UDC 006.91:053.088

Analysis and comparison of Bayesian methods for type A uncertainty evaluation with prior knowledge

I. Lira

Pontificia Universidad Catolica de Chile, Vicuña Mackenna 1460, Santiago, Chile
ilira@ing.puc.cl

Abstract

If a number of observations about a certain quantity may be assumed independent, drawn from a Gaussian distribution, Supplement 1 to the GUM recommends that the standard uncertainty associated with the quantity be obtained by a formula that is applied to more than three observations. Various articles have recently appeared proposing Bayesian methods to surmount this limitation. Some of these methods, which require prior knowledge about the quantity, are reviewed in this article.

Keywords: Bayesian methods; type A uncertainty evaluation; prior knowledge.

Received: 01.12.2022 Edited: 16.12.2022 Approved for publication: 21.12.2022

1. Introduction

When data $\{x_1, x_2, ..., x_n\}$ about a certain quantity are available, the GUM [1] proposes that their mean, $\bar{x}$, be taken as the estimate of the quantity and that their standard uncertainty, u_{GUM}, be calculated as $s/\sqrt{n}$, where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. This procedure is based on frequentist concepts. Instead, based on the Bayesian paradigm, Supplement 1 to the GUM [2] recommends calculating $u_{\mathrm{GUM\text{-}S1}} = \phi \cdot u_{\mathrm{GUM}}$, where $\phi = \sqrt{\frac{n-1}{n-3}}$. This formula is applied only to independent data, obtained from a Gaussian distribution, neglecting any prior knowledge about the measurand that might be available. Obviously, evaluating ϕ requires at least four observations.

Below, we review some of the recently appeared proposals for obtaining the standard uncertainty, which are valid for any number of observations.

2. Proposals – 1

Reference [3] appears to be the first paper to propose solving the problem by taking advantage of the knowledge available before measurements. Its authors applied Bayes' formula to a Gaussian likelihood and informative prior proportional to $1/\upsilon$ for $\upsilon_{\min} \le \upsilon \le \upsilon_{\max}$, where υ is the unknown variance of a Gaussian distribution.

Using clever manipulations, they could derive an analytic expression for factor ϕ involving the upper incomplete gamma function. Results can be easily obtained using the Wolfram Mathematica© software (Fig. 1).

```
n = 2;  x = {0.9551, 0.9537}; mean = Mean[x];
s = StandardDeviation[x]; sM = 0.003; sm = 0.001;
S = (n - 1) s^2;
as = (n - 3)/2; ai = (n - 1)/2; a = S/(2 sM^2); b = S/(2 sm^2);
phi = ((n - 1)/2 (Gamma[as, a] - Gamma[as, b])/(Gamma[ai, a] - Gamma[ai, b]))^(1/2);
uGum = s/Sqrt[n];  uCS = phi uGum;
Print["uCS      = ", uCS]
Print["uGum     = ", uGum]
uGum     = 0.0007
uCS      = 0.0012686
```

Fig. 1. Example 4.3 in [3]

Alternatively, one can use a direct numerical application of Bayes' theorem. In Fig. 2, p0[•] is the prior, l[•] is the likelihood, p1[•] is the joint posterior, p2[•] is the unnormalized marginal posterior and c is the normalization constant.

```
p0[mu_, v_] := If[sm^2 ≤ v ≤ sM^2, 1/v, 0]
l[mu_, v_] := 1/v^(n/2) Exp[-(S + n (mu - mean)^2)/(2 v)]
p1[mu_, v_] := p0[mu, v] × l[mu, v]
p2[mu_] := NIntegrate[p1[mu, v], {v, sm^2, sM^2}]
c = (Quiet[NIntegrate[p2[mu], {mu, -∞, ∞}]])^-1;
expec = Quiet[NIntegrate[mu c p2[mu], {mu, -∞, ∞}]];
stdev =
  (Quiet[NIntegrate[(mu - expec)^2 c p2[mu], {mu, -∞, ∞}]])^(1/2);
unum = stdev;
Print["unum = ", unum]
Print["uCS  = ", uCS]
uCS  = 0.0012686
unum = 0.00126859
```

