bound on the probability that the number of samples needed deviates from k/p by more than a factor of δ. (This also tells us that $1/\bar{\mu}$ is an unbiased estimate for $1/p$). Since we know from Lemma 1 (and the CLT estimate) that k/p samples are needed in any case to ensure the desired accuracy, we have that this estimate uses (with high proability) nearly exactly the right amount of samples, where the 'nearly' is quantified by δ.

This lemma also shows that the estimate $\bar{\mu}$ will be a good one. Note that

$$\begin{aligned}(1-\delta')kp \leq m \leq (1+\delta')kp \quad &\Leftrightarrow \quad (1-\delta')p \leq 1/\bar{\mu} \leq (1+\delta')p \\ &\Leftrightarrow \quad p/(1-\delta') \geq \bar{\mu} \geq p/(1+\delta') \\ &\Leftrightarrow \quad (1+\delta)p \geq \bar{\mu} \geq (1-\delta)p\end{aligned}$$

where $\delta' = \delta/(1+\delta)$. We have proven the following corollary to Lemma 2.

Corollary 1 *Let $\epsilon > 0$, then if we set $k = \ln(3/\epsilon)(1+\delta')/\delta'^2$ then with probability at least $1-\epsilon$ both*

$$|\bar{\mu} - p| < \delta p$$

and

$$m < (1+\delta)k/p$$

where m is the number of samples taken in constructing the quantile estimate.

A variant of this technique for Bernoulli random variables was used in Karp, Luby, and Madras [7] in DNF counting. There they counted the number of geometrically distributed random variables which fit in a fixed amount of samples. In our case the geometric variables all have a common mean, but in their application the variables had different means and which mean was used was chosen according to a separate distribution.

A prime example of where this theorem comes into play is the case of Monte Carlo sampling. For example, consider the network reliability problem, which is as follows. Given a network with edges which fail independently with different probabilities, compute the probability that the graph becomes disconnected. The approximation algorithm for network reliability due to Karger [6] uses direct Monte Carlo sampling in the case that the probability the network fails is above $O(1/n^4)$, where n is the number of nodes in the network. This provides a lower bound on the probability of failure which may be far too low. By using the optimal sampling method this probability may be found directly using with high probability a nearly optimal number of samples.

In another example, Cohen and Lewis [2] used sampling to estimate matrix vector products. Let A be an n by m positive matrix and x be a a positive m vector (where by positive we mean that each entry of the matrix and vector are positive). The basic technique is as follows. First preprocess A, computing each column sum. Then construct a diagonal matrix D so that AD^{-1} has column sums equal to 1. Then $Ax = AD^{-1}Dx$. Setting $y = Dx/||Dx||$ and $\tilde{A} = AD^{-1}$, we have that $Ax = ||Dx||\tilde{A}y$. Since y has 1-norm 1 and the columns of $\tilde{A}$ sum to 1, we can treat the product as taking one step on a Markov Chain with initial distribution y and transition matrix $\tilde{A}^T$. We sample by first taking a random node from 1 to m according to the distribution y, then taking one step on the Markov Chain using the column of $\tilde{A}$ which corresponds to the random node chosen. The distribution of the nodes after one step will be $\tilde{A}y$. In other words, the distribution of our output is one of n nodes corresponding to the n elements of $\tilde{A}y$. Note that if we have the elements of $\tilde{A}y$ to within a relative factor of δ, then we also have the elements of Ax to within a relative factor of δ, since $Ax = \frac{1}{||Dx||}D\tilde{A}y$.

This is just an n-dimensional version of the problem of determining p, and we may directly apply our result to each of the n components individually. The number of samples needed to ensure that the resulting vector elements are within a relative factor of δ of their true values with probability greater than $1-\epsilon$ is $O(log(n/\epsilon)/(v_{\min}\delta^2)$. where $v_{\min}$ is the smallest element of $v = Ax$. Normally the error used is ϵ/n so that, even though the estimates for the values are not independent, the chance that all of the estimates are within a relative factor of δ is ϵ. The preprocessing step for A need only be done once and takes time $O(mn)$.

