10.12663/PSYHUNG.7.2019.2.13.

ESTIMATION METHODS ON STANDARD ERROR OF DIFFERENT STATISTICAL PARAMETERS

Judit T. Kárász[1]

Correspondence Author: Judit T. Kárász (takacsne.karasz.judit@oh.gov.hu)

Abstract

In a statistical process it is important to make estimation on the standard error of our calculated parameters. There are many different algorithms to do it. In our article, we discuss closed formulas on some parameters. We describe the bootstrap and the jackknife general simulation algorithms. We describe the Fay's BRR jackknife method too. Finally, we describe a mixed bootstrap-jackknife algorithm which applied in the National Assessment of Basic Competencies (NABC).

Keywords: estimation on standard error ▪ bootstrap ▪ jackknife

PROLOGUE

In statistical calculations there are many types of errors: error of the sample, error between subject or error of the applied statistical method (e.g. in approximation). The standard error is one of the typical statistical errors: this is the average error of the parameters evaluation.

In the first chapter we show some parameter's standard error's formula (mean, standard deviation, skewness, kurtosis). In the second chapter we describe the jackknife algorithm and its generalization. After this, we describe Fay's BRR jackknife method (Fay, 1984), (Dippo, Fay, & Morganstein, 1984).

In the last two chapters we describe the bootstrap algorithm and the combination of the jackknife and bootstrap methods which we applied in the National Assessment of Basic Competencies (NABC).

[1] Educational Authority, 1055 Budapest, Szalay Street 10-14.

THE FORMULAS OF THE STANDARD ERROR

When we have independent and identically distributed (i.i.d.) sample we can use some close formulas on the standard error of some parameters, like the sample's mean, standard deviation or the skewness and the kurtosis. In this chapter, we present the schemes of the standard errors of the following parameters mean, standard deviation, skewness and kurtosis (Lehmann & Casella, 1998), (Takács, 2010), (Takács, 2012).

If we have an i.i.d. sample: The parameters can be calculated by the following formulas – mean ($\bar{X}$), standard deviation (s), skewness (SK) and kurtosis (K):

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n},$$

$$s = \sqrt[2]{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}},$$

$$SK = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^3}{s^3},$$

$$K = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^4}{s^4} - 3.$$

If we used an i.i.d. sample the standard errors of the parameters are these (Lehmann & Casella, 1998):

$$S_{\bar{X}} = \frac{s}{\sqrt{n}},$$

$$S_{s^2} = s^2\sqrt{\frac{2}{n-1}},$$

With larger sample size (more than 100-200 cases) for the standard error of the standard deviation when we can use approximated calculation:

$$S_s = s\frac{1}{\sqrt{2(n-1)}}.$$

The standard error of the skewness and kurtosis are calculated iteratively from this scheme:

$$S_{skewness} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}},$$

$$S_{kurtosis} = 2S_{skewness}\sqrt{\frac{n^2-1}{(n-3)(n+5)}}.$$

Unfortunately, in the most cases we can not use i.i.d. sample (for example in the National Assessment of Basic Competencies (NABC) the children in each class are connected – so we can not say that this is an independent sample). When we haven't got i.i.d. sample we should use other way to calculate the standard error. In these cases we can choose a simulation method. But we use these formulas for the following calculation: for each parameters we would like to calculate a confidence interval (for example with 95% reliability or in other words, 95% significance level).

On the following chapters we show how we can calculate a confidence interval with specific simulation methods.

JACKKNIFE METHOD

The jackknife method (Efron & Gong, 1983) is the easiest method to fix the dependency in the sample in an estimation. We can use it in every parameter estimation.

We have a sample: $X_1, X_2, \dots, X_n$. We can calculate many parameters from the sample:

$$P(n) = P(X_1, \dots, X_n)$$

We define some more parameters like this:

$$P_i = P(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n).$$

So, we can calculate with the same calculation method „n" number of parameters with the same sample without the „ith" case. Generally, we calculate a weighted parameter with weight 1 on „n-1" case and weight 0 on the „ith" case. But if we use a weighted parameter calculating method we can define many other weights.

For example we can say the following: we use weight (1) on „n-2" cases and use (0) on two cases. Or weight (1) on „n-2" cases and weight (1/2) on two cases, etc.

With these methods there will be many possibilities to calculate the deterministic way many P_i parameters. So, we can order them in ascending order. We trim (for example in 95% level) the upper and lower 2,5% P_i parameters – and we have got the confidence interval with a deterministic (so reproducible) way.

