VERSITA

TATRA MOUNTAINS
Mathematical Publications
DOI: 10.2478/v10127-011-0034-0
Tatra Mt. Math. Publ. **50** (2011), 25–38

A POINT ESTIMATOR OF THE PROBABILITY DIFFERENCE OF OVERALL TREATMENT EFFECT– –SIMULATION STUDY

Pavla Krajíčková

ABSTRACT. One of the aims of the meta-analysis of clinical trials is to determine the efficacy of a new type of treatment. This efficacy is commonly measured by the difference between the efficacy of a standard treatment and the new treatment. For binary data the difference can be measured by a probability difference. We investigate, by simulations and using box plots, the basic statistical properties of the point estimator of the probability difference of overall treatment effects in the meta-analysis based on multicentre trials for various chosen situations. This estimator was suggested in Dokoupilova (2011).

1. Introduction

Let us consider a clinical trial performed in I centers. Suppose that the number of subjects included in the trial in the ith center is $n_{T,i} + n_{C,i}$ for $i = 1, 2, \ldots, I$, where $n_{T,i}$ is the number of subjects who get new treatment and $n_{C,i}$ is the number of subjects in the control group, i.e., subjects who are treated by a standard treatment or a placebo. Subjects in the treated group in the ith center succeed with probability $p_{T,i}$ and subjects in the control group in the ith center succeed with probability $p_{C,i}$ for $i = 1, 2, \ldots, I$. All subjects are considered to be independent. One of the principal questions is: What is the difference between the efficacy of the standard treatment or the placebo and the new treatment?

Let a random variable $X_{T,i}$ denote the number of successes in the treated group in the ith center and the number of successes in the control group in the ith center is denoted by a random variable $X_{C,i}$. Then $X_{T,i}$ has binomial

2010 Mathematics Subject Classification: 62P10.

Keywords: multicenter trial, point estimator, probability of success, probability difference, linear model with random effects.

Supported by the Ministry of Education, Youth and Sports of the Czech Republic under the project MŠMT LC06024.

distribution with the sample size $n_{T,i}$ and the probability of success $p_{T,i}$ and also $X_{C,i}$ has binomial distribution with the sample size $n_{C,i}$ and the probability of success $p_{C,i}$. Random variables $X_{T,1}, \dots, X_{T,I}$, $X_{C,1}, \dots, X_{C,I}$ are stochastically independent. We also work with random variables

$$Y_{l,i} = \frac{X_{l,i}}{n_{l,i}}, \qquad \text{for } l \in \{T, C\} \quad \text{and} \quad i = 1, \dots, I. \tag{1}$$

2. The linear model with random effects

We suppose that the true probabilities of success in the ith center $p_{T,i}$ and $p_{C,i}$ randomly fluctuate around common probabilities of success p_T and p_C, i.e.,

$$p_{l,i} = p_l + b_{l,i}, \qquad \text{for } l \in \{T, C\} \quad \text{and} \quad i = 1, \dots, I, \tag{2}$$

where $b_{l,i}$ is a random effect of the ith center. Further, we suppose that $b_{l,i}$ is normally distributed with the mean 0 and the variance $\sigma^2_{l,0}$, i.e., $b_{l,i} \sim N(0, \sigma^2_{l,0})$[1].

The final situation can be represented by a linear model with random effects

$$Y_{l,i} = p_l + b_{l,i} + \varepsilon_{l,i} \qquad \text{for } l \in \{T, C\} \quad \text{and} \quad i = 1, \dots, I, \tag{3}$$

where $\varepsilon_{l,i}$ are error terms and we suppose that $\varepsilon_{l,i} \sim N(0, \sigma^2_{l,i}/n_{l,i})$.

