Livadiotis *Journal of Statistical Distributions and Applications* 2014, **1**:4
http://www.jsdajournal.com/content/1/1/4

Journal of Statistical Distributions and Applications
a SpringerOpen Journal

RESEARCH Open Access

Chi-*p* distribution: characterization of the goodness of the fitting using L^p norms

George Livadiotis

Correspondence: glivadiotis@swri.edu
Southwest Research Institute, San Antonio, TX, USA

Abstract

This paper derives (1) the Chi-*p* distribution, i.e., the analog of Chi-square distribution but for datasets that follow the General Gaussian distribution of shape *p*, and (2) develops the statistical test for characterizing the goodness of the fitting with L^p norms. It is shown that the statistical test has double role when the fitting method is induced by the L^p norms: For given the shape parameter *p*, the test is rated based on the estimated p-value. Then, a convenient characterization of the fitting rate is developed. In addition, for an unknown shape parameter and if the fitting is expected to be good, then those L^p norms that correspond to unlikely p-values are rejected with a preference to the norms that maximized the p-value. The statistical test methodology is followed by an illuminating application.

1. Introduction

The fitting of a given dataset $\{f_i \pm \sigma_{f_i}\}_{i=1}^N$ to the values $\{V_i\}_{i=1}^N$ of a statistical model $V(X;\alpha)$ in the domain $X \in D_x \subseteq \mathfrak{R}$ (McCullagh 2002; Adèr 2008), involves finding the optimal parameter value $\alpha = \alpha^*$ in $\alpha \in D_\alpha \subseteq \mathfrak{R}$ that minimizes the total square deviations (TSD) between model and data,

$$TSD(\alpha)^2 = \sum_{i=1}^N \sigma_{f_i}^{-2}[f_i - V(x_i;\alpha)]^2, \tag{1}$$

where the inverse of the variance of the data measurements $\left\{w_i = \sigma_{f_i}^{-2}\right\}_{i=1}^N$ is weighting the summation. The deviations may be also defined using the total absolute deviations (TAD),

$$TAD(\alpha) = \sum_{i=1}^N \sigma_{f_i}^{-2}|f_i - V(x_i;\alpha)|. \tag{2}$$

A class of generalized fitting methods has been considered by Livadiotis (2007), using the metric induced by the p-norms L^p, $p \geq 1$, that denotes a complete normalized vector space with finite Lebesgue integral. The total deviations (TD) are now defined by

$$TD(\alpha)^p = \sum_{i=1}^N \sigma_{f_i}^{-p}|f_i - V(x_i;\alpha)|^p. \tag{3}$$

The least square method based on the Euclidean norm, $p = 2$, and the least absolute deviations method based on the "Taxicab" norm, $p = 1$, are some cases of the general fitting methods based on the L^p-norms (see Burden and Faires 1993; for more applications of the fitting methods based on L^p norms, see: Sengupta 1984; Livadiotis and Moussas 2007; Livadiotis 2008; 2012; for fitting methods based on other effect sizes e.g., correlation, see: Livadiotis and McComas 2013a).

The goodness of the least square fitting is typically measured using the estimated Chi-square value, that is the least squared value, $\chi^2_{\text{est}} = TSD(\alpha^*)^2$. Then, this χ^2_{est} is compared with the Chi-square distribution, to examine whether such a value is frequent or not (see next sections). However, this test can apply only to datasets $\{f_i \pm \sigma_{f_i}\}_{i=1}^N$ that follow the normal distribution $f_i \sim N\left(\mu_{f_i}, \sigma_{f_i}\right)$. There is no similar test for cases where the dataset follows the General Gaussian distribution of shape p, $f_i \sim GG\left(\mu_{f_i}, \sigma_{f_i}, p\right)$ (see Section 2 and Appendix A). Livadiotis (2012) showed the connection between the fitting with L^p norms, as in Eq. (3), and datasets that follow the General Gaussian distributions, $f_i \sim GG\left(\mu_{f_i}, \sigma_{f_i}, p\right)$.

The purpose of this paper is to (1) construct the formulation of the Chi-p distribution, the analog of Chi-square distribution but for datasets that follow the General Gaussian distribution of shape p, and (2) develop the statistical test for characterizing the goodness of the fitting with L^p norms, which corresponds to datasets that follow the General Gaussian distribution of shape p. Therefore, in Section 2, we revisit the Chi-square derivation, and following similar steps, we construct the Chi-p distribution. In Section 3, we develop the statistical test for characterizing the goodness of the fitting with L^p norms, using the Chi-p distribution and the p-value. In Section 4, we provide an application of the statistical test. Finally, in Section 5, we summarize the conclusions. Appendix A briefly describes the General Gaussian distribution, while Appendix B shows the mathematical derivation of the surface of the sphere of higher dimensions in L^p space.

2. Chi-*p* distribution

We first revisit the derivation of Chi-square distribution. This distribution is necessary to test the goodness of fitting of measurements that follow the Gaussian distribution. This test applies to datasets $\{x_i \pm \sigma_{x_i}\}_{i=1}^N$ that follow the normal distribution $x_i \sim N\left(\mu_{x_i}, \sigma_{xi}\right)$. The Chi-square is given by

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{x_i - \mu_{x_i}}{\sigma_{x_i}}\right)^2, \tag{4}$$

that is the sum of squares of N independent random variables. The distribution of this sum is given by

$$P(X;N)dX = \frac{2^{-\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)} e^{-\frac{1}{2}X} X^{\frac{N}{2}-1} dX, \quad \text{with} \quad X \equiv \chi^2. \tag{5}$$

The estimated value of the Chi-square for a fitting is given by the minimum at $\alpha = \alpha^*$ of the function $\chi^2(\alpha) = TSD(\alpha)^2$, as shown in Eq. (1) (least squares). Considering that the Chi-square minimum, $\chi^2(\alpha^*)$, is equivalently referred to all the $M = N$-1 degrees of freedom (for N number of data), then each of them contributes to this minimum by a factor of $\frac{1}{M}\chi^2(\alpha^*)$. This is the estimated value of the reduced Chi-square. For multi-parametrical fitting (Livadiotis 2007) of n free parameters, the degrees of freedom are $M = N$-n. In general, the Chi-square distribution in Eq. (5) is referred to M degrees of freedom.