The problem with this method is the following: the jackknife sample is not an i.i.d. sample. But with a simple mathematical trick this problem can be solved. (Miller, 1974). We can define pseudo-parameters with this form:

$$P_i^* \coloneqq nP(n) - (n-1)P_i$$

These pseudo-parameters are almost i.i.d (minimal dependency between them) so we can use them as a new, i.i.d. sample on our parameter.

FAY'S BRR METHOD

The method (BRR – balanced repeated replication) of Fay (Fay, 1984), (Dippo et al., 1984), (Takács, 2010) is analogue with jackknife method. Fay's original method is when we use weight cases with weight (1) on "n-2" cases and (1/2) and (3/2) on two other cases. We use all of the possible pairs and calculate a pseudo-parameter with this weighted method.

Imagine this: in this method every case can be the „pair" of any other case (cases who are similar to them, with whom it has got dependency). But if we imagine for example a school with classes we find that a student in a class can be connected only with one of his or her classmates.

So we will define stratum variables: these stratum variables can be the „gender" – boys only can be connected and replaced with other boys. So the weight depends on the rule of the replacement. Balance is the weight (like on jackkinfe), replace is the stratum variable in the method and repeat means that we will calculate enough pseudo-parameters for the easiest estimation on the confidence interval.

In this case, we define BRR weights for any cases (random, convex combination in each stratum, so the sample size with the weight will be „n") – and in this way this method will be reproducible.

BOOTSTRAP METHOD

The bootstrap method (Efron & Gong, 1983) is based on the jackknife method too.

So, in the jackknife method we use weighted cases with (1,0) weights. When we use the BRR jackknife method we should use stratum variables which variables define groups (with smaller sample size) and we define random weights within the groups – so these are convex combinations of the smaller groups.

In the bootstrap method we do not define weights apriori. We define the repeated measure size (by Efron and Gong the minimal repeated measure is 300-500 repeated samples). So we applied 300-500 resamples with a return packing sample method (imagine that our sample is in a box and we choose one, put it back, choose another one, put it back, etc, etc ect, ect). We do this „n" times in each 300-500 resample.

In each resample we calculate a pseudo-parameter – so from these pseudo-parameters we can create the confidence interval.

CONFIDENCE INTERVAL CALCULATING METHOD IN NABC

In the National Assessment of Basic Competencies (NABC) we combined the BRR jackknife method with the bootstrap method.

The NABC is a whole population survey so there are only a few students who do not write the competencies' tests. In this case, the stratum variable in this survey is the „classroom" so we can connect every student with one of his or her classmates. This will be the rule of the resample replace.

In the second step we define random weights in the classroom with these two rules:

- We replace every student only with his or her classmate;
- We replace every missing student with his or her classmate;

So each student has a random weight and if we have a class with C students, than the final sum of the weights of the students who write the tests will be C. It means that we define the convex combination of the student's weights who wrote the tests – and multiply it „up" with the student number of the class.

Finally, we define 101 different replicant weights with this method (we use these random weights as BRR replicant weights for 101 repeated sample). We calculate with this 101 weights 101 pseudo-parameters and we trim the lowest 3 and highest 3 ones – so we will get the „middle" 95 ones. This will define us the confidence interval.

In this paper, we showed, that a constructive bootstrap method, which takes the population specializations in count, can be a better approach.

The next table contains three different standard error calculation results on the 6th grade NABC mathematics and reading achievement for the whole population, girls and boys. The average achievement point is the same for all three methods, but standard errors are quite different. All the calculations were made in IBM SPSS®. In the first case, we used the first discussed closed formula. The second column contains the results of the IBM SPSS® built-in bootstrap method. The last ones were calculated by our modified bootstrap method, developed for the NABC, which uses class-based weights and it takes the characteristic of each school into account.

Table 1: Standard error on average achievement, 6th grade

Achievement point	Closed formula	Bootstrap	NABC method
Mathematics, mean	0,637	0,629	0,429
Reading, mean	0,682	0,683	0,545
Mathematics, girls	0,865	0,941	0,683
Reading, girls	0,952	1,021	0,732
Mathematics, boys	0,934	0,966	0,720
Reading, boys	0,960	0,949	0,823

The level of the closed formula results and the IBM SPSS® bootstrap results are similar, usually, but not always, the bootstrap is higher. But we accept the NABC method results, because the stratums represent the dependency between the students.