We want to estimate the probability difference $p_T - p_C$. Dokoupilová (2011) described the construction of the point estimator of the difference between common probabilities of success

$$\widehat{p_T - p_C} = \hat{p}_T - \hat{p}_C, \tag{4}$$

where the estimators $\hat{p}_T$ and $\hat{p}_C$ are elements of the point estimator of the vector of the successful treatment common probabilities

$$\begin{pmatrix} \hat{p}_T \\ \hat{p}_C \end{pmatrix} = \begin{pmatrix} \dfrac{\sum_{i=1}^{I} \frac{(X_{T,i}}{n_{T,i}\hat{\sigma}^2_{T,0} + \hat{\sigma}^2_{T,i}}}{\sum_{j=1}^{I} \frac{n_{T,j}}{n_{T,j}\hat{\sigma}^2_{T,0} + \hat{\sigma}^2_{T,j}}} \\ \\ \dfrac{\sum_{i=1}^{I} \frac{X_{C,i}}{n_{C,i}\hat{\sigma}^2_{C,0} + \hat{\sigma}^2_{C,i}}}{\sum_{j=1}^{I} \frac{n_{C,j}}{n_{C,j}\hat{\sigma}^2_{C,0} + \hat{\sigma}^2_{C,j}}} \end{pmatrix} \tag{5}$$

(see also Wimmer and Witkovský (2004)). Wimmer and Witkovský suggested to use in (5) Agresti and Caffo estimator

$$\hat{\sigma}^2_{l,i} = \frac{X_{l,i} + 2}{n_{l,i} + 4}\left(1 - \frac{X_{l,i} + 2}{n_{l,i} + 4}\right) \qquad \text{for } l \in \{T, C\} \quad \text{and} \quad i = 1, \dots, I$$

[1]Of course it is supposed that $\sigma^2_{l,0}$ is such that "practically" $0 < p_l + b_{l,i} < 1$. It is ensured with a proper choice of $\sigma^2_{l,0}$ in simulations.

(see Agresti and Caffo (2000)) and the Mandel and Paule procedure to estimate the unknown variances $\sigma^2_{l,0}$. The estimator $\hat{\sigma}^2_{l,0}$ for $l \in \{T, C\}$ can be obtained as iterative solution of the following equations

$$\hat{\mu}_l^{MP} = \frac{\sum_{i=1}^{I} \frac{X_{l,i}}{n_{l,i}\hat{\sigma}^2_{l,0} + \hat{\sigma}^2_{l,i}}}{\sum_{j=1}^{I} \frac{n_{l,j}}{n_{l,j}\hat{\sigma}^2_{l,0} + \hat{\sigma}^2_{l,j}}},$$

$$\sum_{i=1}^{I} \frac{\left(\frac{X_{l,i}}{n_{l,i}} - \hat{\mu}_l^{MP}\right)^2}{\hat{\sigma}^2_{l,0} + \frac{\hat{\sigma}^2_{l,i}}{n_{l,i}}} = I - 1 \tag{6}$$

(see Paule and Mandel (1982)).

The estimator (5) is the generalized least squares estimator in the linear model with random effects (3), where the unknown parameters $\sigma^2_{l,0}$, $\sigma^2_{l,i}$ are replaced by their estimates given in Agresti and Caffo [1] and Paule and Mandel [3], respectively. Estimators $\hat{\sigma}^2_{l,0}$ and $\hat{\sigma}^2_{l,i}$ are biased, but asymptotically unbiased. So the estimator (5) is biased but asymptotically unbiased.

The estimates of $\sigma^2_{l,0}$ can be considered as an approximation to maximum likelihood estimates, but they are computationally simpler. The left side of the equation (6) is according to Paule and Mandel (1982) with probability one a monotone decreasing function. In simulations, the stopping rule was difference between estimates in particular iteration lower than 0.000 001.

