Journal of Statistical Theory and Practice (2021) 15:35
https://doi.org/10.1007/s42519-020-00156-x

ORIGINAL ARTICLE

Exponential Tail Bounds for Chisquared Random Variables

Malay Ghosh[1]

Accepted: 14 December 2020 / Published online: 4 March 2021
© The Author(s) 2021

Abstract
The paper derives some exponential tail bounds for central and non-central chisquared random variables. The bounds are simple and can easily be applied in statistical analysis. Especially relevant are the tail bounds for non-central chisquares, which are different from some of the other exponential bounds available in the literature, for example the one given in [1].

Keywords F-statistic · Moderate deviations · Non central · Sub-exponential

1 Introduction

The objective of this note is to derive some exponential tail bounds for chisquared random variables. The bounds are non-asymptotic, but they can be used very successfully for asymptotic derivations as well. As a corollary, one can get tail bounds for F-statistics as well. Also, I show how some exact moderate deviation [4] inequalities can be obtained as special cases of these tail bounds.

The chisquared random variables are special cases of sub-exponential random variables. We examine when the bounds obtained here are sharper than the ones that use only the sub-exponentiality of chisquares.

The outline of the next two sections is as follows. Exponential tail bounds for central chisquares are given in Sect. 2. Corresponding bounds for non-central chisquares are given in Sect. 3.

This article is part of the topical collection "Celebrating the Centenary of Professor C. R. Rao" guest edited by, Ravi Khattree, Sreenivasa Rao Jammalamadaka, and M. B. Rao.

✉ Malay Ghosh
ghoshm@stat.ufl.edu

1 Department of Statistics, University of Florida, Box 118545, Gainesville, FL 32611-8545, USA

2 Central Chisquare

We begin with an upper tail bound for central chisquares. The following theorem is proved.

Theorem 1 *Suppose* $X \sim \chi^2_p$. *Then for* $a > p$, $P(X > a) \le \exp[-\frac{p}{2}\{\frac{a}{p} - 1 - \log(\frac{a}{p})\}]$
.

Proof From Markov's inequality in its exponential form (see for example, [2] or [3], one gets

$$\begin{aligned} P(X > a) &\le \inf_{0<t<1/2}[\exp(-ta)E\{\exp(tX)\}] \\ &\le \inf_{0<t<1/2}[\exp(-ta)(1-2t)^{-p/2}]. \end{aligned} \tag{1}$$

Let $g(t) = -ta - (p/2)\log(1-2t)$. Then $g'(t) = -a + p(1-2t)^{-1}$ and $g''(t) = 2p(1-2t)^{-2}(>0)$. Hence, $g(t)$ is minimized at $t = t_0 = (1/2)\left(1 - \frac{p}{a}\right)$. Substitution in (1) yields

$$\begin{aligned} P(X > a) &\le \exp(-t_0 a)(1-2t_0)^{-p/2} \\ &= \exp\left[-\frac{a}{2}\left(1 - \frac{p}{a} + \frac{p}{2}\log\left(\frac{a}{p}\right)\right)\right] \\ &= \exp\left[-\frac{p}{2}\left(\frac{a}{p} - 1 - \log\left(\frac{a}{p}\right)\right)\right]. \end{aligned}$$

This proves the theorem.

Suppose $a = p + c$. Then an equivalent way of writing the above result is

$$P(X - p > c) \le \exp\left[-\frac{p}{2}\left\{\frac{c}{p} - \log\left(1 + \frac{c}{p}\right)\right\}\right]. \tag{2}$$

By the inequality

$$x - \log(1+x) = \int_0^x \left[1 - \frac{1}{1+y}\right]\mathrm{d}y = \int_0^x \frac{y}{1+y}\mathrm{d}y \ge \frac{x^2}{1+x},$$

a weaker version of (2) is given by

$$P(X > p + c) \le \exp\left[-\frac{p}{2}(c/p)^2(1 + c/p)^{-1}\right] \le \exp\left[-\frac{c^2}{4(p+c)}\right]. \tag{3}$$

It may be noted that a chisquare random variable is a special case of a sub-exponential random variable. There are several equivalent definitions of sub-exponential random variables. The one we find convenient is given as follows (see [5], p 26).

