A Short Note on Binomial Tails for Vanishing Mean

John Sylvester*

Department of Computer Science & Technology, University of Cambridge, UK

January 9, 2020

Abstract

In this note we shall consider the upper tail $\mathbb{P}[\mathrm{Bin}(n,p) \geq k]$ for the Binomial distribution $\mathrm{Bin}(n,p)$ when $np \to 0$ and $k > np$. We derive a simple expression for $\mathbb{P}[\mathrm{Bin}(n,p) \geq k]$ which shows that the Chernoff bound for the Binomial is out by a factor $\sqrt{2\pi k}$ aysmtotically .

Let $\mathrm{Bin}(n,p)$ denote a binomially distributed random variable with parameters n and p, that is for $k \in \{0,\dots,n\}$,

$$\mathbb{P}[\mathrm{Bin}(n,p) = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

Our main result is the following exact asymptotic expression for the upper tail of $\mathrm{Bin}(n,p)$.

Theorem 1. *Let $k \in \mathbb{N}$, $k \geq 1$ and $p = p(n) \in (0,1)$ be such that $np \to 0$. Then*

$$\mathbb{P}[\mathrm{Bin}(n,p) \geq k] = \frac{1+o(1)}{\sqrt{2\pi k}} \left(\frac{npe}{k}\right)^k = \Theta\left(\frac{e^{-k(\log k - \log np - 1)}}{\sqrt{k}}\right).$$

Where in the above the $o(1)$ is respect to n.

Recall the classical Chernoff bound [1, Thm 4.4(1)] for the Binomial distribution

$$\mathbb{P}[\mathrm{Bin}(n,p) \geq (1+\delta)np] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{np},$$

for any $\delta > 0$. Thus letting $\delta = k/np - 1$ for $k > np > 0$ we have

$$\mathbb{P}[\mathrm{Bin}(n,p) \geq k] \leq \left(\frac{e^{k/np-1}}{(k/np)^{k/np}}\right)^{np} = e^{-np}\left(\frac{npe}{k}\right)^k. \tag{1}$$

Combining (1) with the result of Theorem 1 shows that the Chernoff bound over estimates the upper tail by a multiplicative $\sqrt{2\pi k}$ factor.

Theorem 1 will follow from the expression for the upper tail below.

Lemma 2. *The following holds for any $0 \leq k \leq n$ and $p = p(n) \in (0,1)$,*

$$\begin{aligned}\mathbb{P}[\mathrm{Bin}(n,p) > k] &= \sum_{j=k+1}^{n} (-p)^{j+1} \binom{n}{j} \sum_{i=0}^{k} (-1)^i \binom{j}{i} \\ &= \binom{n}{k+1} p^{k+1} - \sum_{j=k+2}^{n} (-1)^{k+j} \binom{j-1}{k} \binom{n}{j} p^j.\end{aligned}$$

*The author is currently supported by Thomas Sauerwald's ERC Starting Grant 679660 (DYNAMIC MARCH).

Proof. Recall the identity $\binom{n}{k}\binom{n-k}{j} = \binom{n}{k+j}\binom{k+j}{k}$ and observe the following

$$\begin{aligned}\mathbb{P}[\mathrm{Bin}(n,p)=k] &= \binom{n}{k}p^k(1-p)^{n-k}\\ &= \binom{n}{k}p^k\sum_{j=0}^{n-k}(-1)^j\binom{n-k}{j}p^j\\ &= \sum_{j=0}^{n-k}(-1)^j\binom{n}{k+j}\binom{k+j}{k}p^{k+j}\\ &= \sum_{j=k}^{n}(-p)^j\binom{n}{j}(-1)^k\binom{j}{k}.\end{aligned}$$

Recall the identity $\sum_{i=0}^{k}(-1)^i\binom{k}{i} = 0$. Using the above expression for $\mathbb{P}[\mathrm{Bin}(n,p)=k]$ we shall now prove the first equality in the Lemma's statement by induction on k,

$$\begin{aligned}\mathbb{P}[\mathrm{Bin}(n,p)\le k] &= 1+\sum_{j=k}^{n}(-p)^j\binom{n}{j}\sum_{i=0}^{k-1}(-1)^i\binom{j}{i} + \sum_{j=k}^{n}(-p)^j\binom{n}{j}(-1)^k\binom{j}{k}\\ &= 1+(-p)^k\binom{n}{k}\left(\sum_{i=0}^{k-1}(-1)^i\binom{k}{i}+(-1)^k\binom{k}{k}\right)\\ &\quad+\sum_{j=k+1}^{n}(-p)^j\binom{n}{j}\left(\sum_{i=0}^{k-1}(-1)^i\binom{j}{i}+(-1)^k\binom{j}{k}\right)\\ &= 1+\sum_{j=k+1}^{n}(-p)^j\binom{n}{j}\sum_{i=0}^{k}(-1)^i\binom{j}{i}, \qquad (2)\end{aligned}$$

where the base case holds since $\mathbb{P}[\mathrm{Bin}(n,p)\le 0] = (1-p)^n$.

Recall the identity $\sum_{i=0}^{k}(-1)^i\binom{j}{i} = (-1)^k\binom{j-1}{k}$, combining this with (2) yields

$$\begin{aligned}\mathbb{P}[\mathrm{Bin}(n,p)\le k] &= 1+(-p)^{k+1}\binom{n}{k+1}(-1)^k\binom{k}{k}+\sum_{j=k+2}^{n}(-p)^j\binom{n}{j}(-1)^k\binom{j-1}{k}\\ &= 1-\binom{n}{k+1}p^{k+1}+\sum_{j=k+2}^{n}(-1)^{k+j}\binom{j-1}{k}\binom{n}{j}p^j,\end{aligned}$$

which gives the second identity from the statement. □

Proof of Theorem 1. Notice that

$$\frac{\binom{j}{k}\binom{n}{j+1}p^{j+1}}{\binom{j-1}{k}\binom{n}{j}p^j} = \frac{j(n-j)p}{(j-k)(j+1)} \le np.$$

So if $np\to 0$ then we have the following by Lemma 2

$$\mathbb{P}[\mathrm{Bin}(n,p)>k] = \binom{n}{k+1}p^{k+1}(1+\mathcal{O}(np)) = \frac{1+o(1)}{\sqrt{2\pi k}}\left(\frac{npe}{k}\right)^k,$$

by Stirling's inequality. □

References

[1] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, 2005.