3. Simulation results

We conducted a simulation study to explore the behavior of the point estimator of the difference between common probabilities of success (4). We run this study for three (main) different settings. Assuming the model is true, the values of unknown parameters I, $n_{T,i}$, $n_{C,i}$, $\sigma^2_{l,0}$ ($l \in \{T, C\}$), p_T and p_C were following:

- the number of centers: $I \in \{5, 10, 15, 25, 35\}$,
- the number of subjects: $n_{T,i}, n_{C,i} \in \{5, 10, 15, 30, 50, 100\}$,
- the variance of random effects:

$$\sigma^2_{l,0} \in \left\{0, \frac{1}{4}\left(\frac{p_l}{3}\right)^2, \frac{1}{2}\left(\frac{p_l}{3}\right)^2, \frac{3}{4}\left(\frac{p_l}{3}\right)^2, \left(\frac{p_l}{3}\right)^2\right\} \quad \text{for } p_l \le 0.5$$

and

$$\sigma^2_{l,0} \in \left\{0, \frac{1}{4}\left(\frac{1-p_l}{3}\right)^2, \frac{1}{2}\left(\frac{1-p_l}{3}\right)^2, \frac{3}{4}\left(\frac{1-p_l}{3}\right)^2, \left(\frac{1-p_l}{3}\right)^2\right\} \quad \text{for } p_l > 0.5,$$

respectively,

- both common probabilities of success:

$$p_T, p_C \in \{0.05, 0.1, 0.2 \ldots, 0.8, 0.9, 0.95\}.$$

We made 5000 replications for each situation.

We use box plots to display the results. In all graphs, the black square with the cross is the true value of the difference $p_T - p_C$ and the setting of parameters for given situation is written above the graphs. Simulations were conducted using MATLAB® software.

3.1. Balanced situation across the trial

In balanced situation across the trial, the number of subjects in the ith center in the treated group $n_{T,i}$ was the same as the number of subjects in the ith center in the control group $n_{C,i}$ and also was the same as the number of subjects in the jth center in the treated group $n_{T,j}$ and in the control group

$$n_{C,j} \quad \text{for all } i = 1, \ldots, I \quad \text{and} \quad j = 1, \ldots, I.$$

Figures 1 and 2 show the situations, where the probability p_T is the same and common probability p_C is being changed. In Figure 1 the number of subjects in the all groups is 5 whereas in Figure 2 is 100. We can observe that with growing value of the difference the median gets farther from the true value of the difference. In the last box plot of Figure 2, the true value of the difference even exceeds the box, but this extreme situations is rare in the practice.

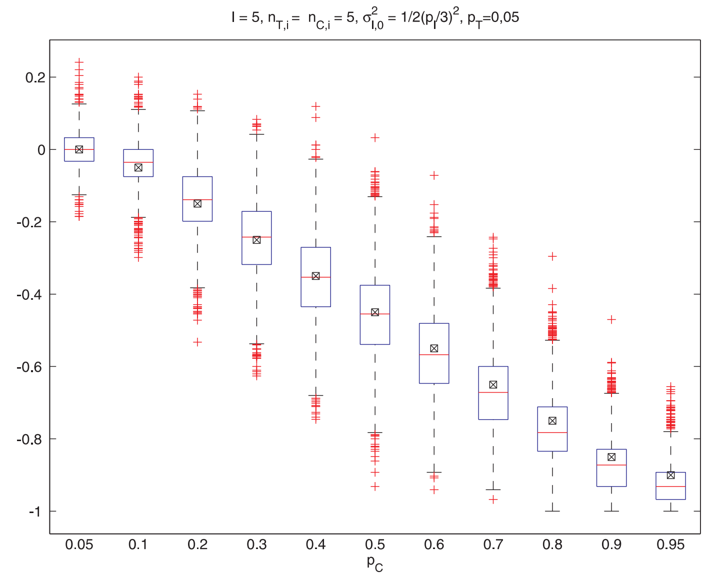

FIGURE 1. Box plots for different $p_T - p_C$, 5 subjects in groups.