Definition A random variable X with mean μ is said to be sub-exponential (ν, α) if $E[\exp\{t(X - i\mu)\}] \le \exp(t^2\nu^2/2)$ whenever $|t| < \alpha^{-1}$.

If $X \sim \chi^2_p$, then X is subexponential $(2p, 4)$. To see this, we note that $E[\exp\{t(X - p)\}] = \exp(-tp)(1 - 2t)^{-p/2} \le \exp(2pt^2)$ with $|t| \le 1/4$. Now $P(X - p > c) \le \inf_{0<t<1/4} \exp(-tc + 2pt^2) = \exp(-c^2/(8p))$. The inequality given in (3) is sharper than the last one when $0 < c < p$. Moreover, as $p \to \infty$, it follows from (2) that $P(X - p > c) \sim \exp(-c^2/(4p))$, while sub-exponentiality continues to yield the same upper bound $\exp(-c^2/(8p))$.

The next inequality is related to the lower tail of a chisquared random variable. The following theorem is proved.

Theorem 2 *Suppose* $X \sim \chi^2_p$. *Then for* $0 < c < p$,

$$P(X - p < -c) \le \exp\left[(p/2)\left\{\frac{c}{p} + \log\left(1 - \frac{c}{p}\right)\right\}\right] \le \exp\left(-\frac{c^2}{4p}\right).$$

Proof The second inequality is an easy consequence of expansion of a logarithmic function. To prove the first inequality, we begin with

$$P(X - p < -c) \le \inf_{t<0}\left[\exp\{-t(p - c)\}(1 - 2t)^{-p/2}\right]. \tag{4}$$

Similar as before, let $g(t) = -t(p - c) - (p/2)\log(1 - 2t)$. Then $g'(t) = -(p - c) + p(1 - 2t)^{-1}$ and $g''(t) = 2p(1 - 2t)^{-2}$. Hence, $g(t)$ is minimized at $t = t_0$ where $1 - 2t_0 = p/(p - c)$, i.e., $t_0 = -c/[2(p - c)]$. Substituting t_0 for t in (4), one gets the inequality

$$P(X - p < -c) \le \exp(c)\left(\frac{p}{p - c}\right)^{-p/2} = \exp\left[\frac{p}{2}\left\{\frac{c}{p} + \log\left(1 - \frac{c}{p}\right)\right\}\right].$$

This proves the theorem.

The exact upper bound given in the rightmost side of Theorem 2 is stronger than the similar sub-exponential bound $\exp\left(-\frac{c^2}{8p}\right)$. Moreover, since $p + c > p$, it is possible to combine (3) with Theorem 2 to get the inequality

$$P(|\chi^2_p - p| > c) \le 2\exp\left[-\frac{c^2}{4(p + c)}\right]. \tag{5}$$

Since χ^2_p/p is the average of p iid χ^2_1 random variables, each with mean 1 and variance 2, the central limit theorem leads to $(\chi^2_p - p)/(\sqrt{2p} \xrightarrow{d} \mathrm{N}\,(0, 1)$. For averages of p iid random variables with nonzero and finite variance, [4] provided an asymptotic two sided tail bound for deviations of the order $\sqrt{\log p}$. Putting $c = \sqrt{2p\log p}$, one gets the asymptotic upper bound $\exp\left(-\frac{2p\log p}{4p}\right) = p^{-1/2}$ which is slightly weaker than $O(p^{-1/2}(\log p)^{-1/2})$, one obtained by Rubin and Sethuraman in conformity with Mill's ratio.