Fig. 2. Example 4.3 in [3] using numerical integration

The method in [3] requires n≥2. However, the numerical procedure is applied even to one observation (Fig. 3).

```
x = 0.9551;
p0[mu_, v_] := If[sm^2 ≤ v ≤ sM^2, 1/v, 0]
l[mu_, v_] := 1/v^(1/2) Exp[-(mu - x)^2/(2 v)]
p1[mu_, v_] := p0[mu, v] × l[mu, v]
p2[mu_] := NIntegrate[p1[mu, v], {v, sm^2, sM^2}]
c = (Quiet[NIntegrate[p2[mu], {mu, -∞, ∞}]])^-1;
expec = Quiet[NIntegrate[mu c p2[mu], {mu, -∞, ∞}]];
stdev =
  (Quiet[NIntegrate[(mu - expec)^2 c p2[mu], {mu, -∞, ∞}]])^(1/2);
Print["u 1 obs = ", stdev]
Print["u 2 obs = ", uCS]
u 2 obs = 0.0012686
u 1 obs = 0.00190813
```

Fig. 3. Example 4.3 in [3] considering only one observation

3. Proposals – 2

In [4], its author proposes using a half-Cauchy prior for the standard deviation, that is, a t-Student with one degree of freedom. Such a prior can be written in the form $p(\mu,\upsilon)\propto(A+\upsilon)^{-1}$, where A is the constant. This median of the distribution can be equal to any available prior estimate of the variance. In this way, a reasonable value for A can be obtained. The resulting posterior does not have any closed form. Of course, it can be evaluated by MCMC, as proposed in [4]. However, as the following example demonstrates, it is simpler (and faster) to use the numerical procedure just described (Fig. 4).

```
x = {0.9551, 0.9537};
m = ((sM sm)/(sM + sm))^2;  A = m/(-1 + Sqrt[e]);  p0[mu_, v_] := (A + v)^-1
l[mu_, v_] := 1/v^(n/2) Exp[-(S + n (mu - mean)^2)/(2 v)]
p1[mu_, v_] := p0[mu, v] × l[mu, v]
p2[mu_] := NIntegrate[p1[mu, v], {v, 0, ∞}]
c = (Quiet[NIntegrate[p2[mu], {mu, -∞, ∞}]])^-1;
expec = Quiet[NIntegrate[mu c p2[mu], {mu, -∞, ∞}]];
stdev =
  (Quiet[NIntegrate[(mu - expec)^2 c p2[mu], {mu, -∞, ∞}]])^(1/2);
uCauchy = stdev;
Print["uCauchy = ", uCauchy]
Print["uCS     = ", uCS]
uCS     = 0.0012686
uCauchy = 0.531866
```

Fig. 4. Example 4.3 in [3] using half-Cauchy prior

4. Proposals – 3

Reference [5] is applied to measurement models of the form $Y=aX+B$, where B represents a linear combination of type B quantities. However, by taking $B = 0$ (and $a = 1$), we recover our measurement model. The authors of [5] assume an inverse gamma prior for the variance of the sampling distribution of repeated measurements of X. If we neglect previous knowledge about the measurand value, the prior becomes

$$p(\mu,\upsilon)\propto\left(\frac{1}{\upsilon}\right)^{1+a}\exp\left(-\frac{b}{\upsilon}\right).$$