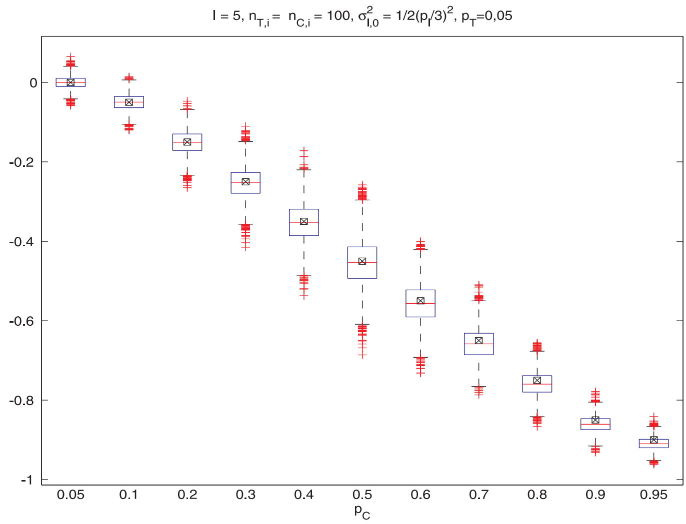

FIGURE 2. Box plots for different $p_T - p_C$, 100 subjects in groups.

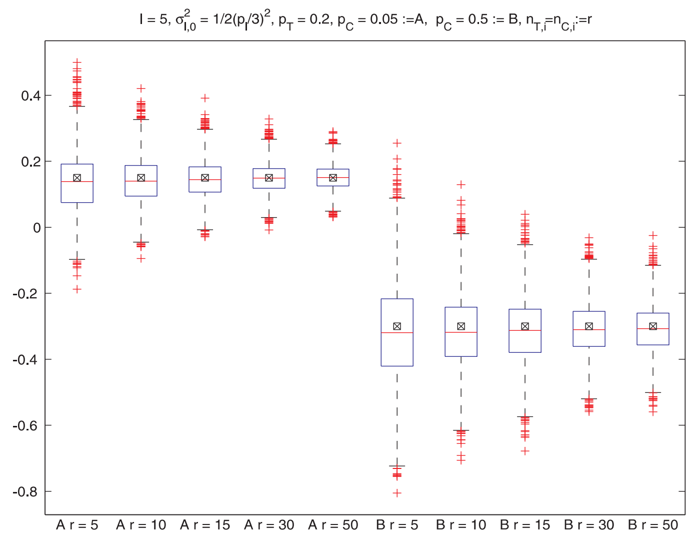

FIGURE 3. Box plots for increasing number of subjects in groups, 5 centers.

Figures 3 and 4 illustrate situations, where the number of subjects in all groups increases from 5 to 50 and is labelled by the letter r. It is shown for two differences 0.15 and minus 0.3.

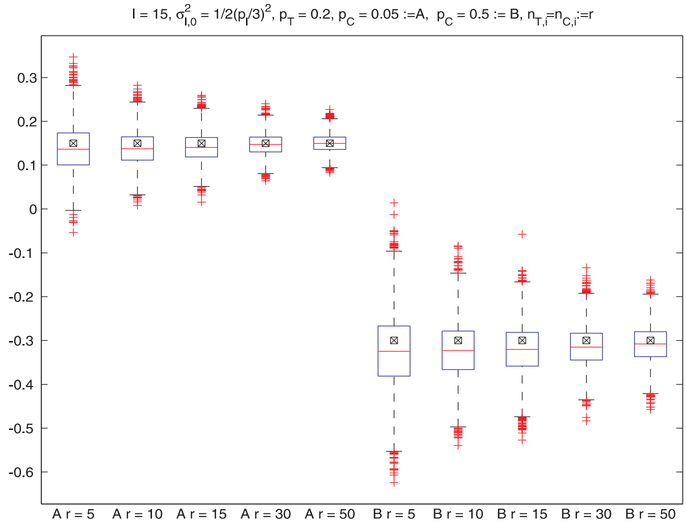

FIGURE 4. Box plots for increasing number of subjects in groups, 15 centers.

We can observe that the median gets closer to the true value of the difference with growing number of subjects and at the same time the interquartile range is getting smaller. So the simulations confirm that the estimator is biased but is asymptotically unbiased. In comparison with Figure 3, where the number of centers I is equal to 5, in Figure 4, where the number of centers I equals 15, we can see slower convergence of the median to the true value of the difference.