It is important that the IBM SPSS® built-in bootstrap method does not work with weights other than integers, so technically we can not make the calculation analog to the NABC method.

SAMPLES IN PLANNED ISSUE

These are the applications of the method above in this issue. Nyitrai and colleagues researched the parental involvement in two papers (Nyitrai et al., 2019a), and (Nyitrai et al., 2019b) used it to calculate confidence intervals on case numbers and average achievement. Different modes of involvement show different achievement levels and different group sizes. Koltói and colleagues discussed the parental involvement into school life (instead of involvement into studying) in (Koltói et al., 2019a) and (Koltói et al., 2019b). Similarly, they calculated confidence intervals on average achievement levels and group sizes on both mathematics and reading.

Beside parental involvement, we researched the impact of family background on achievement. Harsányi and colleagues' papers (Harsányi et al., 2019a), and

(Harsányi et al., 2019b) used the confidence intervals of case numbers and average achievements (calculated as above) of different family indicators to survey the family background impact on mathematical an reading achievement.

Two more papers discuss achievements of different student groups by comparing them along estimated group sizes and ratios. One of the groups is the regular athlete students, surveyed by Smohai and colleagues (Smohai et al, 2019a) (and Smohai et al., 2019b). The other group is the students with special needs (SNI) and BTM students, discussed by Kovács and colleagues in (Kovács et al., 2019) and Kövesdi and colleagues in (Kövesdi et al., 2019) by comparing group ratios and achievements by confidence intervals.

REFERENCES

Dippo, C. S., Fay, R. E., & Morganstein, D. H. (1984). Computing Variances from Complex Samples with Replicate Weights. *American Statistical Association*, 6.

Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, *37*(1), 36–48. https://doi.org/10.2307/2685844

Fay, R. E. (1984). Some Properties of Estimates of Variance Based on Replication Methods. *American Statistical Association*, 495–500.

Lehmann, E. L., & Casella, G. (1998). *Theory of point estimation* (2nd ed). New York: Springer.

Miller, R. G. (1974). The Jackknife–A Review. *Biometrika*, *61*(1), 1–15. https://doi.org/10.2307/2334280

Takács S. (2010). Egy nem hagyományos statisztikai eljárás bemutatása az OECD Pisa adatbázison – Esettanulmány. *Alkalmazott Matematikai Lapok*, *27*(1), 157–174.

Takács S. (2012). Érzékenységvizsgálatok a statisztikai eljárásokban. *Alkalmazott Matematikai Lapok*, 34.

Citations of this tematical edition

Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi And-rea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Születni tudni kell? – Az Országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszeressége, észlelt társadalmi helyzet és a lakókör-nyezet vonatkozásában

Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Smohai Máté, Simon Gabriella, Takács Nándor, Takács Szabolcs: The relationship of school achievement with parents’ employment status, perceived social status, and living environment as reflected in findings of the 2017 National Assessment of Basic Competencies (NABC)

Koltói Lilla, Harsányi Szabolcs Gergő, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: A szülők tanulmányokba való bevonódásának összefüggése az iskolai teljesítménnyel

Koltói Lilla, Harsányi Szabolcs Gergő, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: The relationship between school achievement and paternal involvement in children's school activities as judged by headmasters in the 2017 National Assessment of Basic Competencies (NABC)

Kovács Dóra, Kövesdi Andrea, Harsányi Szabolcs Gergő, Koltói Lilla, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Results of the 2017 National Assessment of Basic Competencies (NABC) –amongst 6th 8th and 10th grade pupils diagnosed with SEN and BTM disorder

Kövesdi Andrea, Kovács Dóra, Harsányi Szabolcs Gergő, Koltói Lilla, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: A 2017. évi Országos kompetenciamérés eredményei Magyarországon – Az SNI és BTM-mel diagnosztizált 6, 8, 10-ik évfolyamos gyermekek körében

Nyitrai Erika, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Iskolai teljesítmény és szülői bevonódottság

Nyitrai Erika, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Relations between Parental Invol-vement and School Performance in the Light of Data from National Assessment of Basic Competencies (NABC) 2017

Smohai Máté, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Takács Nándor, Takács Szabolcs: Az Országos kompetenciamérés során felvett szabadidős sporttevékenységre irányuló adatok elemzése és standardjainak közlése

Smohai Máté, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Pulai-Kottlár Gabriella , Simon Gabriella, Takács Nándor, Takács Szabolcs: Relationship between leisure & sports activities and school performance in the light of data from the National Assessment of Basic Competencies (NABC) 2017 survey