The authors of [5] propose using $a = 1$, which gives a weakly informative prior with neither finite mean nor variance. With this choice, the median of the distribution, which is equal to $b/\ln 2$, can again be taken as the prior estimate of the variance, if available. In this way, the value of parameter b is obtained (Fig. 5).

```
x = {0.9551, 0.9537}; a = 1;
m = (sM sm)/(sM + sm);  b = m Log[2];  p0[mu_, v_] := 1/v^(1+a) Exp[-b/v]
l[mu_, v_] := 1/v^(n/2) Exp[-(S + n (mu - mean)^2)/(2 v)]
p1[mu_, v_] := p0[mu, v] × l[mu, v]
p2[mu_] := NIntegrate[p1[mu, v], {v, 0, ∞}]
c = (Quiet[NIntegrate[p2[mu], {mu, -∞, ∞}]])^-1;
expec = Quiet[NIntegrate[mu c p2[mu], {mu, -∞, ∞}]];
stdev =
  (Quiet[NIntegrate[(mu - expec)^2 c p2[mu], {mu, -∞, ∞}]])^(1/2);
ugamma = stdev;
Print["ugamma  = ", ugamma]
Print["uCauchy = ", uCauchy]
Print["uCS     = ", uCS]
uCS     = 0.0012686
uCauchy = 0.531866
ugamma  = 0.0228108
```

Fig. 5. Example 4.3 in [3] using inverse gamma prior

However, the inverse gamma prior gives a marginal posterior for the measurand equal to a scaled and shifted t-distribution, whose variance can be calculated as $\frac{2b+(n-1)s^2}{n(n+2a-3)}$. In this case, the calculated value of the standard deviation will be equal to **0.0228112**, which practically coincides with the value obtained in Fig. 5.

From this formula, we see that the inverse gamma prior can again be used if there is only one observation, but in that case, the shape parameter a has to be greater than one.

5. Proposals – 4

Reference [6] seems to be the most recent contributor to this discussion. It proposes using a scaled inverse chi-square distribution as an informative prior:

$$p(\mu,\upsilon)\propto\left(\frac{1}{\upsilon}\right)^{1+\nu/2}\exp\left(-\frac{\sigma_0^2\nu}{2\upsilon}\right),$$

where σ_0^2 is the prior variance and v is the associated number of degrees of freedom. It is clear that this distribution describes the same data structure as the inverse gamma does, but with a different parameterization, the relation between the two parameters being $v = 2a$ and $\sigma_0^2 = \dfrac{b}{a}$ (Fig. 6).

```
x = {0.9551, 0.9537}; nu = 2 a; s02 = b/a;
p0[mu_, v_] := 1/v^(1+nu/2) Exp[-s02 nu/(2 v)]
l[mu_, v_] := 1/v^(n/2) Exp[-(S + n (mu - mean)^2)/(2 v)]
p1[mu_, v_] := p0[mu, v] × l[mu, v]
p2[mu_] := NIntegrate[p1[mu, v], {v, 0, ∞}]
c = (Quiet[NIntegrate[p2[mu], {mu, -∞, ∞}]])^-1;
expec = Quiet[NIntegrate[mu c p2[mu], {mu, -∞, ∞}]];
stdev =
  (Quiet[NIntegrate[(mu - expec)^2 c p2[mu], {mu, -∞, ∞}]])^(1/2);
Print["ugamma  = ", ugamma]
Print["uchi2   = ", stdev]
uchi2   = 0.0228108
ugamma  = 0.0228108
```

Fig. 6. Example 4.3 in [3] using scaled inverse chi-square distribution

Therefore, mathematically, it does not matter whether one uses the inverse γ prior or the inverse χ^2 prior. However, the proposal by Carobbi is more intuitive, because as Gelman et. al. point out in their famous book [7], the scaled inverse χ^2 can be thought of as providing the information equivalent to v observations with average squared deviation σ_0^2. These two parameters may be available from prior experiments.

6. Conclusion

1) The Cox-Shirono approach causes comparatively very small uncertainties. This appears to be the result of a strongly informative prior. Perhaps another form of the function $p(\mu, \upsilon)$, which is supported on the interval $\upsilon_{min} \le \upsilon \le \upsilon_{max}$, might yield results that are more reasonable.

2) This comment elicits a query: would there be a way of quantifying the “informativeness” of a prior?

3) For generality, we wrote the prior as $p(\mu, \upsilon)$, even though in none of the proposals it depends on the unknown mean of a Gaussian distribution, μ. That is to say, we have assumed that there is no prior information on the value of the measurand. Evidently, this assumption simplifies matters, but it is not very reasonable. A better approach is the one used by Wübbeler [5], who proposes a NIG prior.