Now, look at the situations with the growing value of the variance of random effects $\sigma^2_{l,0}$, again, shown for three different differences 0.15, 0 and -0.3. s_1 labels $\sigma^2_{l,0} = 0$, s_2 labels that $\sigma^2_{l,0}$ equals the middle value, i.e., $\sigma^2_{l,0} = \frac{1}{2}\left(\frac{p_l}{3}\right)^2$ and finally s_3 labels situations, where $\sigma^2_{l,0}$ is equal to the greatest value from all considered values of $\sigma^2_{l,0}$. While in Figure 5, where the number of subjects in all groups equals 5, we cannot see big differences between results for each variance, if we increase the number of subjects to 100 (Figure 6), we can see larger differences.

Figures 7 and 8 show situations with increasing number of centers I. Again, in Figure 7 the number of subjects is 5 and in Figure 8 is 100. We can see that the results are getting better up to 25 centers and then they stay approximately the same. Generally, we can say that involvement of another center in the study brings considerable improvement of results only to the total number of centers approximately 15 or 25 depending on the number of subjects in groups.

Figure 9 shows comparison of two situations with the same total number of subjects in the study $n = 150$, but in the situation labelled $r1I1$ we had 5 centers and 15 subjects in all groups, while in the situation labelled $r2I2$ was

15 centers and 5 subjects in all groups. We can see that the second situation has slightly better results. It is more observable in Figure 10, where the total number of subjects in the study is 1000.

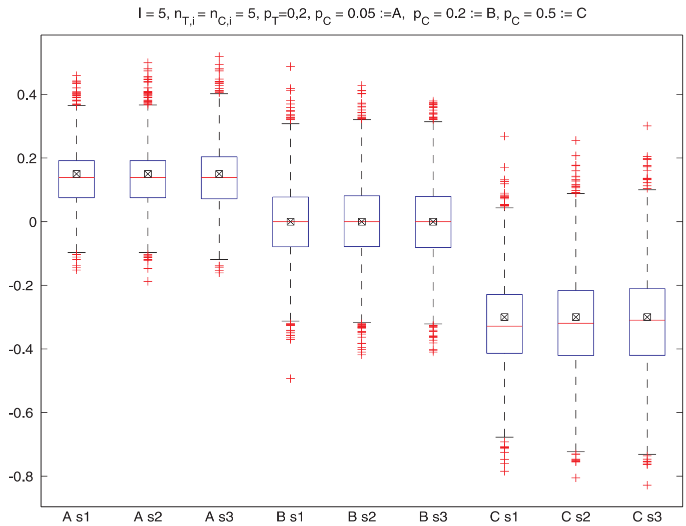

FIGURE 5. Box plots for increasing $\sigma^2_{l,0}$, 5 subjects in groups.

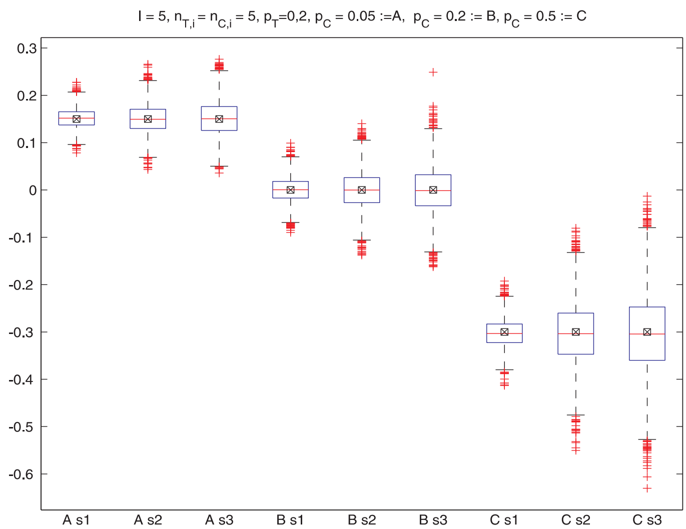

FIGURE 6. Box plots for increasing $\sigma^2_{l,0}$, 100 subjects in groups.

