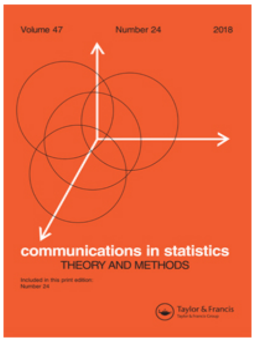

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

A report on the paper “Sungsu Kim. 2019. The probable error in the hypothesis test of normal means using a small sample. *Communications in Statistics - Theory and Methods*. DOI: 10.1080/03610926.2019.1703135.”

Kalimuthu Krishnamoorthy & Arvind K. Shah

To cite this article: Kalimuthu Krishnamoorthy & Arvind K. Shah (2020): A report on the paper “Sungsu Kim. 2019. The probable error in the hypothesis test of normal means using a small sample. *Communications in Statistics - Theory and Methods*. DOI: 10.1080/03610926.2019.1703135.”, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2020.1767786

To link to this article: https://doi.org/10.1080/03610926.2020.1767786

Published online: 18 May 2020.

https://doi.org/10.1080/03610926.2020.1767786

Taylor & Francis
Taylor & Francis Group

LETTER TO THE EDITOR

A report on the paper "Sungsu Kim. 2019. The probable error in the hypothesis test of normal means using a small sample. *Communications in Statistics - Theory and Methods*. DOI: 10.1080/03610926.2019.1703135."

Let $X_1, \ldots, X_n$ be a sample from a $N(\mu, \sigma^2)$ distribution. Consider testing

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_a : \mu > \mu_0$$

The usual t-test is based on the statistic

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

where

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2$$

Kim (2019) has considered the test statistic

$$K_1 = \frac{\bar{X} - \mu_0}{S_0/\sqrt{n}}, \quad \text{with} \quad S_0^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \mu_0)^2$$

and he has concluded that the test based on K_1 is more powerful than the usual t-test based on some graphical and simulation studies.

On the contrary to Kim's claim, the usual t-test and the test based on the K_1 statistic are equivalent, because K_1 can be expressed as a one-to-one function of t. His claims regarding two-sample t test and the test for equality of several means are also wrong.

Here are some details and specific comments on the paper.

1. First, the usual t test is well known and routinely used to test the mean of a normal population. This test is so popular and described in almost all introductory level text books in statistics.
2. It is also known that the t test is uniformly most powerful unbiased (UMPU) test for one-sided and two-sided hypotheses.
3. Lefante and Shah (1986) have shown that the tests based on the K_1 statistic and the t statistic are equivalent. The proof is straightforward and easy to understand.

It should be noted that Kim is aware of the result due to Lefante and Shah, and he cited their paper. Given the above results and facts, it is impossible to find a more powerful test than the t test.

Student's t test and the new test

We first noticed that Kim somehow thinks that if two tests based on different statistics are equivalent, then their null distributions should be the same. This is evident from his analysis of Figure 1 in his paper. However, this is not true. For example, for testing $H_0 : \mu = \mu_0$ vs. $H_a : \mu \neq \mu_0$, the tests based on the statistics

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \quad \text{and} \quad \frac{n(\bar{X} - \mu_0)^2}{S^2}$$

are the same, but the null distribution of the first statistic is t_{n-1}, and the null distribution of the second statistic is $F_{1,n-1}$.

Secondly, we note that Lefante and Shah (1986) have shown that the statistic K_1 can be expressed as a one-to-one function of t statistic as

$$K_1 = K_1(t) = \frac{\sqrt{n}t}{(n - 1 + t^2)^{1/2}} \tag{1}$$

In particular, $K_1(t)$ is an increasing function of t. Thus, if t_q denotes the $100q$ percentile of the t statistic, then $K_1(t_q)$ is the $100q$ percentile of the K_1 statistic. For a given $(n, \bar{x}, s)$, let t_0 be the observed value of the t statistic. Then the observed value of K_1 based on the same $(n, \bar{x}, s)$ is $K_1(t_0)$. For a given level of significance α, let $t_{1-\alpha}$ denotes the upper α quantile of the t distribution with df $= n - 1$. Then

$$t_0 > t_{1-\alpha} \quad \text{if and only if} \quad K_1(t_0) > K_1(t_{1-\alpha})$$

In other words, whenever the t test rejects the H_0 at the level α, the test based on the K_1 statistic also rejects the H_0 at the level α, and vice versa. Thus, the rejection regions of both tests are the same.

It appears that Kim is aware of the fact that these two tests are the same or knows that the one-to-one relation can be used to find the quantiles of K_1 from those of the t distribution. To show some evidence, let us consider the write-up in the second paragraph of page 5 in Kim's paper: "*One may notice that the null distribution(s) of the ratio in the new method is (are) difficult to derive due to the dependency between the numerator and the denominator. The quantiles of the distribution in K_1 are provided in* Table 4." Even though Kim noted that the null distribution of K_1 is difficult to obtain, he has calculated the exact percentiles of K_1, without giving any calculation details, and tabulated in Table 4 of his paper. Our investigation showed that the quantiles of K_1 in Table 4 of Kim's paper can be calculated using the formula

$$K_{1,q} = K_1(t_q) = \frac{\sqrt{n}t_q}{\left(n - 1 + t_q^2\right)^{1/2}}, \quad 0 < q < 1$$

which follows from (1), and is also noted in Lefante and Shah (1986). In the above formula, t_q is the $100q$ percentile of the t distribution with df $= n - 1$. For example, when $n = 2$, $t_{.95} = 6.313752$ and using this number in the above formula, we find $K_{1,.95} = 1.39680$. We calculated the quantiles of K_1 for all values considered in Table 4 of Kim's paper, and they are exactly the same as those reported in Table 4 of Kim's paper. If Kim had used a different formula, other than the above one, to compute the quantiles of K_1, he should have made it available to readers.

Confidence intervals

Since the t test and the test based on the K_1 statistic are the same, the confidence intervals (CIs) that are obtained by inverting these two tests should be the same. On the basis of a flawed argument, Kim claims that the CI based on the t test is wider than the confidence interval based on K_1 statistic, the one in (6) of his paper. The CI for μ in (6) involves a typo, and after fixing the typo, it can be expressed as

$$\left(\bar{x} - K_{1;1-\alpha/2}\frac{1}{\sqrt{n}}\sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{n}},\ \bar{x} + K_{1;1-\alpha/2}\frac{1}{\sqrt{n}}\sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{n}}\right)$$

What is μ_0 here? If μ_0 is the true value of the mean, then why do you want to estimate? This CI is just illogical.

Furthermore, Kim's (2019) claims regarding the improvements of the two-sample t test and the test for equality of several normal means are also wrong. Indeed, Shah and Krishnamoorthy (1993) have shown that the modified tests based on the hypothesis dependent variance estimates and the existing classical tests are equivalent.

Acknowledgement

The authors are thankful to a reviewer for providing some useful comments and suggestions.

References

Lefante, J. J., and A. K. Shah. 1986. A note on one sample t-test. *Journal of Statistical Computation and Simulation* 25 (3–4):295–6. doi:10.1080/00949658608810938.

Shah, A. K., and K. Krishnamoorthy. 1993. Testing means using hypothesis-dependent variance estimates. *The American Statistician* 47 (2):115–7. doi:10.2307/2685190.

Kalimuthu Krishnamoorthy
Department of Mathematics,
University of Louisiana at Lafayette, Lafayette, LA 70508-1010, USA
krishna@louisiana.edu

Arvind K. Shah
Biostatistics and Research Decision Sciences,
Merck Research Laboratories, Rahway, New Jersey, USA