It is possible to use Theorems 1 and 2 to obtain some crude tail bounds for the F-statistic as well. To see this, suppose rXx and Y are two independent chisquared random variables with respective degrees of freedom m_1 and m_2. I write $F = \frac{X/m_1}{Y/m_2}$. Then for $d > 1$, and writing $\delta = (d-1)/(d+1)$,

$$\begin{aligned} P(F > d) &= P[F > (1+\delta)/(1-\delta)] \\ &\leq P[\{X > m_1(1+\delta)\} \cup \{Y < m_2(1-\delta)\}] \\ &\leq Pr([X > m_1(1+\delta))] + P(Y < m_2(1-\delta)). \end{aligned}$$

Putting $d = m_1(1+\delta)$ in (3) and $d = m_2(1-\delta)$ in Theorem 2, one gets the bound

$$P[F > (1+\delta)/(1-\delta)] \leq \exp\left[-\frac{m_1\delta^2}{4(m_1+\delta)}\right] + \exp\left[-\frac{m_2\delta^2}{4}\right]. \tag{6}$$

It is well-known that asymptotically as $m_2 \to \infty$, the F-statistic reduces to a chisquare statistic divided by its degrees of freedom. This is also reflected in (6). In particular, we get the inequality

$$\lim\ \sup_{m_2 \to \infty} P[F > (1+\delta)/(1-\delta)] \leq \exp\left[-\frac{m_1\delta^2}{4(m_1+\delta)}\right].$$

3 Non-Central Chisquare

I find in this section, upper and lower tail bounds for non-central chisquare. These upper bounds are not the sharpest ones that one might get, but they are simple enough for potential use in statistics. I begin with the upper bound.

Theorem 3 *Suppose $X \sim \chi^2_p(\lambda)$. Then for $c > 0$,*

$$\begin{aligned} P(X > p+\lambda+c) &\leq \exp\left[-\frac{p}{2}\left\{\frac{c}{p+2\lambda} - \log\left(1+\frac{c}{p+2\lambda}\right)\right\}\right] \\ &\leq \exp\left[-\frac{pc^2}{4(p+2\lambda)(p+2\lambda+c)}\right]. \end{aligned}$$

Proof The second inequality is based on an argument similar to the one used In Theorem 1. For getting the first inequality, I begin with the moment generating function of a non-central chisquare and get

$$P(X > p+\lambda+c) \leq \ \inf f_{0<t<1/2} \exp[-t(p+\lambda+c)] \exp\left(\frac{\lambda t}{1-2t}\right)(1-2t)^{-p/2}. \tag{7}$$

Let $g(t) = -t(p+\lambda+c) + \frac{\lambda t}{1-2t} - \frac{p}{2}\log(1-2t)$. Then $g'(t) = -(p+\lambda+c) + \lambda(1-2t)^{-2} + p(1-2t)^{-1}$ and $g''(t) = 4\lambda(1-2t)^{-3} + 2p(1-2t)^{-2} > 0$. Thus the infimum in (7) is obtained at $t = t_0$, where $g'(t_0) = 0$. Letting $u = (1-2t)^{-1}$ and noting that u is strictly increasing in t, this amounts to solving the equation $\lambda u^2 + pu - (p + \lambda + c) = 0$. The solution is given by $u_0 = \frac{-p+\sqrt{(p+2\lambda)^2+4\lambda c)}}{2\lambda}$. This solution is not too convenient for use in practice. Instead I use the simple inequality $(1 + z)^{1/2} < 1 + \frac{z}{2}$ to get $u_0 < 1 + \frac{c}{p+2\lambda} = u_1$, say. Correspondingly, $t_0 < t_1 = (u_1 - 1)/(2u_1) = \frac{c}{2(p+2\lambda+c)}$. Substitution of this t_1 for t in (7) yields

$$P(X > p + \lambda + c) \leq \exp\left[-\frac{c(p+\lambda+c)}{p+2\lambda+c)} + \frac{\lambda c}{p+2\lambda} + \frac{p}{2}\log\left(1 + \frac{c}{p+2\lambda}\right)\right]. \quad (8)$$

By the inequality, $(p + \lambda + c)/(p + 2\lambda + c) > (p + \lambda)/(p + 2\lambda)$, it follows on simplification, the right-hand side of (8) is bounded above by $\exp\left[-\frac{p}{2}\left(\frac{c}{p+2\lambda} - \log\left(1 + \frac{c}{p+2\lambda}\right)\right)\right]$. This proves the theorem.