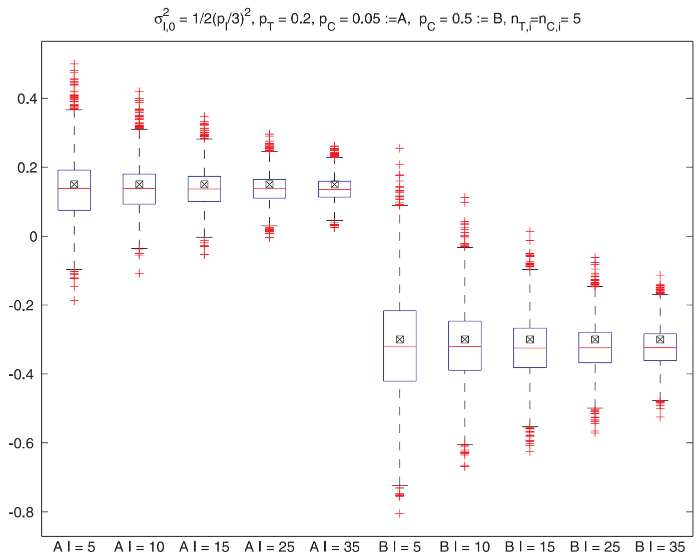

FIGURE 7. Box plots for increasing number of centers, 5 subjects in groups.

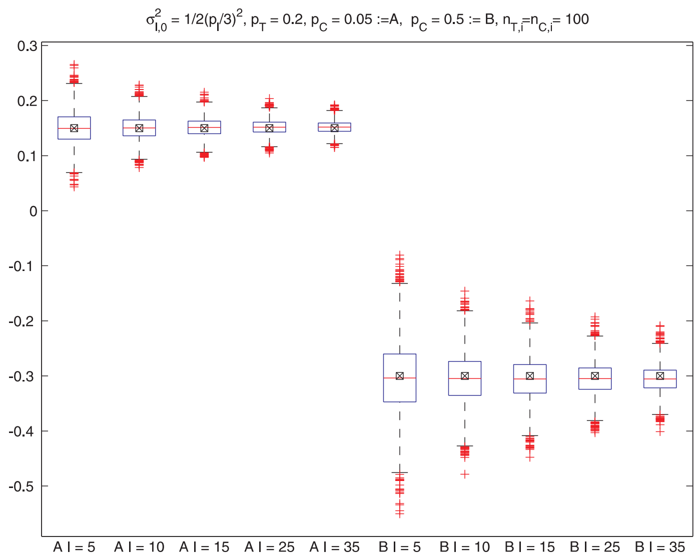

FIGURE 8. Box plots for increasing number of centers, 100 subjects in groups.

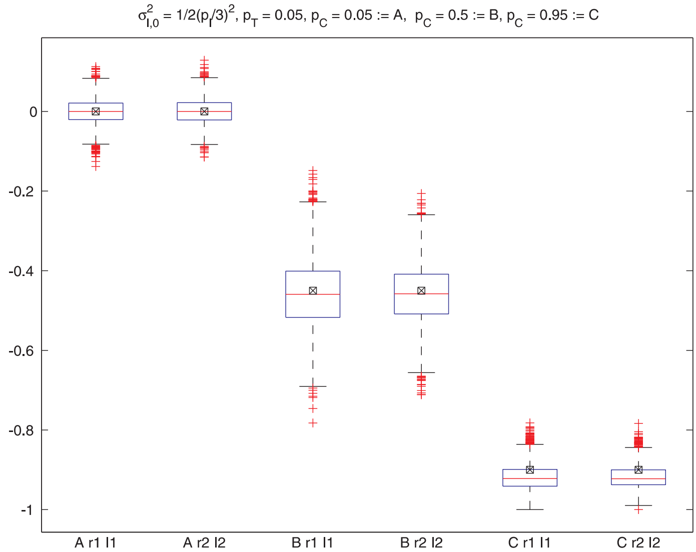

FIGURE 9. Box plots—the same sum of subjects $n = 150$, different number of centers $\mathrm{I} = 5 := \mathrm{I1}$ and $\mathrm{I} = 15 := \mathrm{I2}$ and different number of subjects in groups $n_{l,i} = 15 := \mathrm{r1}$ and $n_{l,i} = 5 := \mathrm{r2}$.

