International Journal of Analysis and Applications

Approximating the Mode of the Non-Central Chi-Squared Distribution

V. Ananiev*, **A. L. Read**

Department of Physics, University of Oslo, Norway

**Corresponding author:* victor.ananyev@gmail.com

Abstract. In this paper we consider the probability density function (pdf) of the non-central χ^2 distribution with arbitrary number of degrees of freedom and non-centrality. For this function we find the approximate location of the maximum and discuss related edge cases of 1 and 2 degrees of freedom. We also use this expression to demonstrate the improved performance of the C++ Boost's implementation of the non-central χ^2 and extend the domain of its applicability.

1. Introduction

Properties of the non-central χ^2 distribution were described before in literature [6–8]. However, the topic of the mode of the non-central χ^2 was significantly underrepresented. We would like to focus on the mode specifically in this paper.

Let $X_1, X_2, ..., X_n$ be normally distributed random variables with unit variance and means $\mu_1, \mu_2, ..., \mu_n$. The sum $X_1^2 + X_2^2 + ... + X_n^2$ follows the non-central χ^2 distribution with $k = n$ degrees of freedom and non-centrality $\lambda = \mu_1^2 + \mu_2^2 + ... + \mu_n^2$. The probability density function of this distribution has a closed form expression:

$$f_{k,\lambda}(x) = \frac{1}{2} \exp^{-\frac{x+\lambda}{2}} \left(\frac{x}{\lambda}\right)^{\frac{k-2}{4}} I_{\frac{k-2}{2}}(\sqrt{\lambda x}) \,, \tag{1.1}$$

where $I_\nu(x)$ is a modified Bessel function of the first kind.

We are interested in the value of x_{mode} that maximizes $f_{k,\lambda}(x)$. Typical shapes of the pdf of the non-central χ^2 distribution are shown in Fig. 1.

(a) Dependency on the number of d.o.f. k.

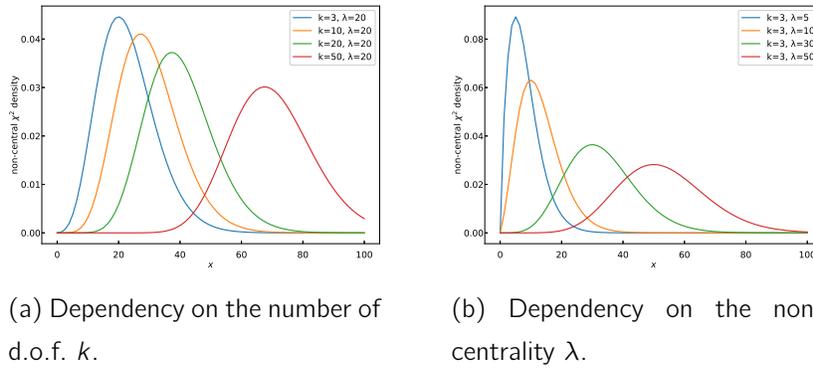(b) Dependency on the non-centrality λ.

Figure 1. Non-central χ^2 distributions and behavior of the mode.

When the number of degrees of freedom k is fixed, we can plot the dependency of the maximum of the pdf as a function of the non-centrality parameter λ, see Fig. 2.
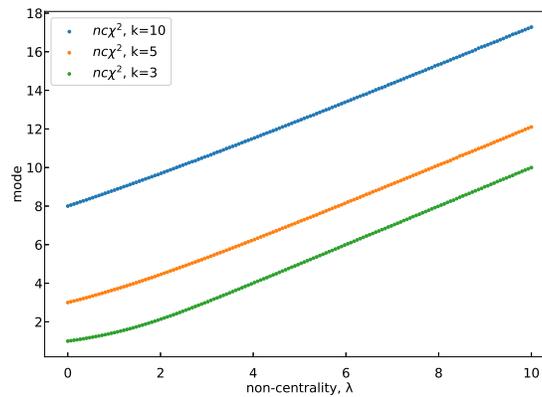

Figure 2. Mode of the non-central χ^2 as a function of the non-centrality parameter λ.

We observe that the bigger λ is the better the mode appears to be approximated with a straight line. The derivation of the line parameters together with the analysis of the edge cases of small number of degrees of freedom, where the mode does not exist, constitute the main results of the paper.

