Measurement 48 (2014) 13–20

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Bayesian inference on a squared quantity

Carlo Carobbi *

Università degli Studi di Firenze, Department of Information Engineering (DINFO), Via S. Marta 3, 50139 Firenze, Italy

ARTICLE INFO

Article history:
Received 23 August 2013
Received in revised form 3 October 2013
Accepted 24 October 2013
Available online 31 October 2013

Keywords:
Bayesian inference
Frequentist inference
Method of moments estimators
Measurement uncertainty
Noncentral chi-squared distribution

ABSTRACT

It is here derived the Bayesian estimator of the noncentrality parameter of the noncentral chi-square distribution. The corresponding frequentist estimator, based on the method of moments, is also derived and its performance is compared with the Bayesian one. The Bayesian estimator is obtained through an analytical derivation which provides insight into the way the estimator works. Reference is also made here to a previously published work on a similar subject by Attivissimo et al. (2012) [1] in order to resolve the paradox there presented. Some defects of the analysis performed in the referenced work are identified and carefully examined. The superiority of the Bayesian estimator is demonstrated although achieved at the price of a greater complexity.

1. Introduction

In the recent paper [1] arguments are presented in favour of frequentist inference and against Bayesian inference on the basis of what the authors of [1] judge to be an apparent paradox. The scope here is to investigate how Bayesian inference works when predicting the square of an unknown quantity. It is also shown that the arguments in [1] against Bayesian inference and in favour of frequentist one are questionable.

The paper is structured as follows. In Section 2 the basic assumptions about the involved quantities are done and the statistical inference problem is defined and solved through a frequentist analysis based on the method of moments (MoM). The paradox described in [1] is discussed in detail and resolved in Section 3. Section 4 is devoted to the solution of the same inference problem but by using a Bayesian instead of a frequentist analysis. The derivation of Bayesian estimators is done in such a way to ease interpretation of how the estimators process available information and in particular observed values. The inference problem essentially consists in finding the estimator of the noncentrality parameter of the noncentral chi-square distribution with one degree of freedom. As far as we know this problem has not been previously dealt with by using a Bayesian approach, either in the scientific or technical literature. In Section 5 the Bayesian estimator is compared with the frequentist one through numerical computation. Section 6 is finally devoted to the conclusive remarks.

Only a few and essential references appear. Attivissimo et al. [1] is the paper that triggered this work, in [2,3] Bayesian inference is dealt with in the light of the most authoritative standards devoted to the evaluation of measurement uncertainty, namely [4] (the so-called GUM) and [5] (Supplement 1 to GUM). Therefore the reader interested in a detailed discussion about the relation between the GUM, its Supplement 1 and Bayesian analysis is directed to [2,3] (rather than to [1], where such relation is also dealt with). References from [6–9] provide specific theoretical support and interpretation to the mathematical analysis developed in the following sections.

The following conventions are here adopted. The name of a quantity and the corresponding random variable are both represented by an italic and capital letter, e.g. Q. A realization or observation of the quantity Q, as obtained from measurement, is represented by the corresponding italic and lower case letter, i.e. q.

* Tel.: +39 055 4796268; fax: +39 055 4796497.
E-mail address: carlo.carobbi@unifi.it

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.10.034

2. Frequentist inference on μ^2

Let $X_1, X_2, ..., X_n$ be n independent, real, normal quantities with expected value μ and variance σ^2. It assumed that μ is unknown while σ^2 is known, as in [1], and we want to predict μ^2. In doing this we consider two different situations A and B, where two distinct sets of observations are obtained from measurements, namely

A. $x_1, x_2, ..., x_n$, or
B. $x_1^2, x_2^2, \ldots, x_n^2$.

Both situations A and B are of interest, in particular in electrical and electronic measurements where μ may represent the value of a constant signal and σ the root-mean-square value of an additive and thermal equivalent (zero-mean, stationary, normal) noise. In situation A the measuring instrument (e.g. an oscilloscope or digital multimeter) can detect the magnitude and sign of X. In B the a square-law detector is implemented in the measuring instrument (e.g. in field and power meters) that therefore detects X^2.

