5-24-2024

Valid Confidence Intervals for μ,σ when There is Only One Observation Available

Anirban Dasgupta
University of Illinois at Urbana-Champaign

Stephen Portnoy
Portland State University

Citation Details

Published as: DasGupta, A., & Portnoy, S. (2024). Valid Confidence Intervals for μ,σWhen There Is Only One Observation Available. Sankhya A.

Valid confidence intervals for μ, σ when there is only one observation available

Steve Portnoy * †‡

Anirban DasGupta §

February 9, 2022

ABSTRACT Portnoy (2019) considered the problem of constructing an optimal confidence interval for the mean based on a single observation $X \sim \mathcal{N}(\mu, \sigma^2)$. Here we extend this result to obtaining 1-sample confidence intervals for σ and to cases of symmetric unimodal distributions and of distributions with compact support. Finally, we extend the multivariate result in Portnoy (2019) to allow a sample of size m from a multivariate normal distribution where m may be less than the dimension.

AMS 2010 subject classifications. Primary- 62F10; secondary- 62C10

*Corresponding author.

†Department of Statistics, University of Illinois at Urbana-Champaign, IL, Email: sportnoy@illinois.edu

‡ and Department of Mathematics and Statistics, Portland, OR

§Department of Statistics, Purdue University, West Lafayette, IN. Email: dasgupta@purdue.edu.

1 Introduction

As noted in Portnoy (2019), the problem of constructing a confidence interval for the mean based on a single observation $X \sim \mathcal{N}(\mu,\, \sigma^2)$ has been considered for some time. The first published version appears to be Abbott and Rosenblatt (1962) who showed that the interval $(X - c_\alpha|X|,\, X + c_\alpha|X|)$ has coverage probability at least $(1-\alpha)$ for appropriately chosen c_α. At about the same time Charles Stein presented a cloassroom example of the form $(-d_\alpha|X|,\, d_\alpha|X|)$, which appeared to have been developed earlier. Statements attributing the idea to Herb Robbins (in Rodriguez (1996) and in a personal communication from Persi Diaconis) suggest that the example was known to theoretical statisticians before 1960.

Portnoy (2019) considered the problem in somewhat more depth and found a legitimate confidence interval that is optimal in the sense of having coverage at least $(1-\alpha)$ and minimizing the maximal expected length. This interval is in fact randomized, and it is strictly better than either interval above. This paper also considered a single observation from a multivariate normal and showed that the set $\{\, ||\mu|| \,\leq\, c_\alpha ||X||^2 \,\}$ has coverage probability at least $(1-\alpha)$ for appropriately chosen c_α, and so is a legitimate confidence set for the multivariate mean.

These results are extended in various directions here. For a single $\mathcal{N}(\mu,\, \sigma^2)$ observation, we find a legitimate confidence interval for σ, and we note that these can be combined via the Bonferroni inequality to form a confidence set for $(\mu,\, \sigma)$ jointly. We also generalize beyond the normal distribution to symmetric unimodal distributions and to distributions with compact support. Finally, we extend the multivariate result to allow a sample $\{X_1,\, \cdots,\, X_m\}$ from the multivariate normal where m may be less than the dimension, p (and, hence, the covariance matrix would not be estimable). While the univariate results are mainly an intriguing curiosity, the multivariate result may actually be useful in some modern large data problems.

Finally, it may be noted that several probability paradoxes concern possible inference based on a single observation. One infamous example is the "Monte Hall" problem: a prize is placed behind one of three doors, and the contestant may choose one of the doors. After making a choice, the MC (Monte Hall) shows the contestant one of the other two doors which is empty and offers to let the contestant switch. Since at least one of the two doors must be empty, many (most?) people assume the MC offered no new information, and take the probability of winning to be the same (1/2) whether or not the contestant switches. A straightforward calculation provides the now well-known result that the probability of winning after switching is 2/3 (assuming all guessing is random).