In general $v_{\min}$ can't be known in advance, but it is possible to estimate it. For example, v_i is bounded below by the minimum element of row i of A (say it is the jth element of the row) times the jth component of x. Estimates of this type may be many times smaller than the true answer, in which case unnecessary samples are being taken. In our method, the samples would be taken until each component was hit at least k times. Since ϵ/n is the error bound, it would only take longer than $(1+\delta)k/v_{min}$ samples with probability ϵ.

In some instances, not all of the vector v is needed. Cohen and Lewis also suggested using this technique for finding those values of v which lie above a threshold value. This kind of approach would also be needed if it was thought that v might have componenets which were 0. The threshold value (by definition) serves as a lower bound on the elements of v. Again the smallest element of v which is above the threshold value may be much larger than the threshold value, meaning that an unnecessary amount of work is being preformed.

4 Further Notes

The main problem in Monte Carlo sampling is that it is more difficult to lower bound p through sampling than it is to upper bound it. To upper bound p, simply set δ very high and use the Chernoff bound. In lower bounding p, δ must be less than 1, limiting our ability to get an accurate estimate.

Lemma 3 in section 3 was for geometric random variables, but a similar lemma holds for exponential random variables.

Lemma 4 *Let $Z_1, \ldots, Z_k$ be exponential random variables with mean $1/\lambda$, and let $S_k = Z_1 + \ldots + Z_k$. Then*

$$P(S_k > (1+\delta)k/\lambda) < e^{-k\delta^2/(2(1+\delta))}$$

and

$$P(S_k < (1-\delta)k/\lambda) < e^{-k\delta^2/2}.$$

One way to use this in quantile estimation is to introduce yet another set of random variables $R_1, R_2, \ldots$, which have an exponential distribution with mean 1, and which are independent of $X_1, X_2, \ldots$. Then pretend that the X_i's form a stochastic process with interarrival times R_i, so that the time interval between X_i and X_{i+1} is R_i. Then the X_i's form a Poisson process, and if we only consider those X_i's which equal 1, these form a thinned Markov Process, with interarrival times which are exponential with mean $1/p$. [9] Call these new interarrival times $Z_1', Z_2', \ldots$. Then if we set m as before, but now use $\tilde{\mu} = \frac{1}{m}\sum_{i=1}^{m} Z_i'$, we have an estimate which is just as good as the bounds we gave for $\bar{\mu}$.

An interesting fact about this approach appears in the multibin case, such as the matrix-vector multiplication example. There we generate n different thinned Poisson processes. The surprising fact here is that each of the processes are independent, unlike the geometric case. If ϵ is small this will not make much difference in the running time, however.

A Appendix–Proof of Chernoff Bounds

We begin by collecting some needed facts from calculus.

Lemma A.1 *For all real x,*

$$1 + x \leq e^x$$
$$x - \frac{x^2}{2} \leq \ln(1+x) \leq x$$

Proof. Note $x = \int_0^x 1\,dy \leq \int_0^x e^y dy = e^x - 1$, so $1 + x \leq e^x$. This gives Inequality 1, and taking the natural logarithm of both sides gives one inequality in 1. Since $1 \geq 1 - y^2$, $\frac{1}{1+y} \geq 1 - y$, and we can integrate to get $\int_0^x \frac{1}{1+y}\,dy \geq \int_0^x 1 - y\,dy$, so $\ln(1+x) \geq x - x^2/2$, which is the second inequality in 1. □

The following lemma embeds the heart of the Chernoff analysis.