2. Derivation

2.1. **Master equation.** In this section we obtain the transcendental equation (Eq. 2.2) that determines the mode of the non-central χ^2 distribution. We reduce it to the ordinary differential equation (Eq. 2.5), where the non-centrality parameter λ is the argument, and the number of degrees of freedom k is a parameter. Finally, we solve the ODE approximately with a Taylor expansion (Eq. 2.10) and investigate edge cases of 1 and 2 degrees of freedom (Sec. 2.3).

We start by setting the derivative of the density of the non-central χ^2 (Eq. 2.1) to zero. This leads us to the transcendental equation (Eq. 2.2) that determines the mode of the distribution:

$$\frac{d}{dx}\chi^2_{k,\lambda}(x) = \frac{1}{2}\chi^2_{d,\lambda}(x) \cdot \left[-1 + \frac{k-2}{2x} + \sqrt{\frac{\lambda}{x}}\frac{I'_{\frac{k-2}{2}}(\sqrt{\lambda x})}{I_{\frac{k-2}{2}}(\sqrt{\lambda x})}\right] , \tag{2.1}$$

$$\frac{d}{dx}\chi^2_{k,\lambda}(x) = 0 \Rightarrow \sqrt{\lambda x}I'_{\frac{k-2}{2}}(\sqrt{\lambda x}) = (x - \frac{k-2}{2})I_{\frac{k-2}{2}}(\sqrt{\lambda x}) . \tag{2.2}$$

We can eliminate the derivative in Eq. 2.2 by using the differential equation for the modified Bessel function [1, Eq. 10.25.1]:

$$t^2\frac{d^2}{dt^2}I_\nu(t) + t\frac{d}{dt}I_\nu(t) - (t^2 + \nu^2)I_{d,\lambda}(t) = 0 . \tag{2.3}$$

To make use of Eq. 2.3, we need the expression for $I''_{\frac{k-2}{2}}$, therefore, we differentiate Eq. 2.2 by λ. Since the mode depends on the non-centrality λ, we should remember that $x = x(\lambda)$, thus $\frac{dx}{d\lambda} = x'$. The resulting expression for $I''_{\frac{k-2}{2}}$ is as follows:

$$\sqrt{\lambda x}I''_{\frac{k-2}{2}}(\sqrt{\lambda x}) = (x - \frac{k}{2})I'_{\frac{k-2}{2}}(\sqrt{\lambda x}) + \frac{2\sqrt{\lambda x}x'}{x + \lambda x'}I_{\frac{k-2}{2}}(\sqrt{\lambda x}) . \tag{2.4}$$

We substitute $I'_{\frac{k-2}{2}}$ (Eq. 2.2) and $I''_{\frac{k-2}{2}}$ (Eq. 2.4) into the differential equation for the modified Bessel function (Eq. 2.3). We then use the property [1, Eq. 10.29.4] to decrease the order of the derivatives of the modified Bessel functions. Assuming that the Bessel function itself is non-zero at the mode, we arrive to the following differential equation for the mode as a function of the non-centrality parameter λ:

$$\lambda x'(x - k - \lambda + 4) + x(x - k - \lambda + 2) = 0 . \tag{2.5}$$

2.2. **Approximate solution.** We observed that the linear approximation works better with growing λ, thus we introduce the asymptotic parameter $t = \frac{k}{\lambda} << 1$ to build the expansion. We expect the solution to be linear in λ, however the asymptotic expansion of $x(t) = C_0 + C_1 t + ...$ won't provide us with a solution linear in λ. Therefore, we reparametrize $x(t)$ with a new function $y(t) = tx(t)$:

$$t = \frac{k}{\lambda} , \tag{2.6}$$

$$y(t) = tx(t) . \tag{2.7}$$

We obtain the following equation after the reparametrization:

$$-(y't - y)(y - kt - k + 4t) + y(y - kt - k + 2t) = 0 . \tag{2.8}$$

To solve Eq. 2.8, we expand $y(t)$ into the Taylor series by the scale parameter $t = \frac{k}{\lambda}$. We would like to find the linear solution and one extra term that estimates the error. Thus, we cut the series at

the third power of t in order to account for the derivative. After solving algebraic equations for the coefficients near each power of t, we arrive to the resulting approximate expression for the mode:

$$y(t) = C_0 + C_1 t + C_2 t^2 + C_3 t^3 + O(t^4) , \qquad (2.9)$$

$$C_0 = k, \quad C_1 = k - 3, \quad C_2 = \frac{k-3}{2k} , \qquad (2.10)$$

$$\boxed{x_{mode} = \lambda + k - 3 + \frac{k-3}{2\lambda} + O\left(\frac{k^2}{\lambda^2}\right)} . \qquad (2.11)$$