Another example is the "two-envelope" problem: two players each receive an envelope: one containing the amount X and the other 2X . By turns, each player may either keep the amount received or switch envelopes. The conundrum is that if a player assumes the envelopes are equally likely, it is always best to switch, which seems paradoxical. However, having observed one value, the problem becomes essentially one of hypothesis testing based on a single observation, and the conditional probabilities will generally fail to be equally likely. Thus, the player must condition on the observed value, making the problem one of standard statistical inference, and not paradoxical. Portnoy (2020) provides a moderately complete treatment of this and related hypothesis testing problems. See also Wapner (2012) for a number of such examples.

2 The Normal Case

Let $X \sim N(\mu, \sigma^2)$, with $-\infty < \mu < \infty$, and $\sigma > 0$. As noted above, there are two classical confidence intervals for the mean, μ, based on the single observation, X: $I_1 = I_1(c) \equiv (X - c|X|,\, X + c|X|)$ (Abbott-Rosenblatt), and $\{I_2 = I_2(c) \equiv (-c|X|,\, c|X|)$ (Stein). The coverage probabilities may be found as direct corollaries of Theorem 1.1 of Portnoy (2019). The probabilities depend only on $\lambda = \mu/\sigma$, and the minimizing values λ^* are direct calculations.

Corollary 2.1. The coverage probability for I_1 depends only on λ, is symmetric about zero in μ (or in λ), and is given by

$$P_1(\lambda,\, c) = \Phi\left(\frac{c}{c+1}\,\lambda\right) + 1 - \Phi\left(\frac{c}{c-1}\,\lambda\right). \tag{1}$$

$P_1(\lambda,\, c)$ is minimized over λ at

$$\lambda^*(c) = \frac{c^2 - 1}{\sqrt{2}\,c^{3/2}}\sqrt{\log\frac{c+1}{c-1}}. \tag{2}$$

Thus, given $\alpha, 0 < \alpha < 0.5$, there exists a constant $c = c(\alpha)$, such that

$$\inf_{\mu,\sigma} P_{\mu,\sigma}(I_1(c(\alpha)) \geq 1 - \alpha$$

.

Corollary 2.2. The coverage probability for I_2 depends only on λ, is symmetric about zero in μ (or in λ), and is given by

$$P_2(\lambda,\, c) = \Phi\left(\frac{c-1}{c}\,\lambda\right) + 1 - \Phi\left(\frac{c+1}{c}\,\lambda\right). \tag{3}$$

$P_2(\lambda,\, c)$ is minimized over λ at

$$\lambda^*(c) = \left[\frac{c}{2}\,\log\frac{c+1}{c-1}\right]^{1/2}. \tag{4}$$

Thus, given $\alpha, 0 < \alpha < 0.5$, there exists a constant $c = c(\alpha)$, such that

$$\inf_{\mu,\sigma} P_{\mu,\sigma}(I_2(c(\alpha)) \geq 1 - \alpha$$

.

Note that since the minimal coverage probability depends only of the length of the interval (by invariance; see Portnoy (2019)), the minimal coverages for $I_1(c)$ and $I_c(c)$ are exactly the same (as functions of c).

Theorem 2.1. Given $\alpha, 0 < \alpha < 1$, there exists a constant $c = c(\alpha)$, such that

$$\inf_{\mu,\sigma} P_{\mu,\sigma}\left(\sigma^2 \leq \frac{X^2}{c^2}\right) = 1 - \alpha.$$

Furthermore, the constant $c = \Phi^{-1}(\frac{1+\alpha}{2})$, where $\Phi(.)$ denotes the standard normal CDF.

This follows since

$$P_{\mu,\sigma}\left(\sigma^2 \leq \frac{X^2}{c^2}\right) = P_{\mu,\sigma}\left(\frac{X^2}{\sigma^2} \geq c^2\right) \geq P_{\mu,\sigma}\left(\frac{(X-\mu)^2}{\sigma^2} \geq c^2\right) = P_{\mu,\sigma}(Z^2 > c^2)$$

where the inequality follows since a non-central chi-square has monotone likelihood ratio, and where Z is a standard unit normal.

Finally, note that intervals for μ and σ may be combined (via Bonferroni inequalities) to provide a rectangular simultaneous confidence set for $\{\mu,\, \sigma\}$. Of course, other simultaneous confidence sets can be constructed, and it seems clear that more circular sets will have smaller area. The problem of finding an minimax confidence set (in analogy with the result in Portnoy (2019)) seems extremely difficult, and it will not be pursued here.

