A Simple Confidence Interval for the Variance

David J. Olive*
Southern Illinois University

October 18, 2025

Abstract

Suppose $Y_1, ..., Y_n$ are independent and identically distributed with mean μ and variance σ^2. Let $\overline{Y}$ be the sample mean, and let $S^2 = S_Y^2$ be the sample variance of the Y_i. Then a $100(1-\alpha)\%$ confidence interval for the mean μ is $\overline{Y} \pm t_{n-1,1-\alpha/2} S/\sqrt{n}$. It will be shown that a $100(1-\alpha)\%$ confidence interval for the variance σ^2 is $S^2 \pm t_{n-1,1-\alpha/2} S_Z/\sqrt{n}$ where S_Z^2 is the sample variance of the $Z_i = Y_i(Y_i - \overline{Y})$.

KEY WORDS: t confidence interval.

1 INTRODUCTION

This section review some confidence intervals that have been suggest for the mean and the variance. Assume $Y_1, ..., Y_n$ are independent and identically distributed (iid) with mean μ and variance σ^2. Let z_δ be the δ percentile of the $N(0,1)$ distribution where $0 < \delta < 1$. Hence $P(Z \le z_\delta) = \delta$ if $Z \sim N(0,1)$. Similarly, let $t_{n-1,\delta}$ be the δ percentile of the t_{n-1} distribution. Hence $P(X \le t_\delta) = \delta$ if $X \sim t_{n-1}$. For a $100(1-\alpha)\%$ confidence interval (CI), take $\delta = 1 - \alpha$. Note that $t_{n-1,1-\alpha/2} > z_{1-\alpha/2}$, but $t_{n-1,1-\alpha/2} \to z_{1-\alpha/2}$ as $n \to \infty$.

Let the sample mean $\overline{Y} = \overline{Y_n} = \frac{1}{n}\sum_{i=1}^n Y_i$, let the sample variance

$S^2 = S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y})^2$, and let the method of moments estimator of the varaiance be $S_M^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^2$. Let the population skewness of the distribution be

$$\gamma = \frac{E[(Y_i - \mu)^3]}{\sigma^3}.$$

*David J. Olive is Professor, Mathematical & Statistical Sciences, Southern Illinois University, Carbondale, IL 62901, USA.

Let $\mu_3 = E[(Y-\mu)^3]$ and

$$\hat{\mu}_3 = S^3 \hat{\gamma} = \frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^3.$$

Let the population (excess) kurtosis of the distribution be

$$\kappa = \frac{E[(Y_i - \mu)^4]}{\sigma^4} - 3.$$

Let

$$\hat{\kappa} = \frac{\frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^4}{S^4} - 3.$$

Let

$$\hat{\psi} = \frac{\hat{\gamma}}{6\sqrt{n}}.$$

Then the large sample $100(1-\alpha)\%$ t CI for μ is

$$\overline{Y} \pm t_{n-1,1-\alpha/2} \;\; S/\sqrt{n}. \tag{1}$$

A competitor is the Johnson (1978) $100(1-\alpha)\%$ CI for μ

$$[\; \overline{Y} + \frac{\hat{\mu}_3}{6S^2 n} - t_{n-1,1-\alpha/2} \;\; S/\sqrt{n}, \;\; \overline{Y} + \frac{\hat{\mu}_3}{6S^2 n} + t_{n-1,1-\alpha/2} \;\; S/\sqrt{n} \;]. \tag{2}$$

Hesterberg (2014) gave the following two competitors of the t CI given by Equation (1.1): the skewness adjusted t interval is

$$[\; \overline{Y} + \frac{S}{\sqrt{n}}[\hat{\psi}(1+2t^2_{n-1,1-\alpha/2}) - t_{n-1,1-\alpha/2}], \; \overline{Y} + \frac{S}{\sqrt{n}}[\hat{\psi}(1+2t^2_{n-1,1-\alpha/2}) + t_{n-1,1-\alpha/2}] \;], \tag{3}$$

and the asymptotic percentile t CI is

$$[\; \overline{Y} + \frac{S}{\sqrt{n}}[\hat{\psi}(t_{n-1,1-\alpha/2} - 1)^2 - t_{n-1,1-\alpha/2}], \; \overline{Y} + \frac{S}{\sqrt{n}}[\hat{\psi}(t_{n-1,1-\alpha/2} - 1)^2 + t_{n-1,1-\alpha/2}] \tag{4}$$

The t–interval (1) may perform better than the three alternatives if the distribution has second moments but does not have third moments. In simulations, these confidence interval performed fairly well for a large variety of distribution with $n = 100$. The lognormal distribution needed $n = 400$. For any $n \geq 50$, distributions can be found where the CIs do not perform well for n but do perform well if the sample size is doubled to $2n$.

