STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **6**, September 2018, pp 409–414.
Published online in International Academic Press (www.IAPress.org)

Limit distributions for asymptotically linear statistics with spherical error

Célia Nunes [1,*], Dário Ferreira [1], Sandra S. Ferreira [1], João T. Mexia [2]

[1] *Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Covilhã, Portugal*
[2] *Center of Mathematics and its Applications, Faculty of Science and Technology, New University of Lisbon, Monte da Caparica, Portugal*

Abstract The aim of this work is to obtain general results for the limit distributions of asymptotically linear statistics when the error is spherical, increasing non-centrality. These results apply directly to homoscedastic normal error thus to high precision measurements. We present a numerical example on cylinder volume to illustrate the usefulness of our approach.

Keywords Asymptotic linearity, Limit distributions, Normal distribution, Spherical densities, Cylinder volume.

AMS 2010 subject classifications 62H10, 62H15, 62J99

DOI: 10.19139/soic.v6i3.580

1. Introduction

Let $r(\boldsymbol{u})$ be the spectral radius of the hessian matrix $\underline{\underline{g}}(\boldsymbol{u})$ of $g(\boldsymbol{u})$, then we take

$$\rho_d(\boldsymbol{u}) = Sup\{r(\boldsymbol{v}) : \|\boldsymbol{v} - \boldsymbol{u}\| \leq d\}. \tag{1}$$

If, whatever $d > 0$,

$$Sup\left\{\frac{\rho_d(\boldsymbol{u})}{\|\underline{g}(\boldsymbol{u})\|};\ \|\boldsymbol{u}\| \geq \ell\right\} \underset{\ell\to\infty}{\longrightarrow} 0, \tag{2}$$

with $\underline{g}(.)$ the gradient of $g(.)$, the function $g(.)$ will be asymptotically linear, see [6], [7] and [8].

In this paper we intend to obtain limit distributions for statistics

$$Y = \frac{g(\boldsymbol{a}+\boldsymbol{e}) - g(\boldsymbol{a})}{\|\underline{g}(\boldsymbol{a})\|}, \tag{3}$$

where $g(.)$ is asymptotically linear and the error $\boldsymbol{e}$ has spherical density, when $\|\boldsymbol{a}\| \to \infty$.

Numerical methods may be used to obtain a lower bound for $\|\boldsymbol{a}\|$ such that the distribution of Y is sufficiently near to the limit distribution for this to be used. Namely, this approach was applied in [6] and [8] leading to the establishing of applications domains for the limit distributions. We point out that those domains are defined from lower bounds for $\|\boldsymbol{a}\|$ and not from minimums sample sizes. Besides this, considering an observation $X = \mu + e$ with mean value μ and variance σ^2 will have non-centrality $\frac{\mu^2}{\sigma^2}$ which decreases with σ^2. In this way high non-centrality will be associated to great precision. We thus may associate the application of these limit distributions to high precision observations.

*Correspondence to: Célia Nunes (Email: celian@ubi.pt). Department of Mathematics, University of Beira Interior, Avenida Marquês d'Ávila e Bolama, 6200 Covilhã, Portugal.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

In the next section we will present the required results on spherical densities. This will be followed by the presentation of the key result that the limit density will be the marginal density of $\boldsymbol{e}$ whose components have identical densities. The case in which $\boldsymbol{e}$ is normal is singled out in Section 4. Namely, we will show how to use additional information to overcome $\boldsymbol{e}$ which will have variance-covariance matrix $\sigma^2\boldsymbol{I}_k$ with unknown σ^2. In Section 5 we apply our results to a numerical study considering the cylinder volume. Finally we present some concluding remarks.

2. Spherical densities

Spherical densities, $f(.)$, are such that

$$f(\boldsymbol{x}) = h(\|\boldsymbol{x}\|),$$

for any nonnegative function $h(.)$, see e. g. [1], [3] and [9]. So, we can establish the following proposition.