The final theorem of this this paper provides a lower tail bound for non-central chisquares.

Theorem 4 *Suppose* $X \sim \chi^2_p(\lambda)$. *Then for* $0 < c < p + \lambda$,

$$P(X < p + \lambda - c) \leq \exp\left[\frac{p}{2}\left\{\frac{c}{p+2\lambda} + \log\left(1 - \frac{c}{p+2\lambda}\right)\right\}\right] \leq \exp\left[-\frac{pc^2}{4(p+2\lambda)^2}\right].$$

Proof Again, the second inequality is obtained by log expansion. To prove the first inequality, we start with

$$P(X < p + \lambda - c) \leq \ \inf_{t<0} \exp\left[-t(p+\lambda-c) + \frac{\lambda t}{1-2t}\right](1-2t)^{-p/2}. \quad (9)$$

Let $g(t) = -t(p + \lambda - c) + \frac{\lambda t}{1-2t} - (p/2)\log(1 - 2t)$. As before, $g(t)$ is minimized at $t = t_0$, where t_0 is a solution of $-(p + \lambda - c) + \lambda(1 - 2t)^{-2} + p(1 - 2t)^{-1} = 0$. Again, writing $u = (1 - 2t)^{-1}$, one needs solving $\lambda u^2 + pu - (p + \lambda - c) = 0$. The solution is given by $u_0 = \frac{-p+\sqrt{(p+2\lambda)^2-4\lambda c}}{2\lambda}$. Now by the inequality $(1 - z)^{1/2} < 1 - \frac{z}{2}$, one gets $u_0 < 1 - \frac{c}{p+2\lambda} = u_1$, say. The corresponding $t_1 = -c/(p + 2\lambda - c)(< 0)$. Substitution of t_1 for t in (8) leads to the inequality

$$P(X < p + \lambda - c) \leq \exp\left[\frac{c(p+\lambda-c)}{p+2\lambda-c} - \frac{\lambda c}{p+2\lambda} + (p/2)\log\left(1 - \frac{c}{p+2\lambda}\right)\right].$$

By the inequality, $(p + \lambda - c)/(p + 2\lambda - c) \leq (p + \lambda)/(p + 2\lambda)$, one gets after simplification,

$$P(X < p + \lambda - c) \leq \exp\left[\frac{p}{2}\left[\frac{c}{p+2\lambda} + \log\left(1 - \frac{c}{p+2\lambda}\right)\right]\right].$$

This proves the theorem.

Remark It is possible to obtain exponential tail-bounds for non-central F-statistic as well. Suppose, for example X and Y are independently distributed with $X \sim \chi^2_{m_1}(\lambda_1)$

and $Y \sim \chi^2_{m_2}(\lambda_2)$. We may recall that $E(X) = m_1 + \lambda_1$ and $E(Y) = m_2 + \lambda_2$. Writing $F = (X/m_1)/(Y/m_2)$, if $d = (1 + \lambda_1/m_1)/(1 + \lambda_2/m_2)(1 + \delta)/(1 - \delta)$, one can as in (6), get the inequality,

$$P(F > d) \leq P(X > (m_1 + \lambda_1)(1 + \delta)) + P(Y < (m_2 + \lambda_2)(1 - \delta)).$$

The exponential bounds are now obtained using $c = (m_1 + \lambda_1)\delta$ in Theorem 3 and $c = (m_2 + \lambda_2)\delta$ in Theorem 4.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Boucheron S, Lugosi G, Massart P (2013) Concentration inequalities: a noasymptotic theory of independence. Oxford University Press, Oxford, England
2. Chernoff H (1952) A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations. Ann Math Stat 23:493–507
3. Hoeffding W (1963) Probability inequalities for sums of bounded random variables. J Am Stat Assoc 58:13–30
4. Rubin H, Sethuraman J (1965) Probabilities of moderate deviations. Sankhya A 27:325–346
5. Wainwright MJ (2019) High-Dimensional statistics: a non-asymptotic viewpoint. Cambridge University Press, Cambridge, England

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.