For testing the goodness of fitting of measurements $\{x_i \pm \sigma_{xi}\}_{i=1}^{N}$ that follow the General Gaussian distribution of shape p, $x_i \sim GG(\mu_{xi}, \sigma_{xi}, p)$, we need to construct the Chi-p distribution connected with L^p fitting methods, where the minimization of $\chi^p(\alpha)$ is given by Eq. (3). The General Gaussian distribution of shape p, $f_i \sim GG\left(\mu_{f_i}, \sigma_{f_i}, p\right)$ (Appendix A). This distribution is parameterized by the mean μ, the variance σ, and the shape parameter p,

$$P(x;\mu,\sigma,p)dx = C_p \cdot e^{-\eta_p \cdot \left|\frac{x-\mu}{\sigma}\right|^p} d\left(\frac{x-\mu}{\sigma}\right), \tag{6}$$

where the involved coefficients are

$$C_p = \sqrt{\frac{p \sin\left(\frac{\pi}{p}\right)}{4\pi(p-1)}}, \eta_p = \left[\frac{\sin\left(\frac{\pi}{p}\right)\Gamma\left(\frac{1}{p}\right)^2}{\pi p(p-1)}\right]^{\frac{p}{2}}. \tag{7}$$

Figure 1 depicts the distribution $P\left(z = \frac{x-\mu}{\sigma}; p\right) \equiv P(x;\mu,\sigma,p)$ for various shape parameters p. Note that the normalized coefficient C_p is derived by setting $\int_{-\infty}^{\infty} P(x;\mu,\sigma,p)\,dx = 1$, while the exponential coefficient η_p is derived so that the L^p-normed variance to equal σ^2. The theory of L^p-normed mean and variance was developed by Livadiotis (2012), which for the case of the General Gaussian distribution (6) leads to the following Propositions:

– *Proposition* 1: The L^p-normed mean of the distribution (6) is $<x>_p = \mu$, $\forall\, p \geq 1$.
– *Proposition* 2: The L^p-normed variance of the distribution (6) is $\sigma_p^2 = \sigma^2$, $\forall\, p \geq 1$.

The proofs of the two Propositions are shown in Appendix A.

We continue with the development of the Chi-p distribution. We start with the following Lemma:

– *Lemma* 1: The surface of the N-dimensional sphere of unit radius in L^p space is given by

$$B_{p,N} = p\left[\left(\frac{2}{p}\right)\Gamma\left(\frac{1}{p}\right)\right]^N / \Gamma\left(\frac{N}{p}\right). \tag{8}$$

The proof is shown in Appendix B.

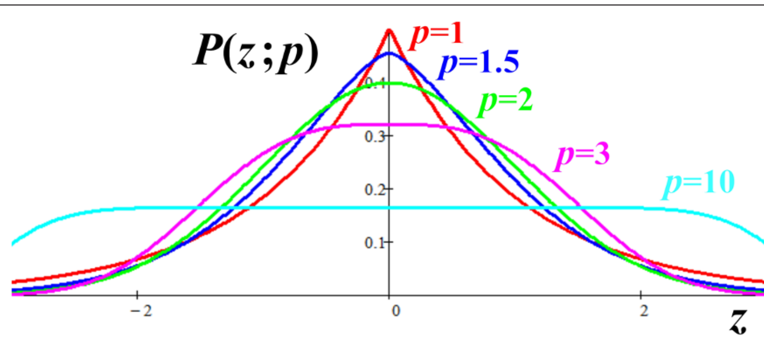

Figure 1 General Gaussian distribution *P(z; p)* for *z* = (*x*-*μ*)/*σ*. This is depicted for various shape parameters p = 1, 1.5, 2, 3, and 10. The larger the value of p, the more flattened the maximum is.

– *Theorem* 1:
The Chi-p is given by the sum of absolute values to the exponent p of N independent random variables,

$$\chi^p = \sum_{i=1}^{N} \left| \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \right|^p. \tag{9}$$

For M degrees of freedom ($M = N\text{-}n$, N number of data, n number of independent variables), the Chi-p distribution is given by

$$P(X; M; p) = \frac{\eta_p^{\frac{M}{p}}}{\Gamma\left(\frac{M}{p}\right)} e^{-\eta_p X} X^{\frac{M}{p}-1}, \tag{10}$$

where the estimated Chi-p value X is given by the minimum at $\alpha = \alpha^*$ of the function $\chi^p(\alpha) = TD(\alpha)^p$, as shown in Eq. (3) (least L^p deviations). Figure 2 plots the Chi-p distribution for various values of the shape parameter p (that correspond to various L^p norms).

– *Proof of Theorem* 1. The distribution of Chi-p can be derived as follows. The normalization of the joint distribution function of all the data is

$$1 = \int_{-\infty}^{+\infty} \prod_{i=1}^{N} \frac{C_p}{\sigma_{x_i}} e^{-\eta_p \left| \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \right|^p} dx_1 \ldots dx_N, \tag{11}$$

where the coefficients (Livadiotis 2012) are given by Eq. (7).

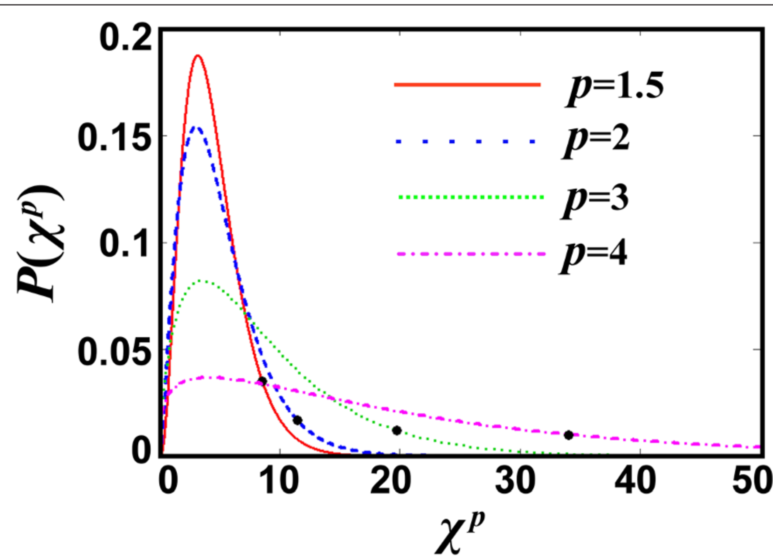

Figure 2 Chi-p distribution function. This is depicted for various norms $p = 1.5$, 2, 3, and 4. The degrees of freedom are $M = 5$. The black points correspond to the estimated values of χ^p for the fitting example in Section 4. Therefore, we observe that by varying the L^p norm, both the Chi-p distribution and the estimated χ^p also vary.