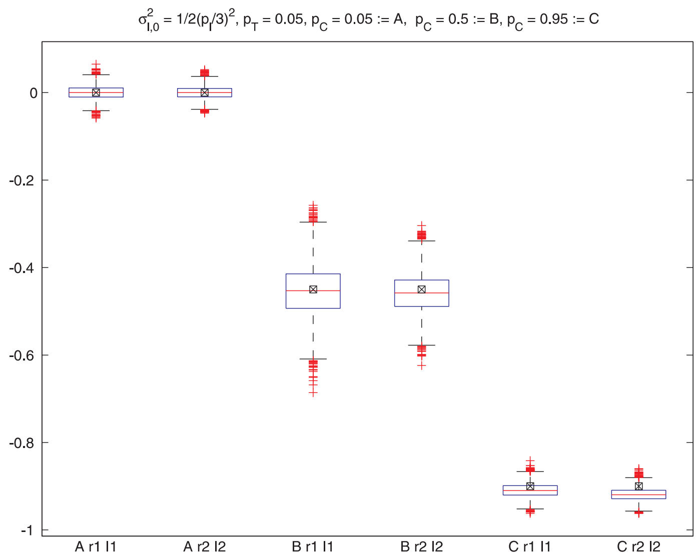

FIGURE 10. Box plots—the same sum of subjects $n = 1000$, different number of centers $\mathrm{I} = 5 := \mathrm{I1}$ and $\mathrm{I} = 10 := \mathrm{I2}$ and different number of subjects in groups $n_{l,i} = 100 := \mathrm{r1}$ and $n_{l,i} = 50 := \mathrm{r2}$.

3.2. Balanced situations across centers

In the balanced situation across centers the number of subjects in the ith center in the treated group $n_{T,i}$ was different from the number of subjects in the ith center in the control group $n_{C,i}$, but it was the same as the number of subjects in the jth center in the treated group $n_{T,j}$ for $i = 1, \ldots, I$ and $j = 1, \ldots, I$, that is

$$n_{T,i} = n_{T,j} \neq n_{C,i} = n_{C,j} \qquad \text{for} \quad i, j = 1, \ldots, I.$$

In Figure 11 and 12, the balanced situations across centers are compared with balanced situations across the trial. In Figure 11 the situation labelled $r1$ has the same number of subjects, $n_{l,i} = 10$, in all groups, while the situation labelled $r2$ has 5 subjects in all treated groups and 15 subjects in all control groups. The sum of subjects in both situations is the same and equals 100. We can see a slight deterioration for the second situation. The same behavior can be seen in Figure 12, where the total number of subjects is 300.

Now, look at the comparison of similar situations in Figure 13, where the number of subjects in all treated groups is 30 in both situations and the number of subjects in all control groups rises from 50 to 100. Even if the total number of subjects increases by about 500 subjects, the results become approximately the same. We can say that it is because of the greater difference between the number of subjects in control and treated groups for the second situation.

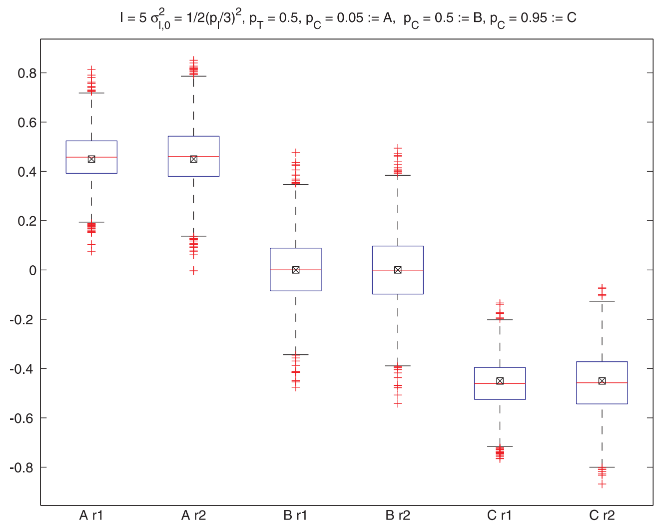

FIGURE 11. Box plots—the same sum of subjects $n = 100$, the same number of subjects in groups $n_{T,i} = n_{C,i} = 10$ (labelled by r1) versus the different number of subjects in treated groups $n_{T,i} = 5$ and control groups $n_{C,i} = 15$ (labelled by r2).