We plot the linear approximation Eq. 2.11 together with the precise numerical solution Fig. 2 in order to verify the approximation is correct, see Fig. 3.
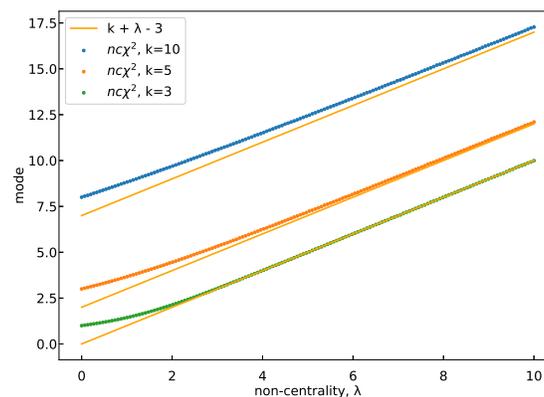

Figure 3. Linear approximation to the mode of the non-central χ^2 compared to the more precise numerical solution as a function of the non-centrality parameter λ.

2.3. **Small number of degrees of freedom.**

2.3.1. *Case $k < 2$.* The asymptotic behavior of the modified Bessel function at $x \to 0$ [1, Eq. 10.30.1] shows that the pdf of the non-central χ^2 diverges, thus it doesn't have a mode:

$$\chi^2_{k,\lambda}(x) \to \frac{1}{2\Gamma(\frac{k}{2})} \frac{1}{(2\lambda)^{\frac{k-2}{2}}} e^{-\frac{\lambda}{2}} \left(\sqrt{\lambda x}\right)^{k-2} , \quad x \to 0 . \qquad (2.12)$$

2.3.2. *Case $k = 2$.* In this case, the pdf at $x = 0$ is finite. If the derivative at $x = 0$ is positive, then the maximum is not there. The expression for the derivative (Eq. 2.13) and its asymptotic behavior at $x \to 0$ (Eq. 2.14) are shown below:

$$\frac{d}{dx}\chi^2_{k,\lambda}(x) = \frac{1}{2}\chi^2_{d,\lambda}(x) \cdot \left[-1 + \sqrt{\frac{\lambda}{x}} \frac{I_{-1}(\sqrt{\lambda x})}{I_0(\sqrt{\lambda x})}\right] , \qquad (2.13)$$

$$\frac{d}{dx}\chi^2_{k,\lambda}(x) \to \frac{1}{2}\chi^2_{d,\lambda}(x) \cdot \left[-1 + \frac{\lambda}{2}\right] , \quad x \to 0 . \qquad (2.14)$$

We observe that when $\lambda > 2$, the pdf of the non-central χ^2 doesn't have its maximum at $x = 0$. In the region $\lambda < 2$, the asymptotic scale $t = \frac{k}{\lambda} > 1$, hence our approximation is inapplicable in this region and we refrain from analysing it.

3. Application

There exist a number of numerical procedures for finding the mode of a distribution [2, Ch. 10]. Some of them require the search region to be specified.

For example, the widely used C++ library Boost [3] identifies the search region based on an initial guess for the mode x_0. Boost iteratively checks regions of the form $[x_0/2, 2x_0]$, $[x_0/2^2, 2^2 x_0]$, etc. When the value of the pdf at both ends of the region becomes smaller than the value at the initial guess point x_0, the algorithm initiates the search for the maximum inside of the region.