By setting $z_i \equiv \frac{x_i - \mu_x}{\sigma_{x_i}}$, we derive

$$1 = \int_{-\infty}^{+\infty} \prod_{i=1}^{N} C_p \, e^{-\eta_p |z_i|^p} dz_1 ... dz_N = \int_{-\infty}^{+\infty} C_p^N \, e^{-\eta_p \sum_{i=1}^{N} |z_i|^p} dz_1 ... dz_N, \tag{12}$$

that is

$$1 = \int_{\vec{z} \in B_{p,N}} d^{N-1}\Omega_N \cdot \int_0^{+\infty} C_p^N e^{-\eta_p Z^p} Z^{N-1} dZ, \tag{13}$$

where we denote $Z^p \equiv \sum_{i=1}^{N} |z_i|^p$, and $B_{p,N} \equiv \int_{\vec{z} \in B_{p,N}} d^{N-1}\Omega_N$ is the surface of the N-dimensional sphere of unit radius in L^p space (*Lemma* 1), so that

$$1 = \int_0^{+\infty} C_p^N B_{p,N} \, e^{-\eta_p Z^p} Z^{N-1} dZ = \int_0^{+\infty} C_p^N \frac{1}{p} B_{p,N} \, e^{-\eta_p X} X^{\frac{N}{p}-1} dX \equiv \int_0^{+\infty} P(X;N;p) dX,$$

where we have used the identity $C_p^N \frac{1}{p} B_{p,N} = \eta_p^{\frac{N}{p}} / \Gamma\left(\frac{N}{p}\right)$. Hence, we find

$$P(X;N;p) dX = \frac{\eta_p^{\frac{N}{p}}}{\Gamma\left(\frac{N}{p}\right)} e^{-\eta_p X} X^{\frac{N}{p}-1} dX, \quad \text{with} \quad X \equiv \chi^p. \tag{14}$$

In general, for M degrees of freedom, the Chi-p distribution is given by Eq. (10).

3. Statistical test of a fitting

In order to estimate the goodness of the fitting, we minimize the Chi-p, χ^p,

$$\chi^p = \sum_{i=1}^{N} \sigma_{f_i}^{-p} [f_i - V(x_i;\alpha)]^p, \tag{15}$$

similar to the minimization of the Chi-square, χ^2, for the case of the Euclidean norm,

$$\chi^2 = \sum_{i=1}^{N} \sigma_{f_i}^{-2} [f_i - V(x_i;\alpha)]^2. \tag{16}$$

We begin with the established method of Chi-square, and then we will proceed to the generalized method of Chi-p.

The goodness of a fitting can be estimated by the reduced Chi-square value, $\chi^2_{\text{red}} = \frac{1}{M} \chi^2_{\text{test}}$, where $M = N$-1 indicates again the degrees of freedom. The meaning of χ^2_{red} is the portion of χ^2 that corresponds to each of the degrees of freedom, and this has to be ~1 for a good fitting. We can easily understand this, for example, when the given data have equal error σ_f, with $\{f_i \pm \sigma_f\}_{i=1}^N$, i.e., $\sigma_{f_i} = \sigma_f$ for all i = 1,...., N. Then, the optimized model value, $V(x_i; \alpha^*)$, gives the expected value of the data point f_i, so that the variance can be approached by $\sigma_f^2 = \frac{1}{M} \sum_{i=1}^{N} [f_i - V(x_i;\alpha^*]^2$ (sample variance). Hence, the derived Chi-square becomes $\chi^2_{\text{est}} = \sigma_f^{-2} \sum_{i=1}^{N} [f_i - V(x_i;\alpha^*)]^2 = M$, and its reduced value $\chi^2_{\text{red}} = \frac{1}{M} \chi^2_{\text{est}} = 1$. Therefore, a fitting can be characterized as "good" when χ^2_{red} ~1, otherwise there is an overestimation, χ^2_{red} <1, or underestimation, χ^2_{red} >1, of the errors. When the deviations of the data $\{f_i\}_{i=1}^N$ from the model values $\{V(x_i;\alpha)\}_{i=1}^N$ are small, the fitting is expected to be good. However, this characterization is meaningless if the errors of the data $\{\sigma_{fi}\}_{i=1}^{N}$ are either (i) quite larger than their deviations from the model values, i.e., if $\sigma_{fi} >> |f_i - V(x_i; \alpha)|$, or (ii) quite smaller, i.e., if $\sigma_{fi} << |f_i - V(x_i; \alpha)|$ (e.g., see Figure 3). Then, a perfect matching between data and model is useless when the errors of the data are comparably large or small.

Furthermore, a better estimation of the goodness is derived from comparing the calculated χ^2 value and the Chi-square distribution, that is the distribution of all the possible χ^2 values for data with normally distributed errors (parameterized by the degrees of freedom M),

$$P(\chi^2; M)\, d\chi^2 = \frac{2^{-\frac{M}{2}}}{\Gamma\left(\frac{M}{2}\right)} e^{-\frac{1}{2}\chi^2} \left(\chi^2\right)^{\frac{M}{2}-1} d\chi^2, \tag{17}$$

(e.g., see Melissinos 1966). The likelihood of having an χ^2 value equal to or smaller than the estimated value χ^2_{est}, is given by the cumulative distribution

$$P\left(0 \le \chi^2 \le \chi^2_{\text{est}}\right) = \int_0^{\chi^2_{\text{est}}} P(\chi^2; M)\, d\chi^2 = 1 - \frac{\Gamma\left(\frac{1}{2}M; \frac{1}{2}\chi^2_{\text{est}}\right)}{\Gamma\left(\frac{1}{2}M\right)}, \tag{18}$$

where $\Gamma(x; b) = \int_x^{\infty} e^{-X} X^{b-1} dX$ is the incomplete Gamma function. In addition, the likelihood of having an χ^2 value equal to or larger than the estimated value χ^2_{est}, is given by the complementary cumulative distribution

$$P\left(\chi^2_{\text{est}} \le \chi^2 < \infty\right) = \int_{\chi^2_{\text{est}}}^{\infty} P(\chi^2; M)\, d\chi^2 = \frac{\Gamma\left(\frac{1}{2}M; \frac{1}{2}\chi^2_{\text{est}}\right)}{\Gamma\left(\frac{1}{2}M\right)}. \tag{19}$$