Next, we discuss confidence intervals for σ^2. Bickel and Doksum (2007, p. 279) suggest that

$$W_n = n^{-1/2}\left[\frac{(n-1)S^2}{\sigma^2} - n\right]$$

can be used as an asymptotic pivot for σ^2 if $E(Y^4) < \infty$. Notice that $W_n =$

$$n^{-1/2}\left[\frac{\sum (Y_i - \mu)^2}{\sigma^2} \; - \; \frac{n(\overline{Y} - \mu)^2}{\sigma^2} \; - \; n\right] =$$

$$\sqrt{n}\left[\frac{\sum\left(\frac{Y_i-\mu}{\sigma}\right)^2}{n} \;-\; 1\right] \;-\; \frac{1}{\sqrt{n}}\; n\; \left(\frac{\overline{Y}-\mu}{\sigma}\right)^2 = X_n - Z_n.$$

Since $\sqrt{n}Z_n \xrightarrow{D} \chi^2_1$, the term $Z_n \xrightarrow{D} 0$. Note that $U_i = [(Y_i - \mu)/\sigma]^2$ has mean $E(U_i) = 1$ and variance

$$V(U_i) = \tau = E(U_i^2) - (E(U_i))^2 = \frac{E[(Y_i-\mu)^4]}{\sigma^4} - 1 = \kappa + 2$$

where κ is the kurtosis of Y_i. Hence $X_n = \sqrt{n}(\overline{U} - 1) \xrightarrow{D} N(0, \tau)$ by the CLT. Thus $W_n \xrightarrow{D} N(0, \tau)$.

Hence

$$1 - \alpha \approx P(-z_{1-\alpha/2} < \frac{W_n}{\sqrt{\tau}} < z_{1-\alpha/2}) = P(-z_{1-\alpha/2}\sqrt{\tau} < W_n < z_{1-\alpha/2}\sqrt{\tau})$$

$$= P(-z_{1-\alpha/2}\sqrt{n\tau} < \frac{(n-1)S^2}{\sigma^2} - n < z_{1-\alpha/2}\sqrt{n\tau})$$

$$= P(n - z_{1-\alpha/2}\sqrt{n\tau} < \frac{(n-1)S^2}{\sigma^2} < n + z_{1-\alpha/2}\sqrt{n\tau}).$$

Hence a large sample $100(1-\alpha)\%$ CI for σ^2 is

$$\left(\frac{(n-1)S^2}{n + z_{1-\alpha/2}\sqrt{n\hat{\tau}}},\ \frac{(n-1)S^2}{n - z_{1-\alpha/2}\sqrt{n\hat{\tau}}}\right)$$

where

$$\hat{\tau} = \frac{\frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^4}{S^4} - 1.$$

Notice that this CI needs $n > z_{1-\alpha/2}\sqrt{n\hat{\tau}}$ for the right endpoint to be positive. It can be shown that $\sqrt{n}$ (length CI) converges to $2\sigma^2 z_{1-\alpha/2}\sqrt{\tau}$ in probability.

Olive (2014, pp. 276-278, 289-290) uses an asymptotically equivalent $100(1-\alpha)\%$ CI of the form

$$\left(\frac{(n-a)S^2}{n + t_{n-1,1-\alpha/2}\sqrt{n\hat{\tau}}},\ \frac{(n+b)S^2}{n - t_{n-1,1-\alpha/2}\sqrt{n\hat{\tau}}}\right)$$

where a and b depend on $\hat{\tau}$. The goal was to make a 95% CI with good coverage for a wide variety of distributions (with 4th moments) for $n \geq 100$. The price is that the CI is too long for some of the distributions with small kurtosis. The $N(\mu, \sigma^2)$ distribution has $\tau = 2$, while the EXP(λ) distribution has $\sigma^2 = \lambda^2$ and $\tau = 8$. The quantity τ is small for the uniform distribution but large for the lognormal LN(0,1) distribution.