Proposition 1
If $f(\boldsymbol{x})$ is spherical

1. is invariant for orthogonal transformations;
2. its marginal densities, $\ddot{f}(.)$, are identical;
3. if it has a dispersion parameter γ so that

$$f(\boldsymbol{x}|\gamma) = \frac{1}{\gamma^k} f(\boldsymbol{x}),$$

where $f(\boldsymbol{x}) = f(\boldsymbol{x}|1)$, and $\boldsymbol{a}'\boldsymbol{X}$ will have density $\ddot{f}(.|\ \|\boldsymbol{a}\|\gamma)$, whenever $\boldsymbol{X}$ has density $f(.|\gamma)$;
4. the marginal densities and $f(.|\gamma)$ are symmetrical.

Proof
Let $\boldsymbol{X}$ have spherical density $f(.)$. Then, with $\boldsymbol{P}$ orthogonal,

$$\boldsymbol{X}^{\bullet} = \boldsymbol{P}\boldsymbol{X}$$

will have density

$$f(\boldsymbol{x}^{\bullet}) = h(\|\boldsymbol{x}^{\bullet}\|),$$

since the jacobian of this transformation is equal to one, so 1. is established.

Let $\boldsymbol{P}_i$ be the orthogonal matrix whose first row has all null elements, except the i-th which is equal to 1, $i = 1, ..., k$. Then $\boldsymbol{X}^{\bullet}{}_i = \boldsymbol{P}_i\boldsymbol{X}$ will have the same density than $\boldsymbol{X}$ and its first marginal density will be the i-th marginal of $f(.)$, $i = 1, ..., k$. Thus all marginal of $f(.)$ will be identical and 2. is established.

Next, let $P(\boldsymbol{a})$ be the orthogonal matrix whose first row vector is $\frac{1}{\|\boldsymbol{a}\|}\boldsymbol{a}$. Thus $\boldsymbol{a}'\boldsymbol{e}$ will be the product by $\|\boldsymbol{a}\|$ of the first component of $P(\boldsymbol{a})\boldsymbol{e}$. This first component has density $\ddot{f}(.|\gamma)$, the marginal density of $f(.|\gamma)$. Since γ is a dispersion parameter, the density of $\boldsymbol{a}'\boldsymbol{X}$ will be $\ddot{f}(.|\ \|\boldsymbol{a}\|\gamma)$.

The last part of the thesis follows from $-\boldsymbol{I}_k$ being an orthogonal matrix. □

3. Limit distributions

We will take the statistics

$$Y = \frac{g(\boldsymbol{a}+\boldsymbol{e}) - g(\boldsymbol{a})}{\|\underline{g}(\boldsymbol{a})\|},$$

and

$$Z = \frac{(\underline{g}(\boldsymbol{a}))'\boldsymbol{e}}{\|\underline{g}(\boldsymbol{a})\|},$$

whatever the random vector $\boldsymbol{e}$. With F_L the distribution of L, we have

$$F_Y \underset{\|\boldsymbol{a}\|\to\infty}{\overset{u}{\longrightarrow}} F_Z, \tag{4}$$

where $\overset{u}{\longrightarrow}$ stands for uniform convergence, whenever F_Z does not depend on

$$\boldsymbol{b} = \frac{1}{\|\underline{g}(\boldsymbol{a})\|}\underline{g}(\boldsymbol{a})$$

(as long as it has norm 1), see [6].

As we saw in the previous section, if $\boldsymbol{e}$ has spherical density, the density f_Z of Z will be $\ddot{f}(.)$, which corresponds to the marginal density of f. If there is a dispersion parameter the density will be $\ddot{f}(.|\gamma)$.

We thus establish the following theorem.

Theorem 1
If $g(.)$ is asymptotically linear and $\boldsymbol{e}$ has spherical density the limit density of Y, when $\|\boldsymbol{a}\| \to \infty$, will be the density of the components of $\boldsymbol{e}$.