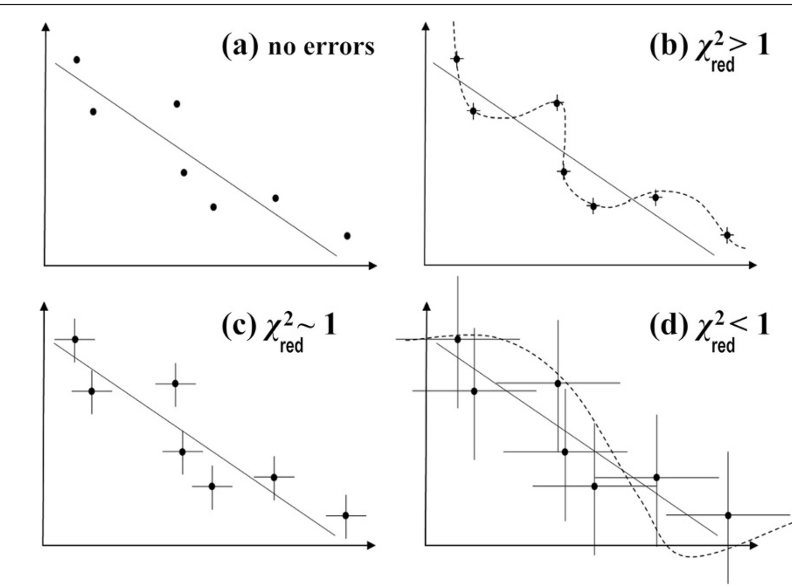

Figure 3 Possible values of the reduced chi-square and their meaning. (a) Seven data points are fitted by a statistical model, here a straight line. **(b)** When the errors are too small (underestimation), the calculated reduced Chi-square is $\chi^2_{\text{red}} > 1$, and the fitted line does not pass through the data points or their error lines. Other more complicated curve can fit better the data (dash line). **(c)** In the case where the errors are similar to the deviations of the data points from the model, the reduced Chi-square is $\chi^2_{\text{red}} \sim 1$, and the fitting is good. **(d)** Finally, when the errors are too large (overestimation), the reduced Chi-square is $\chi^2_{\text{red}} < 1$. In this case, the fitted line does pass through the data points or their error lines, but the curves of any other model can also pass through these, leading to good fitting; hence, the rate of the fitting is meaningless.

The probability of having a result χ^2 larger than the estimated value χ^2_{est}, defines the p-value that equals $P\left(\chi^2_{\mathrm{est}} \le \chi^2 < \infty\right)$. The larger the p-value, the better the fitting is (e.g., Melissinos 1966). However, the p-value test fails when p > 0.5. Indeed, p-values larger than 0.5 correspond to $\chi^2_{\mathrm{est}} < M$ or $\chi^2_{\mathrm{red}} < 1$. Even larger p-values, up to p = 1, correspond to even smaller Chi-squares, down to $\chi^2_{\mathrm{red}} \sim 0$. Thus, an increasing p-value above the threshold of 0.5 cannot lead to a better fitting but to a worse, similar to the indication $\chi^2_{\mathrm{red}} < 1$. For this reason, we use the "p-value of the extremes". According to this, the probability of taking a result χ^2, more extreme than the observed value is given by the p-value that equals the minimum between $P\left(0 \le \chi^2 \le \chi^2_{\mathrm{est}}\right)$ and $P\left(\chi^2_{\mathrm{est}} \le \chi^2 < \infty\right)$, i.e.,

$$\mathsf{p}\text{–value} = \min\left[\frac{\Gamma\left(\frac{1}{2}M; \frac{1}{2}\chi^2_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{2}M\right)} \quad , \quad 1 - \frac{\Gamma\left(\frac{1}{2}M; \frac{1}{2}\chi^2_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{2}M\right)}\right], \tag{20}$$

(see some applications in Livadiotis and McComas 2013b; Frisch et al. 2013; Funsten et al. 2013). Note that the maximum p-value is 0.5, and this corresponds to the estimated Chi-square $\chi^2_{\mathrm{est},1/2} \cong M - \frac{2}{3}$. This is larger than the Chi-square that maximizes the distribution, $\chi^2_{\mathrm{est,\,max}} = M - 2$. Hence, $\chi^2_{\mathrm{est,\,max}} < \chi^2_{\mathrm{est},1/2}$, i.e., the Chi-square that corresponds to p-value = 0.5, is located always at the right of the maximum.

The statistical test of the fitting for the evaluation of its goodness comes from the null hypothesis that the given data are described by the fitted statistical model. If the derived p-value is smaller than the significance level of ~0.05, then the hypothesis is typically rejected, and the hypothesis that the data are described by the examined statistical model is characterized as unlikely.

A convenient rate for a statistical test is to give more detailed characterization than "likely" when p-value > 0.05, or "unlikely" when p-value < 0.05. For this reason, it is necessary to ascribe an 1–1 relation between the domain of p-values $\{\mathsf{p} \in [0, 0.5]\}$ and the range of a rating values $\{\mathsf{T} \in [-1,\ 1]\}$, with the correspondence: 1) Impossible $\mathsf{p} = 0 \leftrightarrow \mathsf{T} = -1$; 2) indefinite $\mathsf{p} = 0.05 \leftrightarrow \mathsf{T} = 0$; 3) certain $\mathsf{p} = 0.5 \leftrightarrow \mathsf{T} = 1$. Choosing a power-law function, $(\mathsf{T} + 1)/2 = (\mathsf{p}/\mathsf{p}_0)^{\gamma}$, we find $\mathsf{p}_0 = 0.5$ and γ = log 2, i.e.,

$$(\mathsf{T} + 1)/2 = (2\mathsf{p})^{\log 2}. \tag{21}$$

We can easily now characterize the testing rates by a linear separation of the values of T, as shown in Table 1.

Table 1 Testing rates and characterizations

p-value	Rate T	Characterization
p ~ 0	T ~ −1	Impossible
0 < p <0.005	−1 < T < −0.5	Highly unlikely
0.005 ≤ p <0.05	−0.5 ≤ T <0	Unlikely
0.05 ≤ p <0.19	0 ≤ T <0.5	Likely
0.19 ≤ p <0.5	0.5 ≤ T <1	Highly likely
p ~ 0.5	T ~ 1	Certain