By the binomial theorem, if $E(Y^4)$ exists and $E(Y) = \mu$ then

$$E(Y-\mu)^4 = \sum_{j=0}^{4} \binom{4}{j} E[Y^j](-\mu)^{4-j} =$$

$$\mu^4 - 4\mu^3 E(Y) + 6\mu^2(V(Y) + [E(Y)]^2) - 4\mu E(Y^3) + E(Y^4).$$

This fact can be useful for computing

$$\tau = \frac{E[(Y_i-\mu)^4]}{\sigma^4} - 1 = \kappa + 2.$$

2 Large Sample Theory for the New Confidence Interval

Part a) of the following theorem can be derived from the results in Bickel and Doksum (2007, p. 279). Theorem 1 with the V_i is a special case of Theorem 2 from Olive et al. (2025) which finds the limiting distribution of $\sqrt{n}(\hat{\boldsymbol{c}} - \boldsymbol{c})$ where $\hat{\boldsymbol{c}}$ stacks k distinct elements of the sample covariance matrix into the vector $\hat{\boldsymbol{c}}$. Note that the proof of Theorem 1 shows that if $W_i = Z_i + O_P(n^{-1/2})$, then $S^2_W = S^2_Z + O_P(n^{-1/2})$.

Theorem 1. Assume the cases $Y_1, ..., Y_n$ are iid with $E(Y_i) = \mu$, $V(Y_i) = \sigma^2$, and that $E(Y_i^4)$ exists. Let $W_i = (Y_i - \mu)^2$. Then a)

$$\sqrt{n}(\overline{W} - \sigma^2) \xrightarrow{D} N(0, \sigma^4(\kappa + 2)), \ \ \sqrt{n}(S^2 - \sigma^2) \xrightarrow{D} N(0, \sigma^4(\kappa + 2)), \tag{5}$$

$$\text{and} \ \ \sqrt{\mathrm{n}}(\mathrm{S}^2_{\mathrm{M}} - \sigma^2) \xrightarrow{\mathrm{D}} \mathrm{N}(0, \sigma^4(\kappa + 2)).$$

b) Let $Z_i = Y_i(Y_i - \overline{Y})$ and $V_i = (Y_i - \overline{Y})^2$. Then $S^2_W = S^2_Z + O_P(n^{-1/2}) = S^2_V + O_P(n^{-1/2})$. Denote the method of moments estimator of σ^2 by $\tilde{\sigma}^2$. Then $\tilde{\sigma}^2_W = \tilde{\sigma}^2_Z + O_P(n^{-1/2}) = \tilde{\sigma}^2_V + O_P(n^{-1/2})$.

Proof. a) $E(W) = \sigma^2$ and $V(W) = E(W^2) - [E(W)]^2 =$

$$(E[(Y - \mu)^4] - \sigma^4)\frac{\sigma^4}{\sigma^4} = \left(\frac{E[(Y - \mu)^4]}{\sigma^4} - 1\right)\sigma^4 = \sigma^4(\kappa + 2).$$

Hence $\sqrt{n}(\overline{W}_n - \sigma^2) \xrightarrow{D} N(0, \sigma^4(\kappa + 2))$ by the central limit theorem. By the delta method,

$$\frac{n(\overline{Y} - \mu)}{\sigma^2} \xrightarrow{D} \chi^2_1.$$

Then $nS^2_M = n\tilde{\sigma}^2_Y = \sum_{i=1}^n (Y_i - \overline{Y})^2 = \sum_{i=1}^n (Y_i - \mu + \mu - \overline{Y})^2 = \sum_{i=1}^n (Y_i - \mu)^2 +$ $2\sum_{i=1}^n (Y_i - \mu)(\mu - \overline{Y}) + \sum_{i=1}^n (\mu - \overline{Y})^2 = \sum_{i=1}^n W_i - n(\overline{Y} - \mu)^2$. Thus $S^2_M = \overline{W} - (\overline{Y} - \mu)^2$. Hence $\sqrt{n}(S^2_M - \sigma^2) =$

$$\sqrt{n}(\overline{W} - \sigma^2) - \frac{n(\overline{Y} - \mu)}{\sigma^2}\frac{\sigma^2}{\sqrt{n}} = \sqrt{n}(\overline{W} - \sigma^2) + O_P(n^{-1/2}).$$

Hence $\sqrt{n}(S^2_M - \sigma^2) \xrightarrow{D} N(0, \sigma^4(\kappa + 2))$.

b) $W_i = (Y_i - \overline{Y} + \overline{Y} - \mu)(Y_i - \overline{Y} + \overline{Y} - \mu) =$

$$V_i + (Y_i - \overline{Y})(\overline{Y} - \mu) + (\overline{Y} - \mu)(Y_i - \overline{Y}) + (\overline{Y} - \mu)(\overline{Y} - \mu).$$