4. Normal case

Let us put $\boldsymbol{K} \sim N(\boldsymbol{\eta}, \sigma^2\boldsymbol{V})$ to indicate that $\boldsymbol{K}$ is normal with mean vector $\boldsymbol{\eta}$ and variance-covariance matrix $\sigma^2\boldsymbol{V}$. If $\boldsymbol{e} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I}_k)$, its components will have distribution $N(0, \sigma^2)$ so, from Theorem 1, we can conclude that, $N(0, \sigma^2)$ will also be the limit distribution of Y, whatever the asymptotically linear function $g(.)$.

Let us consider an example. We will take $\boldsymbol{a} = \boldsymbol{\mu}$, assuming that $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{e} \sim N(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}_k)$ and the asymptotically linear function

$$g(\boldsymbol{u}) = \|\boldsymbol{u}\|^2.$$

We obtain

$$\begin{cases} \underline{g}(\boldsymbol{u}) = 2\boldsymbol{u} \\ \underline{\underline{g}}(\boldsymbol{u}) = 2\boldsymbol{I}_k \end{cases},$$

and, according to the Theorem 1, the limit density of

$$Y = \frac{\|\boldsymbol{\mu} + \boldsymbol{e}\|^2 - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|}, \tag{5}$$

when $\|\boldsymbol{\mu}\| \to \infty$, will be the density of the components of $\boldsymbol{e}$. So, for large values of $\|\boldsymbol{\mu}\|$,

$$Y = \frac{\|\boldsymbol{X}\|^2 - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|} \overset{\sim}{\approx} N(0, \sigma^2), \tag{6}$$

where $\overset{\sim}{\approx}$ indicates "approximately distributed".

With y a value taken by Y and $\|\boldsymbol{x}\|^2$ the value taken by $\|\boldsymbol{X}\|^2$ we have an equation on $\|\boldsymbol{\mu}\|$ where the solution is

$$\|\widetilde{\boldsymbol{\mu}}\| = -y + \sqrt{y^2 + \|\boldsymbol{x}\|^2}. \tag{7}$$

If we have additional information, for instance that $\sigma^2 = \ddot{\sigma}^2$, we can generate samples

$$\ddot{Y}_1, ..., \ddot{Y}_n \text{ iid} \sim N(0, \ddot{\sigma}^2),$$

where iid indicates independent and identical distributed, and from these obtain the samples

$$\|\widetilde{\boldsymbol{\mu}}\|_1, ..., \|\widetilde{\boldsymbol{\mu}}\|_n.$$

According to the reverse Glivenko-Cantelli theorem, in whatever interval $[q, 1-q]$, with $q \leq p \leq 1-p$,

$$Sup\{|u_{n,p} - u_p|\} \underset{n\to\infty}{\longrightarrow} 0,$$

where $u_{n,p}$ $[u_p]$ is the p-th empirical [exact] quantile for $\|\boldsymbol{\mu}\|$, see [4] and [5].

Another interesting situation is when, instead of additional information, we have $\boldsymbol{X}$ independent of S, where S is the product by σ^2 of a central chi-square with r degrees of freedom, $S \sim \sigma^2\chi^2_r$. Then, see [4], with s the value taken by S, the q-th quantile for the distribution induced by s/χ^2_r for σ^2 is

$$\sigma^2_q = \frac{s}{\chi_{r,1-q}}, \tag{8}$$

where $\chi_{r,1-q}$ denote the $(1-q)$-th quantile for the distribution of χ^2_r.

Moreover, we can replace the expression for $\|\widetilde{\boldsymbol{\mu}}\|$ by

$$\begin{aligned} \widetilde{\delta}^{1/2} &= -y\sqrt{\frac{w}{s}} + \sqrt{y^2\frac{w}{s} + \|\boldsymbol{x}\|^2\frac{w}{s}} \\ &= \sqrt{\frac{w}{s}}\left(-y + \sqrt{y^2 + \|\boldsymbol{x}\|^2}\right), \end{aligned} \tag{9}$$

where w is a value taken by a χ^2_r and $\widetilde{\delta}$ will be a simulated value for the non-centrality parameter

$$\delta = \frac{\|\boldsymbol{\mu}\|^2}{\sigma^2}.$$

The q-th quantile of $\widetilde{\delta}^{1/2}$ will be given by

$$\widetilde{\delta}^{1/2}_q = \sqrt{\left(\frac{w}{s}\right)_q}\left(-y + \sqrt{y^2 + \|\boldsymbol{x}\|^2}\right), \tag{10}$$

where $\left(\frac{s}{w}\right)_q$ denotes the q-th quantile for σ^2. So we can conclude that $\widetilde{\delta}^{1/2}_q$ decreases with $\frac{s}{w}$.