In the case of data that follow the General Gaussian distribution of shape p, the derived p-value is dependent on the shape p. Indeed, we have

$$P(\chi^p;M;p)\,d\chi^p = \frac{\eta_p^{\frac{M}{p}}}{\Gamma\left(\frac{M}{p}\right)}\cdot\left(\chi^p\right)^{\frac{M}{p}-1}\cdot e^{-\eta_p\chi^p}\,d\chi^p, \quad (22)$$

and

$$P\left(0\le\chi^p\le\chi^p_{\mathrm{est}}\right) = \int_0^{\chi^p_{\mathrm{est}}} P(\chi^p;M;p)\,d\chi^p = 1-\frac{\Gamma\left(\frac{1}{p}M;\eta_p\chi^p_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{p}M\right)}, \quad (23)$$

$$P\left(\chi^p_{\mathrm{est}}\le\chi^p<\infty\right) = \int_{\chi^p_{\mathrm{est}}}^{\infty} P(\chi^p;M;p)\,d\chi^p = \frac{\Gamma\left(\frac{1}{p}M;\eta_p\chi^p_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{p}M\right)}, \quad (24)$$

and the p-value that equals the minimum between $P\left(0\le\chi^p\le\chi^p_{\mathrm{est}}\right)$ and $P\left(\chi^p_{\mathrm{est}}\le\chi^p<\infty\right)$, i.e.,

$$\text{p-value} = \min\left[\frac{\Gamma\left(\frac{1}{p}M;\eta_p\chi^p_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{p}M\right)}\ ,\ 1-\frac{\Gamma\left(\frac{1}{p}M;\eta_p\chi^p_{\mathrm{est}}\right)}{\Gamma\left(\frac{1}{p}M\right)}\right]. \quad (25)$$

Note that the maximum p-value = 0.5 corresponds to the estimated Chi-square $\chi^p_{\mathrm{est},1/2}\cong\frac{1}{p\eta_p}M-\frac{1}{3\eta_p}$. This is larger than the Chi-square that maximizes the distribution, $\chi^p_{\mathrm{est},1/2}\cong\frac{1}{p\eta_p}M-\frac{1}{\eta_p}$. Hence, again we find $\chi^p_{\mathrm{est,\,max}}<\chi^p_{\mathrm{est},1/2}$.

The statistical test has double role in the case of L^p norms. If the shape parameter p is known, then the test can be rated by deriving the p-value and according to Table 1. If the shape parameter is unknown and the fitting is expected to be good, then all the shape values p that correspond to unlikely p-values can be rejected. In fact, the largest p-value corresponds to the most-likely shape parameter p of the examined data. These are shown in the following applications.

4. Applications

Table 2 contains a dataset of observations of the ratio of the umbral area to the whole sunspot area, $\{f_i\}_{i=1}^N$, $N = 6$ (Edwards 1957). Assuming that each of them follows a General Gaussian distribution about their mean, $f_i \sim GG(\mu_i, \sigma_i, p)$, what is the likelihood of these measurements to represent a constant physical quantity? Let this constant be indicated by μ_p, which can be derived from the fitting of $\left\{f_i \pm \sigma_{f_i}\right\}_{i=1}^N$, and thus, it is typically depended on the p-norm. However, different values of the p-norm lead to

Table 2 Testing rates and characterizations

Heliographic latitude (degrees)	Ratio of umbral area to whole sunspot area f_i (%)	Standard deviation σ_{f_i} (%)
0-5	0.1708	0.0053
5-10	0.1677	0.0019
10-15	0.1624	0.0016
15-20	0.1610	0.0019
20-25	0.1594	0.0026
>25	0.1627	0.0040

different estimated values of the Chi-p, χ^p_{est}. Thus, the p-value of the null hypothesis ($\mathrm{H_o}$) depends also on the p-norm.

We apply a statistical test to examine whether the data of the sunspot area ratios are dependent with heliolatitude on not. Therefore, the null hypothesis is that the dataset is described by the statistical model of constant value, i.e., $\{V(x_i;\alpha)=\alpha\}_{i=1}^{N}$. We construct and minimize the Chi-p, given by

$$\chi^p(\alpha)=\sum_{i=1}^{N}\left|\frac{f_i-\alpha}{\sigma_{f_i}}\right|^p, \tag{26}$$

so that the L^p-mean value $\alpha_p=\alpha_p(p)$ is implicitly given by

$$\sum_{i=1}^{N}\left|\frac{f_i-\alpha_p}{\sigma_{f_i}}\right|^p sign\left(f_i-\alpha_p\right)=0, \tag{27}$$

and the estimated Chi-p is

$$\chi^p(p)=\sum_{i=1}^{N}\left|\frac{f_i-\alpha_p}{\sigma_{f_i}}\right|^p. \tag{28}$$

Figure 4(a) shows the six data points co-plotted with four values of α_p, that correspond to $p\to 1$, $p\to\infty$, and the two shape parameter values p_1, p_2 for which the p-value is equal to 0.05. The whole diagram of $\alpha_p=\alpha_p(p)$ is shown in Figure 4(b) and the p-value as a function of p is shown in Figure 4(c).

We observe that the function α_p is monotonically increasing converging to some constant value for $p\to\infty$. The corresponding mean value, α_∞, is given by

$$\alpha_\infty=\frac{\frac{x_{\min}}{\sigma_{x_{\min}}}+\frac{x_{\max}}{\sigma_{x_{\max}}}}{\frac{1}{\sigma_{x_{\min}}}+\frac{1}{\sigma_{x_{\max}}}}\cong 0.166. \tag{29}$$

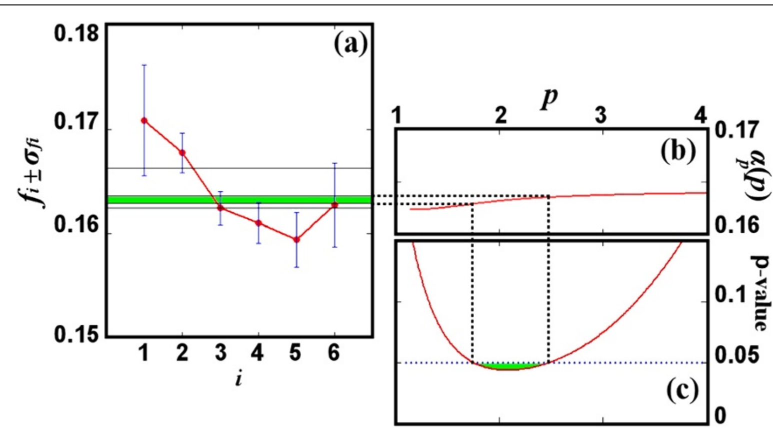

Figure 4 Statistical test for the rate of fitting based on L^p norms. (a) The dataset of Table 2 is co-plotted with four values of α_p, that correspond to $p\to 1$, $p\to\infty$, and the two shape parameter values $p_1\sim 1.7$ and $p_2\sim 2.5$ for which the p-value is equal to 0.05. **(b)** The diagram of L^p mean values, $\alpha_p=\alpha_p(p)$. **(c)** The p-value as a function of p. We observe that for the Euclidean norm $p=2$, the null hypothesis is rejected, i.e., the sunspot area ratio data are not invariant with the heliolatitude. However, if the examined data are expected to be invariant, and thus the null hypothesis to be accepted, then the norms between p_1 and p_2 (green) are rejected because lead to p-value < 0.05.