In this section we deal with the problem of estimating μ^2 by using frequentist inference based on the method of moments (MoM), see [6 , p. 301]. Let $E_X[g(X)]$ and $V_X[g(X)]$ be respectively the expectation and variance of $g(X)$, where $g(X)$ represents a general function of X. Expectation and variance are calculated by using the probability density function (pdf) of X (normal in this case). In situation A the true power μ^2 is estimated by $(\overline{X})^2$, where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. $\overline{X}$ is normal with expected value μ and variance σ^2/n. Therefore $[\overline{X}/(\sigma/\sqrt{n})]^2$ is a noncentral chi-square random variable with $\nu = 1$ degrees of freedom and noncentrality parameter $\lambda = [\mu/(\sigma/\sqrt{n})]^2$. Then [7, p. 943]

$$E_X\left[(\overline{X})^2\right] = \frac{\sigma^2}{n}(\nu + \lambda) = \mu^2 + \frac{\sigma^2}{n}, \tag{1}$$

and

$$V_X\left[(\overline{X})^2\right] = \frac{\sigma^4}{n^2}2(\nu + 2\lambda) = 2\frac{\sigma^2}{n}\left(2\mu^2 + \frac{\sigma^2}{n}\right) \tag{2}$$

MoM estimator in situation A (MoM estimator A) is obtained from (1) and (2), in terms of the estimate μ_A^2 and its variance $u^2(\mu_A^2)$, as

$$\mu_A^2 \approx (\bar{x})^2 - \frac{\sigma^2}{n}, \tag{3}$$

$$u^2(\mu_A^2) \approx 2\frac{\sigma^2}{n}\left[2(\bar{x})^2 + \frac{\sigma^2}{n}\right], \tag{4}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Eqs. (3) and (4) need some comments. For the validity of (3) it is needed that $E_X[(\overline{X})^2] \approx (\bar{x})^2$ and that $(\bar{x})^2 - \frac{\sigma^2}{n} > 0$. Both conditions are verified when n is so large that σ^2/n is small when compared with μ^2. Then for small sample size and small signal-to-noise ratio MoM estimate is expected to fail, in the sense that μ_A^2 may be largely deviated from μ^2 or may even be negative (which is evidently absurd). Note that (4) has been derived substituting $(\bar{x})^2$ to μ^2 in (2). Actually, the logical estimate of the variance of μ_A^2 should be derived substituting μ_A^2 to μ^2 in (2), thus obtaining $u^2(\mu_A^2) \approx 2\frac{\sigma^2}{n}\left[2(\bar{x})^2 - \frac{\sigma^2}{n}\right]$. However this estimate of $u^2(\mu_A^2)$ is not strictly positive and therefore unacceptable (variance is positive by definition). This lack of transparency of the MoM estimator is not incidental but intrinsic to frequentist inference, contrarily to Bayesian inference. Estimator (3) is consistent, in that it converges (in probability) to the true value μ^2 when n increases. Correspondingly its variance (4) tends to zero.

In situation B the estimator of μ^2 (MoM estimator B) is $\overline{(X^2)} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$. $\overline{(X^2)}/(\sigma^2/n)$ is a noncentral chi-squared random variable with $\nu = n$ degrees of freedom and noncentrality parameter $\lambda = [\mu/(\sigma/\sqrt{n})]^2$. We then have

$$E_X[\overline{(X^2)}] = \frac{\sigma^2}{n}(\nu + \lambda) = \mu^2 + \sigma^2, \tag{5}$$

and

$$V_X[\overline{(X^2)}] = \frac{\sigma^4}{n^2}2(\nu + 2\lambda) = 2\frac{\sigma^2}{n}(2\mu^2 + \sigma^2). \tag{6}$$

The MoM estimator for situation B is derived similarly to the one for situation A as

$$\mu_B^2 \approx \overline{(x^2)} - \sigma^2, \tag{7}$$

$$u^2(\mu_B^2) \approx 2\frac{\sigma^2}{n}\left[2\overline{(x^2)} + \sigma^2\right] \tag{8}$$

where $\overline{(x^2)} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

Analogous considerations to those done above for μ_A^2 apply to μ_B^2 and concerning the validity of the estimator, its variance and consistency. When $n = 1$ we have $(\bar{x})^2 = \overline{(x^2)} = x_1^2$ and

$$\mu_A^2 = \mu_B^2 \approx x_1^2 - \sigma^2, \tag{9}$$

and

$$u^2(\mu_A^2) = u^2(\mu_B^2) \approx 2\sigma^2(2x_1^2 + \sigma^2). \tag{10}$$

Finally note that when $n > 1$ MoM estimator A has less bias and variance than MoM estimator B.

3. Resolution of the paradox [1]

The analysis in [1] is done in the case $n = 1$ where, as it will be shown in Section 4, Bayesian inference leads to the same results in both situations A and B, namely

$$E_\mu[\mu^2] = x_1^2 + \sigma^2, \tag{11}$$

and

$$V_\mu[\mu^2] = 2\sigma^2(2x_1^2 + \sigma^2). \tag{12}$$

It is important to observe that in the Bayesian approach μ is viewed as a random variable whose known probability density function is a mathematical description of the state of knowledge about its unknown and constant true value.