We can also use the reverse Glivenko-Cantelli theorem to obtain confidence intervals for $\delta^{1/2}$ and δ. These intervals can be used to test, through duality, the hypothesis

$$H_0 : \delta = \delta_0.$$

Namely we may be interested in certain applications for STATIS methodology, see e.g. [10], on testing H_0 against

$$H_1 : \delta > \delta_0$$

since only when H_0 is rejected we can be confident in certain model formulation applying.

5. Numerical example: Cylinder volume

In this section we will apply the proposed methodology to the cylinder volume, see [2] and [8]. Now, the asymptotically linear function involved is

$$g(\boldsymbol{u}) = \frac{\pi}{4}u_1^2u_2,$$

that corresponds to the volume of a cylinder with diameter u_1 and height u_2. So we have

$$\underline{g}(\boldsymbol{u}) = \frac{\pi}{4}\begin{bmatrix} 2u_1u_2 \\ u_1^2 \end{bmatrix}$$

and

$$\underline{\underline{g}}(\boldsymbol{u}) = \frac{\pi}{2}\begin{bmatrix} u_2 & u_1 \\ u_1 & 0 \end{bmatrix}.$$

Considering $\boldsymbol{e} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_2)$ we obtain

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{e} \sim N(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I}_2)$$

and, for large values of $\|\boldsymbol{\mu}\|$,

$$Y = \frac{\frac{\pi}{4}X_1^2 X_2 - \frac{\pi}{4}\mu_1^2\mu_2}{\frac{\pi}{4}\sqrt{4\mu_1^2\mu_2^2 + \mu_1^4}} = \frac{X_1^2 X_2 - \mu_1^2\mu_2}{\sqrt{4\mu_1^2\mu_2^2 + \mu_1^4}} \overset{\sim}{\approx} N(0, \sigma^2), \tag{11}$$

where X_1, X_2 are the components of $\boldsymbol{X}$ and μ_1, μ_2 the components of $\boldsymbol{\mu}$.

We will consider the data used in Nunes et al. [8]. In this research the authors generated samples with size 30 using *R* software, assuming the diameters and heights to be normal distributed with mean values 2 and 4, respectively, and standard deviation 0.01. The results are presented in Tables 1 and 2. The corresponding volumes are presented in Table 3 and the values of Y in Table 4.

Table 1. Values of diameters

1.998453	1.988018	2.009190	2.011452	2.007588	2.015884
2.000025	1.985956	2.007079	2.001321	2.003486	1.989625
1.996678	2.008400	1.990175	1.994048	2.008467	1.999402
1.997339	1.995141	1.990827	1.997725	1.996254	2.001284
2.016228	2.008911	2.009408	1.996999	2.015621	2.022904

Table 2. Values of heights

3.993445	4.002777	4.007956	3.992823	3.990726	3.998639
3.985255	3.997894	4.002110	4.009253	3.987402	3.989987
3.979206	4.015282	3.988615	4.010963	3.997377	3.994258
3.991256	3.993969	3.996335	3.996736	4.018685	3.996539
3.999213	3.987138	3.982822	3.982265	3.996033	3.998660

Table 3. Cylinders volumes

12.52638	12.42487	12.70734	12.68789	12.63255	12.76243
12.52036	12.38398	12.66216	12.61209	12.57050	12.40520
12.45956	12.72057	12.40780	12.52592	12.66469	12.54082
12.50555	12.48653	12.43995	12.52757	12.57782	12.57162
12.76861	12.63783	12.63040	12.47314	12.75078	12.85154

Table 4. Values of Y

-0.003088	-0.010924	0.010883	0.009381	0.005109	0.015136
-0.003552	-0.014080	0.007395	0.003529	0.000319	-0.012443
-0.008247	0.011904	-0.012242	-0.003123	0.007590	-0.001972
-0.004695	-0.006164	-0.009759	-0.002996	0.000884	0.000406
0.015613	0.005517	0.004943	-0.007198	0.014237	0.022015

The *p-value* of the *Kolmogorov-Smirnov* test for normality, considering null mean value and variance $\sigma^2 = 0.01^2$, was 0.9052. So we don't reject the hypothesis of normality of Y for the usual levels of significance.