3 More General Symmetric Unimodal Families

Theorem 3.1. Let $X \sim \frac{1}{\sigma} f_0(\frac{x-\mu}{\sigma})$, where $f_0(-z) = f_0(z)$ for all real z and $f_0(.)$ is strictly decreasing on $[0, \infty)$. Then, given $\alpha, 0 < \alpha < 1$, there exists $c = c(\alpha)$ such that

$$\inf_{\mu,\sigma} P_{\mu,\sigma}\left(\sigma \leq \frac{|X|}{c}\right) = 1 - \alpha.$$

Furthermore, this constant c satisfies

$$\int_0^c f_0(z)\,dz = \frac{\alpha}{2}. \tag{5}$$

Theorem 3.2. Let $X \sim \frac{1}{\sigma} f_0(\frac{x-\mu}{\sigma})$, where $f_0(.)$ is a continuous function and $f_0(-z) = f_0(z)$ for all real z. Let $F_0(.)$ denote the CDF corresponding to f_0, $F_0(x) = \int_{-\infty}^x f_0(z)dz$. Let $g(\theta|\alpha) = \alpha\, f_0(\alpha\,\theta), \alpha, \theta > 0$. Assume further that

(a) $g(\theta|\alpha)$ is strictly MLR in θ.

(b) For all α_1, α_2 with $\alpha_2 > \alpha_1, \log \frac{g(\theta|\alpha_2)}{g(\theta|\alpha_1)}$ is convex in θ.

Then,

(i) Given $c > 1$, there exists a unique root $\theta = \theta(c)$ of the equation

$$\frac{f_0(\frac{c}{c+1}\,\theta)}{f_0(\frac{c}{c-1}\,\theta)} = \frac{c+1}{c-1}. \tag{6}$$

(ii) $\theta(c)$ is continuous in c.

(iii) Moreover, for every $c > 1$,

$$\inf_{\mu,\sigma} P\Big(X - c|X| \leq \mu \leq X + c|X|\Big) = \psi(c) = F_0(\frac{c}{c+1}\,\theta(c)) + 1 - F_0(\frac{c}{c-1}\,\theta(c)). \tag{7}$$

(iv) $\psi(c)$ is continuous in c.

Example 3.1. The Cauchy Case: Suppose $X \sim C(\mu, \sigma)$, the Cauchy distribution with location parameter μ and scale parameter $\sigma, -\infty < \mu < \infty, \sigma > 0$. Therefore, $f_0(z) = \frac{1}{\pi\,(1+z^2)}$, and direct calculation gives that $\theta(c)$ of part (i) of Theorem 3.2 is given by $\theta(c) = \frac{\sqrt{c^2-1}}{c}, c > 1$. It follows that $\theta(c) \to 0$ as $c \to 1$ and $\theta(c) \to 1$ as $c \to \infty$, and, $\frac{c}{c+1}\,\theta(c) \to 0$ and $\frac{c}{c-1}\,\theta(c) \to \infty$ if $c \to 1$, while both $\frac{c}{c+1}\,\theta(c)$ and $\frac{c}{c-1}\,\theta(c) \to 1$ if $c \to \infty$. Together, these imply that $\psi(c)$, the infimum coverage probability of part (iii), equation (5), in Theorem 3.2 satisfies $\psi(c) \to 0.5$ as $c \to 1$ and $\psi(c) \to 1$ as $c \to \infty$. Hence, by the continuity of $\psi(c)$ (part (iv), Theorem 3.2)), given α such that $0.5 < 1 - \alpha < 1$, there is a $c = c(\alpha)$ such that $\psi(c) = 1 - \alpha$. Thus, in the Cauchy case, any nominal confidence level $1 - \alpha > .5$ can be exactly attained by a confidence interval of the form $X \pm c|X|$.