Lemma A.2 *Let $X_1, X_2, \ldots, X_n$ be independent random variables with moment generating functions $m_1(t), m_2(t), \ldots, m_n(t)$, and expectations $p_1, \ldots, p_n$. (Here the moment generating function is the usual $m_i(t) = E(e^{tX})$.) Then*

$$P\left(\sum_{i=1}^n X_i > (1+\delta)\sum_{i=1}^n p_i\right) < \frac{\prod_{i=1}^n m_i(t)}{\exp\left(t(1+\delta)\sum p_i\right)}$$
$$P\left(\sum_{i=1}^n X_i < (1-\delta)\sum_{i=1}^n p_i\right) < \frac{\prod_{i=1}^n m_i(t)}{\exp\left(t(1-\delta)\sum p_i\right)}$$

for all $t > 0$.

PROOF: The function e^{tx} is 1-1 and increasing for $t > 0$, so we have that:

$$P\left(\sum_{i=1}^n X_i > (1+\delta)\sum_{i=1}^n p_i\right) = P\left(\exp\left(t\sum_{i=1}^n X_i\right) > \exp\left(t(1+\delta)\sum_{i=1}^n p_i\right)\right)$$

Utilizing Markov's Inequality, we have that

$$\begin{aligned} P\left(\sum_{i=1}^n X_i > (1+\delta)\sum_{i=1}^n p_i\right) &\leq E\left[\frac{\exp\left(t\sum_{i=1}^n X_i\right)}{\exp\left(t(1+\delta)\sum_{i=1}^n p_i)\right)}\right] \\ &= \frac{E\left[\prod_{i=1}^n \exp(tX_i)\right]}{\exp\left(t(1+\delta)\sum_{i=1}^n p_i\right)} \\ &= \frac{\prod_{i=1}^n E\left[\exp(tX_i)\right]}{\exp\left(t(1+\delta)\sum_{i=1}^n p_i\right)} \\ &= \frac{\prod_{i=1}^n m_i(t)}{\exp\left(t(1+\delta)\sum_{i=1}^n p_i\right)}. \end{aligned}$$

The result for $P\left(\sum_{i=1}^n X_i < (1-\delta)\sum_{i=1}^n p_i\right)$ is similar.

Throughout the rest of this paper, we will let

$$g(\delta) = \frac{1}{(1+\delta)^{(1+\delta)p}(1 - \frac{p}{1-p}\delta)^{1-p-p\delta}}.$$

The reason for this definition becomes clear in the following lemma.

Lemma A.3 *Let $X_1, \ldots, X_n$ be random variables with mean p which lie in $[0,1]$, and let $S_n = X_1 + \ldots + X_n$. Then for all $\delta > 0$,*

$$P(S_n > (1+\delta)np) < g(\delta)^n$$

and

$$P(S_n < (1-\delta)np) < g(-\delta)^n.$$

Proof. Bernoulli random variables are 1 with probability p and 0 otherwise, so trivially the moment generating functions are identically

$$(1-p)e^0 + pe^t = 1 - p + pe^t.$$

We now show that this is an upper bound for the moment generating function of any $[0,1]$ random variable with mean p. The technique we use is due to Hoeffding [5]. Consider that e^{tx} is a convex function, so the line segment from $(0, e^{t\cdot 0})$ to $(1, e^{t\cdot 1})$ lies entirely above the function e^{tx}. As a formula, this means for a $[0,1]$ random variable X:

$$\begin{aligned} e^{tX} &\leq 1 + X(e^t - 1) \\ E\left[e^{tX}\right] &\leq E\left[1 + X(e^t - 1)\right] \\ &= 1 + E[X](e^t - 1) \\ &= 1 - p + pe^t \end{aligned}$$

Using this result together with Lemma A.2 gives

$$\begin{aligned} P(S_n > (1+\delta)np) &< (1-p+pe^t)^n / \exp(t(1+\delta)np) \\ &= \left(\frac{1-p+pe^t}{\exp(t(1+\delta))p}\right)^n \end{aligned}$$