The p-value has a minimum value at $p \sim 2.08$ and increases for larger shape values p until it reaches $p \sim 5.77$ where becomes p-value ~ 0.5 (not shown in the figure). If the shape p of the dataset is known, e.g., $p = 2$, then the null hypothesis is rejected, i.e., the sunspot area ratio data are dependent on the heliolatitude. On the other hand, if the data are expected to be invariant with the heliolatitude, and thus the null hypothesis to be accepted, then all the norms between $p_1 \sim 1.7$ and $p_2 \sim 2.5$ are rejected, and the norm L^p with $p \sim 5.77$ characterizes better these data points; the respective mean value is given by $\alpha_p(5.77){\sim}0.164$. Therefore, if we know the shape/norm p that characterizes the data, we can proceed and rate the goodness of the fitting. However, if p is unknown, at least we could detect those values of p for which the null hypothesis is accepted or rejected.

One of the most intriguing questions regarding the L^p-normed fitting is how can we determine the characteristic p-norm of the data. This is the suitable norm that should be used for the fitting of those data (Livadiotis 2007). The maximization of the p-value is one promising method. We demonstrate this as follows. We construct $N = 10^4$ data, $\{f_i\}_{i=1}^{N}$, of a random variable that follows the General Gaussian distribution of shape p, $f_i \sim GG(\mu = 0, \sigma = 1, p = 3)$. Figure 5(a) shows that the normalized histogram of these values matches this General Gaussian distribution. The p-value is approximated using the asymptotic behavior of (complete and incomplete) Gamma functions for large degrees of freedom, $M = 9999$. Hence, in order to derive the maximum p-value, it is sufficient to maximize

$$\text{p-value} \sim \left(\frac{e}{M} p \eta_p X_{\text{est}}^p\right)^{\frac{M}{p}} e^{-\eta_p X_{\text{est}}^p}. \tag{30}$$

This is shown in Figure 5(b), where the peak is at $p \cong 2.95 \pm 0.08$. Therefore, the p-value is maximized at the same value of p-norm as the shape of the General Gaussian distribution.

5. Conclusions

This paper (1) presented the derivation of the Chi-p distribution, the analog of Chi-square distribution but for datasets that follow the General Gaussian distribution of

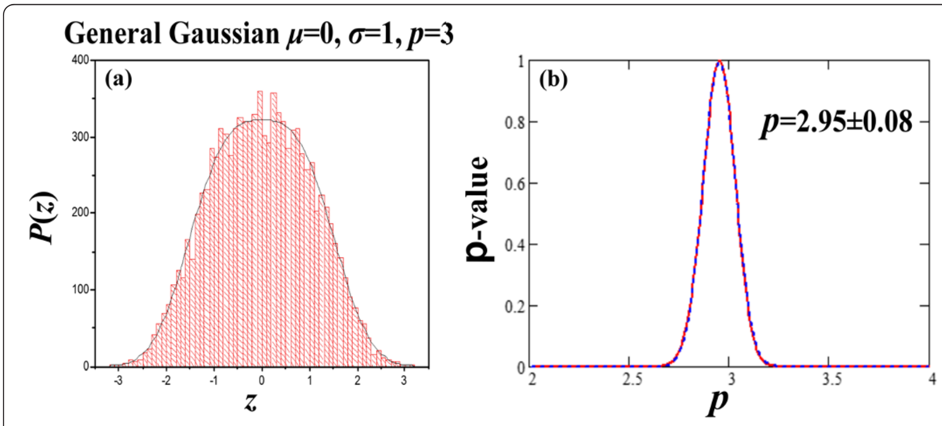

Figure 5 Method for determining the characteristic p-norm of the data. (a) Normalized histogram of $N = 10^4$ data of a random variable that follows the General Gaussian distribution of zero mean, unity variance, and shape $p = 3$. **(b)**. The fitting of the data by a line at $z = 0$ is characterized by a p-value that is maximized at the p-norm $p \cong 2.95 \pm 0.08$, that coincides with the characteristic shape parameter of the data $p = 3$.

shape p, and (2) developed the statistical test for characterizing the goodness of the fitting with L^p norms, which corresponds to datasets that follow the General Gaussian distribution of shape p.

It was shown that the statistical test has double role in the case of L^p norms: (1) If the shape parameter p is fixed and known, then the test can be rated by deriving the p-value. A convenient characterization of the fitting rate was developed. (2) If the shape parameter is unknown and the fitting is expected to be good for some shape parameter value p, a method for estimating p was given by fitting a General Gaussian distribution of shape p to the data, and then use this estimated shape parameter p to the Chi-p distribution to characterize the goodness of fitting. In particular, all the shape values p that correspond to unlikely p-values can be rejected, while the largest p-value corresponds to the most-likely shape parameter p of the examined data. This was verified by an illuminating example where the method of the fitting based on L^p norms was applied.

Appendix A: General Gaussian distribution

According to the theory of L^p-normed mean and variance, developed by Livadiotis (2012), the L^p-normed mean $<x>_p$ of the random variable X with probability distribution $P(x)$, is implicitly defined by

$$\int_{-\infty}^{\infty} P(x) \left|x- < x>_p\right|^{p-1} \mathrm{sign}\left(x- < x>_p\right) dx = 0, \tag{A1}$$

where sign(u) returns the sign of u. The L^p-normed variance σ_p^2 is given by

$$\sigma_p^2 = \frac{\displaystyle\int_{-\infty}^{\infty} P(x) \left|x- < x>_p\right|^{p} dx}{\displaystyle(p-1)\int_{-\infty}^{\infty} P(x) \left|x- < x>_p\right|^{p-2} dx}. \tag{A2}$$

Next, we derive the L^p-normed mean and variance of the General Gaussian distribution (6), which are *Propositions* 1 and 2, stated in Section 2.

- *Proposition 1*: Given the distribution (6), we have that the L^p-normed mean is $<x>_p = \mu$, $\forall\, p \geq 1$.