4) If one decides to keep with the assumption of constant μ, the inverse γ and χ^2 priors produce a simple algebraic formula for the standard uncertainty, usable by all practitioners. However, the latter prior may be preferred because its parameters have an intuitive interpretation.

5) In simple cases, the Mathematica software allows direct numerical exploration of marginal posteriors, such as computing credible intervals, without the need to use MCMC or other alternatives.

Аналіз та порівняння Байєсівських методів для оцінки невизначеності типу А з попередніми знаннями

І. Ліра

Папський католицький університет Чилі, Vicuña Mackenna 1460, Сантьяго, Чилі
ilira@ing.puc.cl

Анотація

Якщо можна припустити, що ряд спостережень за певною величиною є незалежним, узятим із розподілу Гаусса, то доступний Додаток 1 GUM рекомендує, щоб стандартну невизначеність, пов’язану з цією величиною, було отримано з формули, яка застосовується до більш ніж трьох спостережень. На практиці це обмеження може призводити до значних витрат на проведення додаткових вимірювань. Однак наведена формула ігнорує будь-які попередні знання про вимірювану величину. Нещодавно з’явилися різні статті, у яких пропонується застосування Байєсівських методів для подолання цього обмеження. Деякі з цих методів, для яких вимагається попереднє знання величини, обговорюються в цій статті. У процесі дослідження з’ясувалося, що підхід Кокса-Широно обумовлює порівняно дуже невеликі невизначеності. Цей результат здається дуже інформативним апріорним. Можливо, інша форма функції, що підтримується на відповідному інтервалі, може надати більш прийнятні результати. Для загальності апріорне було записано залежним від невідомого середнього розподілу Гаусса, хоча й в жодній із пропозицій він від нього не залежить. Тобто було зроблено припущення, що немає абсолютно ніякої попередньої інформації про значення вимірюваної величини. Кращим підходом є той, що використовував Вюббелер, який запропонував апріорне як нормальний обернений гамма-розподіл. Для певних умов отримано просту алгебраїчну формулу оцінювання стандартної невизначеності, яку можна використовувати для всіх практик. Перевагу можна віддати останньому апріорному, оскільки його параметри мають інтуїтивно зрозумілу інтерпретацію. Показано, що в простих випадках програмне забезпечення Mathematica дозволяє пряме числове дослідження граничних апостеріорних ймовірностей, наприклад, обчислення довірчих інтервалів, не потребуючи використання Монте-Карло марковських ланцюгів чи інших альтернативних методів.

Ключові слова: Байєсівські методи; оцінка невизначеності типу А; попередні знання.

References

1. JCGM 100:2008 (GUM 1995 with minor corrections). Evaluation of measurement data – Guide to the expression of uncertainty in measurement. 134p.
2. JCGM 101:2008. Evaluation of measurement data – Supplement 1 to the “Guide to the expression of uncertainty in measurement” – Propagation of distributions using a Monte Carlo method. 82 p.
3. Cox M., Shirono K. Informative Bayesian Type A uncertainty evaluation, especially applicable to a small number of observations. *Metrologia*, 2017, vol. 54, no. 5, p. 642. doi: 10.1088/1681-7575/aa787f
4. Adriaan M.H. van der Veen. Evaluating measurement uncertainty in fluid phase equilibrium calculations. *Metrologia*, 2018, vol. 55, no. 2, p. S60. doi: 10.1088/1681-7575/aaa6dd
5. Wübbeler G., Marschall M., Elster C. A simple method for Bayesian uncertainty evaluation in linear models. *Metrologia*, 2020, vol. 57, no. 6, p. 065010. doi: 10.1088/1681-7575/aba3b8
6. Carobbi C. An informed type A evaluation of standard uncertainty valid for any sample size greater than or equal to 1. *Acta Imeko*, 2022, vol. 11, no. 2, pp. 1–5.
7. Gelman A., Carlin J., Stern H., Rubin D. Bayesian Data Analysis. Second Edition. Chapman and Hall/CRC, 2003. 690 p.