At the time of writing, Boost used $x_0 = k + 1$ as the initial guess. We already know, based on the approximate solution (Eq. 2.11), that the chosen guess will undershoot at large non-centrality values λ. Let's estimate λ above which the method will require the second iteration for the region to cover the mode. For this we compare the linear estimate for the location of the mode (Eq. 2.11) to the initial guess x_0 used by Boost:

$$k + \lambda - 3 > 2 \cdot (k + 1) , \tag{3.1}$$

$$\lambda > k + 5 . \tag{3.2}$$

With Eq. 3.2, for any number of d.o.f. k we are able to specify the threshold α, defined by $\lambda = \alpha k$, at which the original initial guess starts undershooting:

$$\alpha > \frac{5}{k} + 1 . \tag{3.3}$$

We see that large k corresponds to small thresholds α. The most conservative estimate for the threshold would be at the smallest k possible: $k = 2$. Thus, $\alpha = 3.5$ is the threshold that approximately works for $k = 2$ and is the overestimated threshold for bigger values of k.

The threshold α (Eq. 3.3) is closely related to the asymptotic scale t (Eq. 2.6) that we used for finding the approximate solution, specifically: $t = \frac{k}{\lambda} = \frac{1}{\alpha}$. For example, the conservative threshold $\alpha \approx 3.5$ corresponds to the asymptotic scale $t \approx 0.25 < 1$. It means that the region where the original guess of Boost undershoots, is, at the same time, the region where our approximate solution for the mode becomes applicable and can be used as a corrected initial guess. However, the fact that we use the conservative threshold may lead to the situation where the original method has already started undershooting but λ is not yet big enough to turn on the corrected regime.

3.1. **Dependency on** λ**.** We fix the threshold to the conservative value $\frac{k}{\lambda} = 0.25$. We then plot the dependency of the run time on the non-centrality λ for a set of d.o.f. k: 2, 15, 50, see Fig. 4. For benchmarking we use the Google benchmark library [4]. The benchmarking script itself became a part of the Boost.Math [5]. Using this script we measure the run time 100 times and use the mean as a central value. The error bar is computed as a standard deviation. We add noise with standard deviation $\sigma = 10^{-6}$ to parameters k and λ to avoid caching effects. The vertical line on the plots shows the threshold where the original initial guess for smaller λ is switched to the corrected value at

bigger λ. Therefore, we expect that both lines coincide below the threshold and the improved solution would lie lower above the threshold. One can notice missing values on the curve representing the original initial guess. The reason for this is the numerical instability of the algorithm in Boost, that has been resolved after we corrected the initial guess.
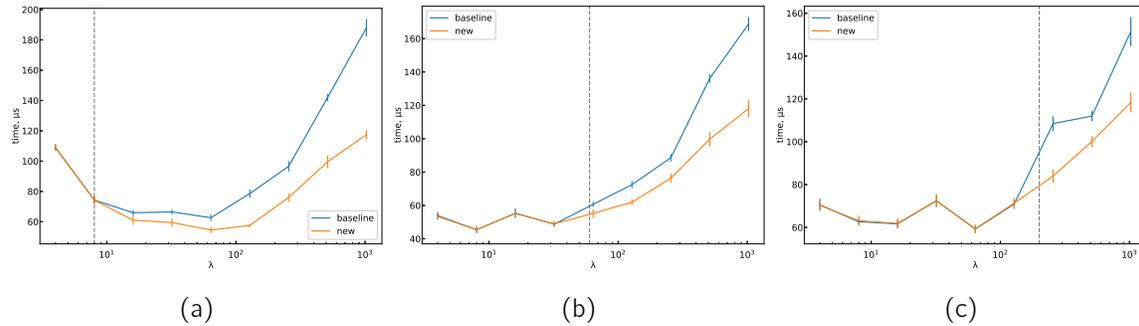

(a)   (b)   (c)

Figure 4. Run time as a function of the non-centrality λ for d.o.f. $k = 2$ (4a), $k = 15$ (4b), $k = 50$ (4c). Vertical line shows the threshold at which the corrected expression replaces the original initial guess.

3.2. **Dependency on d.o.f. (k).** In the set of plots in Fig. 5, we fix the asymptotic scale to values $\frac{k}{\lambda} = 0.25, 0.15, 0.05$ and investigate the dependency of the run time on the number of d.o.f. Since the threshold is fixed, the difference in the run time is caused by the actual position where the original initial guess starts to undershoot, the non-conservative threshold. The farther the fixed threshold is from the non-conservative threshold, the more significant the effect of undershooting at the test point will be. Therefore, we expect the difference in the run time to grow with number of d.o.f, as follows from Eq. 3.3. For each value of the asymptotic scale, in addition to the full plot, we also show a zoomed version that shows the region where both original and improved methods were able to converge (Fig. 5).