Let $f(t) = (1-p+pe^t)/(\exp(t(1+\delta)p)$. We wish to choose t to make $f(t)$ as small as possible. When $t = 0, f(t) = 1$. Note that if $(1+\delta)p \geq 1$, then $P(S_n > (1+\delta)np) < P(S_n > n) = 0$ so the bound is trivial. If $(1+\delta)p < 1$, then $f(t) \to \infty$ as $t \to \infty$. Differentiating $f(t)$ with respect to t yields:

$$f'(t) = -(1+\delta)pf(t) + \frac{pe^t}{1-p+pe^t}f(t)$$

Setting equal to 0 and solving for e^t gives

$$e^t = \frac{1+\delta}{1 - \frac{p}{1-p}\delta}.$$

Of course we could set t to any positive value and still have a valid inquality. By finding the minimum of $f(t)$, however, we make our bound as strong as possible.

We can now substitute our value for e^t into the inequality.

$$P(S_n > (1+\delta)np)^{1/n} < \frac{1 - p + \frac{p(1+\delta)}{1-\frac{p}{1-p}\delta}}{\left(\frac{1+\delta}{1-\frac{p}{1-p}\delta}\right)^{(1+\delta)p}}$$

$$
\begin{aligned}
&= \frac{(1-p-p\delta)+p(1+\delta)}{\left[1-\frac{p}{1-p}\delta\right]^{1-(1+\delta)p}[1+\delta]^{(1+\delta)p}} \\
&= \frac{1}{\left[1-\frac{p}{1-p}\delta\right]^{(1-p-p\delta)}[1+\delta]^{(1+\delta)p}} \\
&= g(\delta)
\end{aligned}
$$

□

In light of this lemma, it makes sense to develop some facts about $g(y)$.

Lemma A.4 *Let* $\delta > 0$. *Then*

$$
e^{-\delta^2 p/(2(1-p))} < g(\delta) < \left(\frac{1+\delta}{e^{\delta}}\right)^{p/(1-p)} < e^{-\delta^2 p/(2(1-p)(1+\delta))},
$$

$$
e^{-\delta^2 p/(2(1-p))} > \left(\frac{1-\delta}{e^{-\delta}}\right)^{p/(1-p)} > g(-\delta)
$$

Proof. Set $g(y) = \left((1-\frac{p}{1-p}y)\right)^{-(1-p-py)}\left((1+y)^{-(1+y)}\right)$. Then

$$
\begin{aligned}
g'(y) &= \left[\frac{\frac{p}{1-p}(1-p-py)}{1-\frac{p}{1-p}} - p\ln\left(1-\frac{p}{1-p}\right)\right]g(y) - \left[\frac{(1+y)p}{1+y} + p\ln(1+\delta)\right]g(y) \\
&= g(y)p\ln\left(\frac{1-\frac{p}{1-p}y}{1+y}\right) \\
&= g(y)p\ln\left(1-\frac{(1+y)-(1-\frac{p}{1-p}y)}{1+y}\right) \\
&= g(y)p\ln\left(1-\frac{\frac{1}{1-p}y}{1+y}\right) \\
&\leq g(y)p\left(-\frac{\frac{1}{1-p}y}{1+y}\right) \\
&= -\frac{py}{(1-p)(1+y)}g(y)
\end{aligned}
$$

Now we take look at the logarithm of $g(y)$.

$$
\begin{aligned}
\frac{d}{dy}\ln(g(y)) &= \frac{g'(y)}{g(y)} \\
&\leq \frac{-py}{(1-p)(1+y)}
\end{aligned}
$$