Thus $W_i - \overline{W} = V_i - \overline{V} + a_i$ where

$$a_i = 2(Y_i - \overline{Y})(\overline{Y} - \mu) = O_P(n^{-1/2}).$$

Thus

$$\tilde{\sigma}^2_W = \frac{1}{n}\sum_{i=1}^n (W_i - \overline{W})^2 = \frac{1}{n}\sum_{i=1}^n (V_i - \overline{V})^2 + O_P(n^{-1/2}) = \tilde{\sigma}^2_V + O_P(n^{-1/2}).$$

Similarly, $W_i = (Y_i - \mu)(Y_i - \overline{Y} + \overline{Y} - \mu) = Y_i(Y_i - \overline{Y}) + Y_i(\overline{Y} - \mu) + \mu(Y_i - \mu) = Z_i + (Y_i - \mu)(\overline{Y} - \mu) = Z_i + a_i = Z_i + O_P(n^{-1/2})$. Thus

$$\frac{1}{n}\sum_{i=1}^n (W_i - \overline{W})^2 = \frac{1}{n}\sum_{i=1}^n (Z_i + a_i - (\overline{Z} + \overline{a}))^2 = \frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z} + a_i - \overline{a})^2 =$$

$$\frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z})^2 + \frac{2}{n}\sum_{i=1}^n (Z_i - \overline{Z})(a_i - \overline{a}) + \frac{2}{n}\sum_{i=1}^n (a_i - \overline{a})^2 =$$

$$\frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z})^2 + \frac{1}{n} n O_P(1) O_P(n^{-1/2}) + \frac{1}{n} n O_P(n^{-1/2}) O_P(n^{-1/2}) = \frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z})^2 + O_P(n^{-1/2}).$$

$\square$

Hence a $100(1-\alpha)\% CI$ for σ^2 is

$$S^2 \pm t_{n-1,1-\alpha/2} S_Z/\sqrt{n} \tag{6}$$

where S_Z^2 is the sample variance of the $Z_i = Y_i(Y_i - \overline{Y})$. If the CI is $[L_n, U_n]$, then $[max(0, L_n), U_n]$ is shorter with the same coverage since $\sigma^2 \geq 0$.. S_Z can be replaced by S_V, but S_V tends to be larger than S_Z in small samples, and hence gives better coverage.

3 Conclusions

Outliers affect the confidence interval (6) even more that outliers effect the t CI for μ. It is useful to check for outliers by making a dot plot of the data.

McKinney (2021) gave some more competitors for the t CI for μ. The Johnson (1978) CI (2) appeared to be best, but only very slightly better than the usual t CI (1).

Simulations were done in *R*. See R Core Team (2024). The collection of *R* functions *lspack*, available from (http://parker.ad.siu.edu/Olive/lspack.txt), has some useful functions. The function `varci` computes the CI (6). The function `varcisim3` simulates CI (6). The function `varcisim` was used to produce Table 1.

References

Bickel, P.J., and Doksum, K.A. (2007), *Mathematical Statistics: Basic Ideas and Selected Topics,* Vol. 1., 2nd ed., Updated Printing, Pearson Prentice Hall, Upper Saddle River, NJ.

Hesterberg, T., (2014), "What Teachers Should Know about the Bootstrap: Resampling in the Undergraduate Statistics Curriculum," available from (http://arxiv.org/pdf/1411.5279v1.pdf).

Hesterberg, T., (2015), "What Teachers Should Know about the Bootstrap: Resampling in the Undergraduate Statistics Curriculum," *The American Statistician*, 69, 371-386.)

Johnson, N.J. (1978), "Modified t Tests and Confidence Intervals for Asymmetrical Populations," *Journal of the American Statistical Association*, 73, 536-544.

McKinney, D. (2021), "Some t–Type Confidence Intervals," MS paper, Southern Illinois University, at
(https://opensiuc.lib.siu.edu/cgi/ viewcontent.cgi?article=2403&context=gs_rp).

Olive, D.J. (2014), *Statistical Theory and Inference*, Springer, New York, NY.

Olive, D.J., Alshammari, A.A., Pathiranage, K.G., and Hettige, L.A.W. (2025), “Testing with the One Component Partial Least Squares and the Marginal Maximum Likelihood Estimators,” *Communications in Statistics: Theory and Methods*, to appear. https://doi.org/10.1080/03610926.2025.2527340