4. Conclusion

In this paper we present an approximate expression for the mode of the non-central χ^2 distribution: $x_{mode} \approx k + \lambda - 3$, where k is the number of degrees of freedom and λ is the non-centrality parameter. The approximation is based on an asymptotic expansion and is valid in the region where the scale parameter $\frac{k}{\lambda} << 1$ and where the mode exists $k > 2$. The approximate formula can be used as the initial guess for iterative procedures searching for a precise solution. Run time performance and the domain of applicability of the Boost implementation of the mode search was improved using the presented approximate expression. The improvement became a part of the Boost.Math library [5].
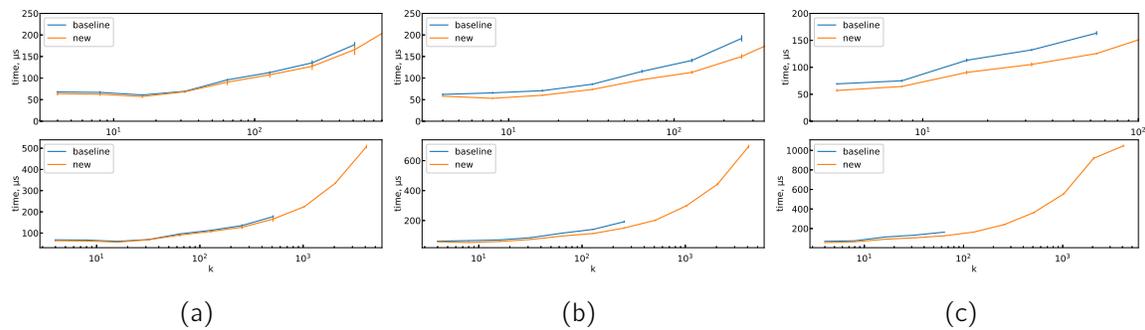
Figure 5. Run time as a function of the number of d.o.f. (k) for the asymptotic scale values $\frac{k}{\lambda} = 0.25$ (5a), $\frac{k}{\lambda} = 0.15$ (5b), $\frac{k}{\lambda} = 0.05$ (5c). The upper plot in each pair shows the zoomed version, focused on the region where both original and improved methods were able to converge.

5. Acknowledgements

References

[1] F.W.J. Olver et al., NIST Digital Library of Mathematical Functions, (2021). http://dlmf.nist.gov/.

[2] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes 3rd Edition, (2007). http://numerical.recipes/.

[3] Boost C++ Libraries, v1.76.0, (2021). https://www.boost.org/.

[4] Google benchmark, (2021). https://github.com/google/benchmark.

[5] Boost Math, (2021). https://github.com/boostorg/math/pull/645.

[6] S. András, Á. Baricz, Properties of the Probability Density Function of the Non-Central Chi-Squared Distribution, J. Math. Anal. Appl. 346 (2008), 395–402. https://doi.org/10.1016/j.jmaa.2008.05.074.

[7] D. Horgan, C.C. Murphy, On the Convergence of the Chi Square and Noncentral Chi Square Distributions to the Normal Distribution, IEEE Commun. Lett. 17 (2013), 2233–2236. https://doi.org/10.1109/LCOMM.2013.111113.131879.

[8] L. Saulis, Asymptotic Expansion for the Distribution and Density Functions of the Quadratic Form of a Stationary Gaussian Process in the Large Deviation Cramer Zone, Nonlinear Anal.: Model. Control. 6 (2001), 87–101. https://doi.org/10.15388/NA.2001.6.1.15218.

[9] S.S. Sawant, D.A. Levin, V. Theofilis, Analytical Prediction of Low-Frequency Fluctuations Inside a One-Dimensional Shock, arXiv:physics.flu-dyn (2021). https://arxiv.org/abs/2101.00664.

[10] V. Pereyra, Iterated Deferred Corrections for Nonlinear Operator Equations, Numer. Math. 10 (1967), 316–323. https://doi.org/10.1007/BF02162030.