Using partial fractions gives:

$$
\begin{aligned}
\ln(g(x)) &= \ln(g(0)) + \int_0^x \frac{-py}{(1-p)(1+y)}dy \\
&\leq \frac{p[\ln(1+x)-x]}{2(1-p)},
\end{aligned}
$$

and so

$$g(\delta) \leq \left(\frac{1+\delta}{e^{\delta}}\right)^{p/(1-p)},$$

and

$$g(-\delta) \leq \left(\frac{1+\delta}{e^{\delta}}\right)^{p/(1-p)}.$$

Similarly, we may use the other half of Inequality 1 and say that

$$\begin{aligned}\frac{g'(y)}{g(y)} &= p\ln\left(1-\frac{\frac{1}{1-p}y}{1+y}\right)\\ &\geq p\left(\left(\frac{\frac{1}{1-p}y}{1+y}\right)-\frac{1}{2}\left(\frac{\frac{1}{1-p}y}{1+y}\right)^2\right)\end{aligned}$$

Since $p \leq 0.5$, $2(1-p)y \leq 1$, and

$$\begin{aligned}\frac{g'(y)}{g(y)} &\geq \frac{p}{1-p}\left[\left(\frac{y}{1+y}\right)-\left(\frac{y^2}{(1+y)^2}\right)\right]\\ &= -\frac{p}{1-p}\left[\frac{y}{(1+y)^2}\right]\end{aligned}$$

Integrating yields

$$\begin{aligned}\ln(g(y)) &\geq -\frac{p}{1-p}\left(\ln(1+x)+\frac{x}{1+x}\right)\\ g(y) &\geq \left(\frac{e^{x/(1+x)}}{1+x}\right)^{\frac{p}{1-p}}\end{aligned}$$

□

In the case that $X_1, \ldots$ are Bernoulli random variables with p, we can show that the tail probabilities are not only $O(g(\delta))$, they are also $\Omega(g(\delta))$.

Lemma A.5 *Let $X_1, \ldots, X_n$ be 0,1 random variables with mean $p \leq 1/2$, $\delta > 0$, $k = \frac{np\delta^2}{1-p}$, $C = \frac{1+\delta}{\sqrt{2\pi k(1-\frac{p}{1-p}\delta)}}$ and $S_n = X_1 + \ldots + X_n$. Then if $k > 1$,*

$$\frac{1}{1+\lceil\frac{1}{\sqrt{k}}\rceil}e^{-\frac{31}{180n}}\left(1-e^{-\sqrt{k}}\right)Cg(\delta)^n < P(S_n > (1+\delta)np) < Cg(\delta)^n$$

and

$$\frac{1}{1+\lceil\frac{1}{\sqrt{k}}\rceil}e^{-\frac{31}{180n}}\left(1-e^{-\sqrt{k}}\right)Cg(-\delta)^n < P(S_n < (1-\delta)np) < Cg(-\delta)^n.$$

Note that as $k \to \infty$, the function of k in front of the less than inequalities goes to 1.

Proof. For convienence we will assume that δnp is an integer. Note that

$$P(S_n > (1+\delta)np) = \sum_{j=(1+\delta)np}^{n}\binom{n}{j}p^j(1-p)^{n-j}.$$

In order to bound this quantity, we shall use Stirling's formula, the form of which we use is [4]

$$e^{-n+\frac{1}{12n}-\frac{1}{360n^3}}\sqrt{2\pi}n^{(n+1/2)} \leq n! \leq e^{-n+\frac{1}{12n}}\sqrt{2\pi}n^{(n+1/2)}.$$

For the binomial coefficients, this means:

$$\begin{aligned}\binom{n}{j} &= \frac{n!}{j!(n-j)!}\\ &\leq \frac{e^{\frac{1}{12n}}e^{-n}\sqrt{2\pi}n^{(n+1/2)}}{e^{\frac{1}{12n}-\frac{1}{360n^3}}e^{-j}\sqrt{2\pi}j^{(j+1/2)}e^{\frac{1}{12n}-\frac{1}{360n^3}}e^{n-j}\sqrt{2\pi}(n-j)^{(n-j+1/2)}}\\ &\leq e^{\frac{1}{180n^3}-\frac{1}{12n}}\frac{n^{(n+1/2)}}{\sqrt{2\pi}j^{(j+1/2)}(n-j)^{(n-j+1/2)}}\end{aligned}$$