- *Proof.* We have

$$\int_{-\infty}^{\infty} e^{-\eta_p \cdot |z|^p} \left|z- < z>_p\right|^{p-1} \mathrm{sign}\left(z- < z>_p\right) dz = 0, \tag{A3}$$

for $z \equiv (x-\mu)/\sigma$, $<z>_p \equiv (<x>_p - \mu)/\sigma$. Let's assume that $<z>_p = 0$. Then, the left-hand side of Eq.(A3) is

$$\int_{-\infty}^{\infty} e^{-\eta_p \cdot |z|^p} |z|^{p-1} \mathrm{sign}(z)\, dz = 0, \tag{A4}$$

because the integrant is a product of symmetric and antisymmetric function. Then, (A3) is true for $<z>_p = 0$, and given the uniqueness of the L^p-normed mean for each p, we end up with proposition 1. (Note that it is not surprising that the mean, $<x>_p = \mu$, is independent of p. Livadiotis (2012) showed that symmetric probability distributions lead to L^p-normed means that are independent of p.)

– *Proposition 2*: Given the distribution (6), we have that the L^p-normed variance is $\sigma_p^2 = \sigma^2$, $\forall\, p \geq 1$.

– *Proof.* We have $< |z|^q >= \int_{-\infty}^{\infty} P(z)\,|z|^q\,dz = 0$, i.e.,

$$\int_{-\infty}^{\infty} e^{-\eta_p\cdot|z|^p}|z|^q dz = 2\int_0^{\infty} e^{-\eta_p\cdot z^p} z^q dz = 2\eta_p{}^{-\frac{q+1}{p}}\int_0^{\infty} e^{-w} w^{\frac{q+1}{p}-1} dw$$
$$= 2\eta_p{}^{-\frac{q+1}{p}}\Gamma\left(\frac{q+1}{p}\right), \tag{A5a}$$

or,

$$< z^q >= C_p \frac{2}{p}\eta_p{}^{-\frac{q+1}{p}}\Gamma\left(\frac{q+1}{p}\right). \tag{A5b}$$

Hence, from (A2) we obtain

$$\sigma_p^2 = \frac{\int_{-\infty}^{\infty} P(z)\,|z|^p\,dz}{(p-1)\int_{-\infty}^{\infty} P(z)\,|z|^{p-2}dz}\cdot\sigma^2 = \frac{\eta_p{}^{-\frac{2}{p}}\Gamma\left(1+\frac{1}{p}\right)}{(p-1)\,\Gamma\left(1-\frac{1}{p}\right)}\cdot\sigma^2 = \sigma^2. \tag{A6}$$

Appendix B: Surface of the N-dimensional sphere in L^p space, $B_{p,N}$

This appendix shows the proof of *Lemma* 1, stated in Section 2.

– *Lemma* 1: The surface of the N-dimensional sphere of unit radius in L^p space, $B_{p,N}$, is given by Eq.(8). This is involved in the proof of Chi-p distribution (10), as shown below.

– *Proof of Lemma* 1.
Let the integral

$$1 = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} F\left(\vec{z}\right) dz_1 \ldots\ dz_N, \tag{B1}$$

where $\vec{z} = (z_1, \ldots, z_N)$, $Z^p \equiv \sum_{i=1}^N |z_i|^p$. The magnitude Z is the only quantity with dimensions the same as each of the components z_i. Indeed, if we define $c_i \equiv z_i/\zeta$, where $\zeta \equiv \sqrt{\sum_{i=1}^N z_i{}^2}$ is the Euclidean magnitude of $\vec{z}$, then, $Z = \left(\sum_{i=1}^N |z_i|^p\right)^{\frac{1}{p}} = \zeta\cdot\left(\sum_{i=1}^N |c_i|^p\right)^{\frac{1}{p}}$, i.e., Z and ζ have the same dimensions. (In the previous sections the components z_i were dimensionless by definition, i.e., $z_i \equiv \frac{x_i-\mu_x}{\sigma_{x_i}}$. However, we can still use this dimension analysis, since the components z_i may have dimensions in the generic case). Hence, we write Eq.(B1) as $dz_1 \ldots dz_N = Z^{N-1} dZ\, d^{N-1}\Omega_N$, i.e.,

$$1 = \int_0^{+\infty}\int_{\vec{z}\in B_{p,N}} F(Z;\Omega_N)\, Z^{N-1} dZ\, d^{N-1}\Omega_N, \tag{B2}$$

where $F\left(\vec{z}\right)=F(Z;\Omega_N)$; Ω_N symbolizes all the angular dependence, and $d^{N-1}\Omega_N$ denotes the angular infinitesimal. Since $F(Z;\Omega_N)=F(Z)$, we have $B_{p,N}\equiv\int_{\vec{z}\in B_{p,N}} d^{N-1}\Omega_N$, or

$$1=\int_{\vec{z}\in B_{p,N}} d^{N-1}\Omega_N\cdot\int_0^{+\infty}F(Z)\,Z^{N-1}dZ=B_{p,N}\cdot\int_0^{+\infty}F(Z)\,Z^{N-1}dZ$$
$$=C_p^N\frac{1}{p}B_{p,N}\cdot\int_0^{+\infty}F\left(X^{\frac{1}{p}}\right)X^{\frac{N}{p}-1}dX,$$

where $F(Z)=C_p^N e^{-\eta_p Z^p}$, $F\left(X^{\frac{1}{p}}\right)=C_p^N e^{-\eta_p X}$. Therefore,

$$1=\int_0^{+\infty}C_p^N\frac{1}{p}B_{p,N}\,e^{-\eta_p X}X^{\frac{N}{p}-1}dX\equiv\int_0^{+\infty}P(X;N;p)dX,$$

or,

$$P(X;N;p)=C_p^N\frac{1}{p}B_{p,N}\,e^{-\eta_p X}X^{\frac{N}{p}-1}. \tag{B3}$$

The normalization $\int_0^{+\infty}P(X;N;p)dX=1$ gives $C_p^N\frac{1}{p}B_{p,N}=\eta_p^{\frac{N}{p}}/\Gamma\left(\frac{N}{p}\right)$, or

$$B_{p,N}=p\eta_p^{\frac{N}{p}}/\left[C_p^N\Gamma\left(\frac{N}{p}\right)\right]=p\left[\left(\frac{2}{p}\right)\Gamma\left(\frac{1}{p}\right)\right]^N/\Gamma\left(\frac{N}{p}\right). \tag{B4}$$