4 General Distributions with Compact Support

Theorem 4.1. Let $X \sim F$ and suppose that $P_F(a \leq |X| \leq b) = 1$, where $0 < a < b < \infty$. Let $\sigma^2 = \sigma^2(F) = \mathrm{Var}_F(X)$.
Let $K^2 = \frac{4}{(\frac{b}{a}+\frac{a}{b})^2}$, and $\alpha > 1 - K^2$. Then,

$$P_F\left(\sigma^2 \leq \frac{X^2}{c^2}\right) \geq 1-\alpha,$$

where $c^2 = 1 - \frac{\sqrt{1-\alpha}}{K}$.

5 A confidence set for μ based on a sample of size m from $\mathcal{N}_p(\mu,\, \Sigma)$

Theorem 5.1. Let $\{X_1,\, \cdots,\, X_m\}$ be a sample from $\mathcal{N}_p(\mu,\, \Sigma)$. Then to achieve

$$\inf_{\mu,\,\Sigma} P\left\{||\mu|| \leq \frac{c\,||X||}{\sqrt{m}}\right\} \geq 1-\alpha \tag{8}$$

it suffices to take $c = 3.85\,\alpha^{-1/(pm)}$.

6 Proofs

(Theorem 3.1). If $X \sim \frac{1}{\sigma}\, f_0(\frac{x-\mu}{\sigma})$, then $Y = \frac{X}{\sigma} \sim f_0(y-\theta)$ where $\theta = \frac{\mu}{\sigma}$, and hence, $|Y| \sim f_0(y-\theta) + f_0(y+\theta)$ under the assumptions made on $f_0(.)$. Therefore,

$$P_{\mu,\sigma}\left(\frac{|X|}{\sigma} \leq c\right) = P_{\mu,\sigma}\left(|Y| \leq c\right)$$

$$= \int_0^c [f_0(y+\theta) + f_0(y-\theta)]\, dy = \int_\theta^{\theta+c} f_0(z)\, dz + \int_{-\theta}^{c-\theta} f_0(z)\, dz$$

$$= \int_{\theta-c}^{\theta+c} f_0(z)\, dz \tag{9}$$

(since $f_0(-z) = f_0(z)$ for all z)

$$\leq \int_{-c}^{c} f_0(z)\, dz = 2\int_0^c f_0(z)\, dz.$$

(since $f_0(.)$ is strictly decreasing on $(0, \infty)$)
Therefore, if c is chosen such that $\int_0^c f_0(z)\,dz = \frac{\alpha}{2}$, then, we have

$$P_{\mu,\sigma}\left(\sigma \geq \frac{|X|}{c}\right) \leq \alpha \Rightarrow P_{\mu,\sigma}\left(\sigma \leq \frac{|X|}{c}\right) \geq 1-\alpha,$$

and the infimum of $P_{\mu,\sigma}\left(\sigma \leq \frac{|X|}{c}\right) = 1-\alpha$ by construction of c. This proves Theorem 3.1.

(**Theorem 3.2**). Following exactly the same lines as in Theorem 2.1, one has that

$$P_{\mu,\sigma,F_0}\left(\frac{|X|}{\sigma} \leq c\right) = F_0(\frac{c}{c+1}\,\theta) + 1 - F_0(\frac{c}{c-1}\,\theta), \tag{10}$$

where $\theta = \frac{\mu}{\sigma}$. The minimum must be at a critical point, which would satisfy

$$\frac{c}{c+1}\,f_0(\frac{c}{c+1}\,\theta) - \frac{c}{c-1}\,f_0(\frac{c}{c-1}\,\theta) = 0$$

$$\Leftrightarrow \frac{f_0(\frac{c}{c+1}\,\theta)}{f_0(\frac{c}{c-1}\,\theta)} = \frac{c+1}{c-1} \Leftrightarrow \frac{c}{c+1}\,f_0(\frac{c}{c+1}\,\theta) = \frac{c}{c-1}\,f_0(\frac{c}{c-1}\,\theta). \tag{11}$$