Similarly

$$e^{-\frac{1}{360n^3}-\frac{1}{12n}}\frac{n^{(n+1/2)}}{\sqrt{2\pi}j^{(j+1/2)}(n-j)^{(n-j+1/2)}} \leq \binom{n}{j}.$$

Put another way, we have that for a binomial variable S_n with parameters p, n, that

$$e^{-\frac{31}{360}}\frac{\sqrt{n}}{\sqrt{2\pi j(n-k))}}\cdot\left(\frac{np}{j}\right)^j\left(\frac{n(1-p)}{n-j}\right)^{(n-j)} \leq P(S_n=j) \leq \frac{\sqrt{n}}{\sqrt{2\pi nj(n-j)}}\cdot\left(\frac{np}{j}\right)^j\left(\frac{n(1-p)}{n-j}\right)^{(n-j)}.$$

In our case $j=(1+\delta)np$, so

$$\begin{aligned}\frac{1}{\sqrt{2\pi np(1-p)}}\cdot\left(\frac{np}{k}\right)^k\left(\frac{n(1-p)}{n-k}\right)^{(n-k)} &= \frac{1}{\sqrt{2\pi np(1-p)}}\cdot\frac{1}{(1+\delta)^{(1+\delta)np}(\frac{n-(1+\delta)np}{1-p})^{(n-(1+\delta)np)}}\\ &= \frac{1}{\sqrt{2\pi np(1-p)}}\cdot\left(\frac{1}{(1+\delta)^{(1+\delta)p}(1+\frac{p}{1-p}\delta)^{(1-p-p\delta)}}\right)^n\\ &= g(\delta)^n\end{aligned}$$

So our inequality becomes

$$e^{-\frac{31}{360}}\frac{1}{\sqrt{2\pi np(1-p)}}g(\delta)^n \leq P(S_n=j) \leq \frac{1}{\sqrt{2\pi np(1-p)}}g(\delta)^n.$$

Now consider the ratio

$$\begin{aligned}P(S_n=j+1)/P(S_n=j) &= \frac{\frac{n!}{(j+1)!(n-j-1)!}p^{(j+1)}(1-p)^{(n-j-1)}}{\frac{n!}{j!(n-j)!}p^j(1-p)^{(n-j)}}\\ &= \frac{(n-j)p}{(j+1)(1-p)}\\ &\leq \frac{(n-j)p}{j(1-p)}\end{aligned}$$

If $j > (1+\delta)np$, then

$$P(S_n=j+1)/P(S_n=j) \leq \frac{1-\frac{p}{1-p}\delta}{1+\delta},$$

which is a decreasing function in δ. Thus we may bound $P(S_n \geq (1+\delta)np)$ above by a decreasing infinite geometric series starting at $P(S_n = (1+\delta)np)$ and with ratio $r = \frac{1-\frac{p}{1-p}\delta}{1+\delta}$. Through elementary algebra, one can show that

$$\frac{1}{1-r} = \frac{(1+\delta)(1-p)}{\delta}.$$

Hence

$$\begin{aligned} P(S_n \geq (1+\delta)np) &\leq \frac{(1+\delta)(1-p)}{\delta\sqrt{2\pi n(1+\delta)p(1-p-p\delta)}} g(\delta)^n \\ &\geq \sqrt{\frac{(1+\delta)(1-p)}{2\pi np(1-\frac{p}{1-p}\delta)}} g(\delta)^n \end{aligned}$$