Another way to show Eq.(B4) is through the integration of all the components,

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}F\left(\vec{z}\right)dz_1\ldots\ dz_N=2^N\cdot\int_0^{+\infty}\cdots\int_0^{+\infty}F\left(\vec{z}\right)dz_1\ldots\ dz_N$$
$$=2^N\cdot\int_0^{+\infty}F(Z)\int_{\substack{\vec{z}\in B_{p,N}\\ z_i\geq 0}}\left(Z^p-z_2^p-z_3^p\ldots-z_N^p\right)^{\frac{1}{p}-1}Z^{p-1}\,dZdz_2\ldots dz_N,$$

by substituting $F\left(\vec{z}\right)=F(Z)$ and $z_1=\left(Z^p-z_2^p-z_3^p\ldots-z_N^p\right)^{\frac{1}{p}}$ (for $z_i\geq 0$). The integration range $\vec{z}\in B_{p,N}$, $z_i\geq 0$, means $0\leq z_i\leq\left(Z^p-\sum_{i+1}^N z_j^p\ldots-z_N^p\right)^{\frac{1}{p}}$ for $i=1,\ldots,N$-1, and $0\leq z_N\leq Z$. Similar, we have

$$\int_{\substack{\vec{z}\in B_{p,N}\\ z_i\geq 0}}\left(Z^p-z_2^p-z_3^p\ldots-z_N^p\right)^{\frac{1}{p}-1}dz_2\ldots dz_N=a_{1,p}\int_{\substack{\vec{z}\in B_{p,N}\\ z_i\geq 0}}\left(Z^p-z_3^p\ldots-z_N^p\right)^{\frac{2}{p}-1}dz_3\ldots dz_N$$
$$=a_{1,p}a_{2,p}\int_{\substack{\vec{z}\in B_{p,N}\\ z_i\geq 0}}\left(Z^p-z_4^p\ldots-z_N^p\right)^{\frac{3}{p}-1}dz_4\ldots dz_N=\prod_{i=1}^{N-2}a_{i,p}\cdot\int_{\substack{\vec{z}\in B_{p,N}\\ z_i\geq 0}}\left(Z^p-z_N^p\right)^{\frac{N-1}{p}-1}dz_N$$
$$=\prod_{i=1}^{N-1}a_{i,p}\cdot Z^{N-p},$$

where

$$a_{i,p}\equiv\int_0^1\left(1-t^p\right)^{\frac{i}{p}-1}dt. \tag{B5}$$

Hence, we derive

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}F\left(\vec{z}\right)dz_1\ldots\ dz_N=2^N\cdot\prod_{i=1}^{N}a_{i,p}\cdot\int_0^{+\infty}F(Z)\,Z^{N-1}\,dZ, \tag{B6}$$

while, on the other hand, we have

$$\begin{aligned}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}F\left(\vec{z}\right)dz_1\ldots\ dz_N&=\int_{\vec{z}\in B_{p,N}}d^{N-1}\Omega_N\cdot\int_0^{+\infty}F(Z)\,Z^{N-1}dZ\\&=B_{p,N}\cdot\int_0^{+\infty}F(Z)\,Z^{N-1}dZ,\end{aligned} \tag{B7}$$

thus,

$$B_{p,N}=2^N\cdot\prod_{i=1}^{N-1}a_{i,p}. \tag{B8}$$

We easily find that

$$a_{i,p}=\frac{1}{p}\int_0^1 y^{\frac{1}{p}-1}(1-y)^{\frac{i}{p}-1}\,dy=\frac{1}{p}B\left(\frac{1}{p},\frac{i}{p}\right), \tag{B9}$$

where $B(x,y)\equiv\Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the Beta function. Hence, we have

$$B_{p,N}=p\left(\frac{2}{p}\right)^N\Gamma\left(\frac{1}{p}\right)^{N-1}\cdot\prod_{i=1}^{N-1}\Gamma\left(\frac{i}{p}\right)/\Gamma\left(\frac{i+1}{p}\right). \tag{B10}$$

Since, $\prod_{i=1}^{N-1}\Gamma\left(\frac{i}{p}\right)/\Gamma\left(\frac{i+1}{p}\right)=\Gamma\left(\frac{1}{p}\right)/\Gamma\left(\frac{N}{p}\right)$, finally, we end up with Eq.(B4).

Competing interests
The authors declare that they have no competing interests.

Received: 6 June 2013 Accepted: 24 December 2013
Published: 11 June 2014

References
Adèr, HJ: Modelling (Chapter 12). In: Adèr, HJ, Mellenbergh, GJ, with contributions by D.J. Hand (eds.) Advising on Research Methods: A consultant's companion, pp. 271–304. Johannes van Kessel Publishing, Huizen, The Netherlands (2008)
Burden, RL, Faires, JD: Numerical Analysis, pp. 437–438. PWS Publishing Company, Boston, MA (1993)
Edwards, AWF: The proportion of umbra in large sunspots, 1878–1954. The Observatory **77**, 69–70 (1957)
Frisch, PC, Bzowski, M, Livadiotis, G, McComas, DJ, Möbius, E, Mueller, HR, Pryor, WR, Schwadron, NA, Sokól, JM, Vallerga, JV, Ajello, JM: Decades-long changes of the interstellar wind through our solar system. Science **341**, 1080 (2013)
Funsten, HO, Frisch, PC, Heerikhuisen, J, Higdon, DM, Janzen, P, Larsen, BA, Livadiotis, G, McComas, DJ, Möbius, E, Reese, CS, Reisenfeld, DB, Schwadron, NA, Zirnstein, E: The circularity of the IBEX Ribbon of enhanced energetic neutral atom flux. Astrophys. J. **776**, 30 (2013)
Livadiotis, G: Approach to general methods for fitting and their sensitivity. Physica A **375**, 518–536 (2007)
Livadiotis, G: Approach to the block entropy modeling and optimization. Physica A **387**, 2471–2494 (2008)
Livadiotis, G: Expectation values and Variance based on L^p norms. Entropy **14**, 2375–2396 (2012)
Livadiotis, G, McComas, DJ: Fitting method based on correlation maximization: Applications in Astrophysics. J. Geophys. Res. **118**, 2863–2875 (2013a)
Livadiotis, G, McComas, DJ: Evidence of large scale phase space quantization in plasmas". Entropy **15**, 1116–1132 (2013b)
Livadiotis, G, Moussas, X: The sunspot as an autonomous dynamical system: A model for the growth and decay phases of sunspots. Physica A **379**, 436–458 (2007)
McCullagh, P: What is statistical model? Ann. Stat. **30**, 1225–1310 (2002)
Melissinos, AC: Experiments in Modern Physics, pp. 464–467. Academic Press Inc, London, UK (1966)
Sengupta, A: A rational function approximation of the singular eigenfunction of the monoenergetic neutron transport equation. J. Phys. A **17**, 2743–2758 (1984)

doi:10.1186/2195-5832-1-4
Cite this article as: Livadiotis: Chi-*p* distribution: characterization of the goodness of the fitting using L^p norms. *Journal of Statistical Distributions and Applications* 2014 **1**:4.