Since $g(\theta|\alpha) = \alpha\,f(\alpha\,\theta)$ is strictly MLR, it follows that (10) has at most one root. However, since for any $\alpha_1 < \alpha_2, \log\frac{g(\theta|\alpha_2)}{g(\theta|\alpha_1)}$ is convex, it follows that $\frac{f_0(\frac{c}{c+1}\,\theta)}{f_0(\frac{c}{c-1}\,\theta)} \to \infty$ as $\theta \to \infty$, and hence (10) must have a root. This establishes part (i) of Theorem 3.2. Continuity of this unique root, $\theta(c)$ follows from joint continuity of $\frac{c}{c+1}\,f_0(\frac{c}{c+1}\,\theta) - \frac{c}{c-1}\,f_0(\frac{c}{c-1}\,\theta)$ in c and θ, as $f_0(z)$ has been assumed to be continuous in z. The continuity of the infimum $\psi(c)$ follows from continuity of $F_0(.)$ and continuity of $\theta(c)$.

(**Theorem 4.1**). Since $0 \leq a \leq |X| \leq b < \infty$, by the reverse Cauchy-Schwarz inequality,

$$E(X^4) \leq \frac{(\frac{b}{a}+\frac{a}{b})^2}{4}\,[E(X^2)]^2 \Rightarrow [E(X^2)]^2 \geq K^2\,E(X^4), \tag{12}$$

with K defined as in the statement of the theorem.
On the other hand, since $E(X^2) = \mu^2 + \sigma^2$, for any $c, 0 < c < 1$, by the Paley-Zygmund inequality,

$$P(X^2 > c^2\,\sigma^2) \geq P(X^2 > c^2\,(\mu^2+\sigma^2)) \geq (1-c^2)^2\,\frac{[E(X^2)]^2}{E(X^4)} \geq (1-c^2)^2\,K^2. \tag{13}$$

Hence, if α is such that $1-\alpha < K^2$, then

$$P(\sigma^2 < \frac{X^2}{c^2}) \geq 1-\alpha, \tag{14}$$

with c^2 being chosen as $1 - \frac{\sqrt{1-\alpha}}{K}$.

(**Theorem 5.1**). Follow the proof of Theorem 3 in Portnoy (2018) almost exactly. Noting that $\sum_{j=1}^{m} ||X_j||^2$ is a non-central chi-square, the coverage probability (CP) of the set (8) can be written exactly as in equation 8 of Portnoy (2018):

$$CP = P\left\{||m\,\nu||^2 \leq c^2\,\chi^2_{pm+2K}\right\} \tag{15}$$

where K is Poisson with mean $||\nu||^2/2$. Note that the only difference here is the appearance of m.

The development and calculations in Portnoy (2018) now go through without any change except that p is replaced by mp. This again provides:

$$1 - CP \leq \alpha\, a^{-pm/2} \leq \alpha \tag{16}$$

where $a = 1/(1-\exp(-2\pi\, e^{p/4}+1))$, and where c is defined by

$$c^2 = 2e^2\,\alpha^{-2/p}\,a\ . \tag{17}$$

Theorem 5.1 follows trivially by dividing both sides of (15) by m.

References

[1] Abbott, J. H., and Rosenblatt, J. I. (1962), Two Stage Estimation with One Observation in the First Stage, it Annals of the Institute of Statistical Mathematics, 14, 229–235.

[2] Portnoy, S. (2020). The Two-Envelope Problem for General Distributions. *J Stat Theory Pract 14*, Article number 21.

[3] Portnoy, S. (2019) Invariance, Optimality, and a 1-Observation Confidence Interval for a Normal Mean, *The American Statistician, 73:1*, 10-15.

[4] Portnoy, S. (2018), Some Theorems on Optimality of a Single Observation Confidence Interval for the Mean of a Normal Distribution, *arXiv:* 1702.05545 [math.ST].

[5] Rodríguez C.C. (1996) Confidence Intervals from one Observation. In: Skilling J., Sibisi S. (eds) *Maximum Entropy and Bayesian Methods. Fundamental Theories of Physics* (An International Book Series on The Fundamental Theories of Physics: Their Clarification, Development and Application), vol 70. Springer, Dordrecht, 175–182.

[6] Wapner, Leonard M. (2012). Unexpected Expectations: The Curiosities of a Mathematical Crystal Ball, CRC Press: Taylor & Francis Group, Boca Raton, FL.