Taking the value $y = 0.022015$ of Y (randomly selected) we obtained

$$\|\widetilde{\boldsymbol{\mu}}\| = 4.925884.$$

In this case, $S \sim \chi^2_{58}$ and $s = 0.005$.

The quantiles of $\widetilde{\delta}^{1/2}$, $\widetilde{\delta}_q^{1/2}$, are presented in Table 5.

Table 5. Quantiles of $\widetilde{\delta}^{1/2}$

Values of q	0.9	0.95	0.99
$\widetilde{\delta}_q^{1/2}$	645.8365	610.4035	591.7619

The high values obtained for these quantiles are due to the fact that we worked with small variance. So we can conclude that we are in a non-central situation in which the limit distributions, obtained through the asymptotic linearity, apply.

6. Final Remarks

With this research it was shown that the general results of the limit distributions apply when the error has spherical density, namely if it is normal. The numerical application on cylinder volume illustrates the usefulness of our approach. Moreover the approach presented for the normal case can be applied to Wishart matrices. Namely, we intend to publish results on limit distributions for these matrices, their trace and determinant. Others applications may be found in [2], [6] and [8].

Acknowledgement

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID/MAT/00212/2013 and UID/MAT/00297/2013.

REFERENCES

1. C. Fernandez, J. Osiewalski, M. Steel, *Modelling & inference with ν-spherical distributions*. J. Amer. Statist. Assoc., vol. 90, no. 432, pp 1331 – 1340, 1995.
2. D. Ferreira, S.S Ferreira, C. Nunes, L. Ramos and J.T. Mexia, *Approximate normality of low degree polynomials in normal independent variables*. Far East J. Math. Sci., vol. 68, no. 2, pp. 287 - 296, 2012.
3. D. Ferreira, S.S. Ferreira, C. Nunes and S. Incio, *Inducing pivot variables and non-centrality parameters in elliptical distributions*. 11th International Conference on Numerical Analysis and Applied Mathematics. AIP Conf. Proc., vol. 1558, pp. 833-836, 2013.
4. D. Ferreira, S.S. Ferreira, C. Nunes, M. Fonseca and J.T. Mexia, *Chisquared and Related Inducing Pivot Variables: An Application to Orthogonal Mixed Models*. Communications in Statistics - Theory and Methods, 2013. doi:10.1080/03610926.2013.770532
5. M. Loève, *Probability Theory*, 4th ed., New York, NY: Springer, 1977.
6. J.T. Mexia and M.M. Oliveira, *Asymptotic linearity and limit distributions, approximations*. J. Stat. Plan. Infer., vol. 140, no. 2, pp. 353-357, 2010.
7. J.T. Mexia, C. Nunes and M.M. Oliveira, *Multivariate Application Domains for the Delta Method*. In Proceedings of 9th International Conference on Numerical Analysis and Applied Mathematics. AIP Conf. Proc. vol. 1389, pp. 1486-1489, 2011.
8. C. Nunes, M.M. Oliveira and J.T. Mexia, *Application domains for the Delta method*. Statistics, vol. 47, no. (2), pp. 317-328, 2013.
9. R.N. Rattihalli and P.Y. Patil, *Generalized υ-Spherical Densities*. Communications in Statistics - Theory and Methods, vol. 39, no. 19, pp. 3568-3583, 2013.
10. M.M. Oliveira, *Modelação de séries emparelhadas de estudos com estrutura comum*. Unpublished PhD thesis, Universidade de Évora, 2002.