On the other hand,

$$P(S_n = j+1)/P(S_n = j) \geq \frac{1-\frac{p}{1-p}z\delta}{1+z\delta},$$

for $(1+\delta)np \leq j \leq \lceil(1+z\delta)np\rceil$. Then

$$\frac{1}{1-r} \geq \frac{1+z\delta}{(1-p)z\delta}.$$

and

$$\begin{aligned} P(S_n \geq (1+\delta)np) &\geq \frac{(1+z\delta)(1-p)}{z\delta\sqrt{2\pi n(1+z\delta)p(1-p-p\delta)}}(g(\delta)^n - g(2\delta)^n) \\ &\leq \frac{1}{2}\sqrt{\frac{(1+\delta)(1-p)}{2\pi np(1-\frac{p}{1-p}\delta)}}\left(1-\left(\frac{g(2\delta)}{g(\delta)}\right)\right) g(\delta)^n \end{aligned}$$

Recall from the proof of Lemma A.4 that

$$\frac{d}{dy}\ln(g(y)) \leq -\frac{py}{(1-p)(1+y)}.$$

So,

$$\begin{aligned} \ln(g((1+z)\delta)) &\leq \ln(g(\delta)) + \int_{\delta}^{(1+z)\delta} -\frac{py}{(1-p)(1+y)}dy \\ \ln(\frac{g((1+z)\delta)}{\delta}) &\leq -\frac{p}{1-p}\int_{\delta}^{(1+z)\delta} \frac{y}{1+z\delta}dy \\ &\leq -\left(\frac{p}{(1-p)(1+z\delta)}\right)\left(\frac{(z\delta)^2-\delta^2}{2}\right) \\ \frac{g(z\delta)}{g(\delta)} &\leq e^{-\frac{p\delta^2}{2(1-p)}} \\ \left(\frac{g(z\delta)}{g(\delta)}\right)^n &\leq e^{-3k/2} \end{aligned}$$

Let $h(y) = \left(\frac{1+y}{e^y}\right)$. It is easy to verify that $h(y)$ satisfies the differential equation

$$\frac{d}{dy}l\ln(h(y)) = -\frac{y}{1+y}$$

Hence for $\delta \geq 0$,

$$\begin{aligned}
\ln(h(\delta)) &= \ln(h(0)) + \int_0^\delta -\frac{y}{1+y}\,dy \\
&< \int_0^\delta -\frac{y}{1+\delta}\,dy \\
&= -\delta^2/[2(1+\delta)] \\
h(y) &< e^{-\delta^2/[2(1+\delta)]}
\end{aligned}$$

and

$$\begin{aligned}
\ln(h(-\delta)) &= \ln(h(0)) + \int_0^\delta -\frac{y}{1+y}\,dy \\
&> \int_0^\delta -\frac{y}{-}dy \\
h(-\delta) &> e^(- \delta^2/2)
\end{aligned}$$

completing the proof.

□

References

[1] Edith Cohen. Size-estimation framework with applications to transitive closure and reachability. *J. Comput. System Sci.*, 55(3):441–453, 1997.

[2] Edith Cohen and David D. Lewis. Approximating matrix multiplication for pattern recognition tasks. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 682–691, 1997.

[3] Paul Dagum, Richard Karp, Michael Luby, and Sheldon Ross. An optimal algorithm for Monte Carlo estimation (extended abstract). In *Proceeding of the 36th Annual Symposium on Foundations of Computer Science*, pages 142–149, 1995.

[4] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics.* Addison-Wesley Pub. Co., 1994.

[5] Wassily Hoeffding. On probabilities of large deviations. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics*, pages 203–219. Univ. California Press, Berkeley, Calif., 1967.

[6] David R. Karger. A randomized fully polynomial time approximation scheme for the all terminal network reliability problem. In *Proccedings of the Twenty-Seventh Annual ACM Symposium on the Theory of Computing*, pages 11–17, 1995.

[7] Richard M. Karp, Michael Luby, and Neal Madras. Monte carlo approximation algorithms for enumeration problems. *J. Algorithms*, 10(3):429–448, 1989.

[8] Sidney C. Port. *Theoretical Probability for Applications.* Wiley-Interscience, 1994.

[9] Sidney Resnick. *Adventures in Stochastic Processes.* Birkhäuser, 1992.