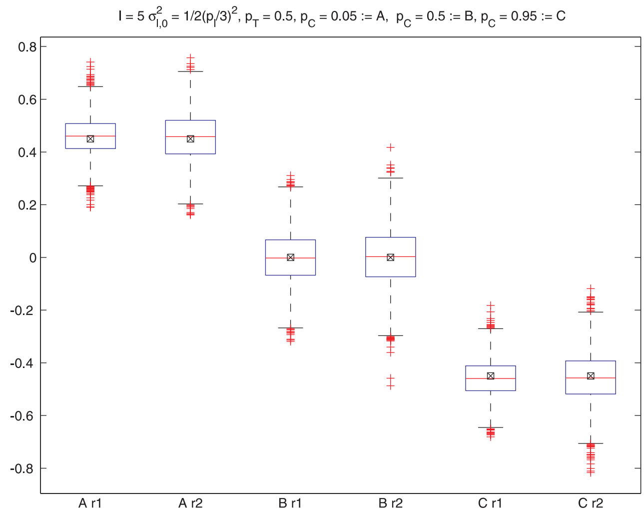

FIGURE 12. Box plots—the same sum of subjects $n = 300$, the same number of subjects in groups $n_{T,i} = n_{C,i} = 30$ (labelled by r1) versus the different number of subjects in treated groups $n_{T,i} = 10$ and control groups $n_{C,i} = 50$ (labelled by r2).

3.3. Balanced situations across groups

In the balanced situation across groups, the number of subjects in the ith center in the treated group $n_{T,i}$ was the same as the number of subjects in the ith center in the control group $n_{C,i}$, but was different from the number of subjects in the jth center in the treated group $n_{T,j}$ and in the control group $n_{C,j}$ for some $i \neq j$, $i = 1, \dots, I$ and $j = 1, \dots, I$, that is

$$n_{T,i} = n_{C,i} \quad \text{for all } i, j = 1, \dots, I \wedge n_{T,i} \neq n_{T,j} \quad \text{for at least one couple } i \neq j.$$

Figures 14 and 15 show the comparison of a situation with the different number of subjects across the centers labelled by $r2$ with the balanced situations across the trial labelled by $r1$. In Figure 14, the total sum of subjects is 150, $n_{l,i}$ is equal to 15 in the situation $r1$ and $\mathbf{n}_T = \mathbf{n}_C$ are equal to $[5,\ 30,\ 5,\ 30,\ 5]'$ in the situation $r2$. We cannot see big differences in results, even if we add subjects to the study to get a total sum of 600 subjects (see Figure 15). $n_{l,i}$ is now 30 in the situation $r1$ and $\mathbf{n}_T$, $\mathbf{n}_C$ are both equal to $[10,\ 50,\ 10,\ 50,\ 10,\ 50,\ 10,\ 50,\ 10,\ 50]'$ in the situation $r2$.

Finally, Figure 16 shows comparison of three previous situations: $r1$ denotes balanced situation across the trial ($n_{l,i} = 30$), $r2$ denotes balanced situations across centers ($n_{T,i} = 10$ and $n_{C,i} = 50$) and $r3$ denotes balanced situations across groups ($\mathbf{n}_T = \mathbf{n}_C = [10,\ 50,\ 10,\ 50,\ 10,\ 50,\ 10,\ 50,\ 10,\ 50]'$). We can see that the best results are for the situation $r1$. Slightly worse results are in the $r3$ situation and the worst-case is the situation $r2$.

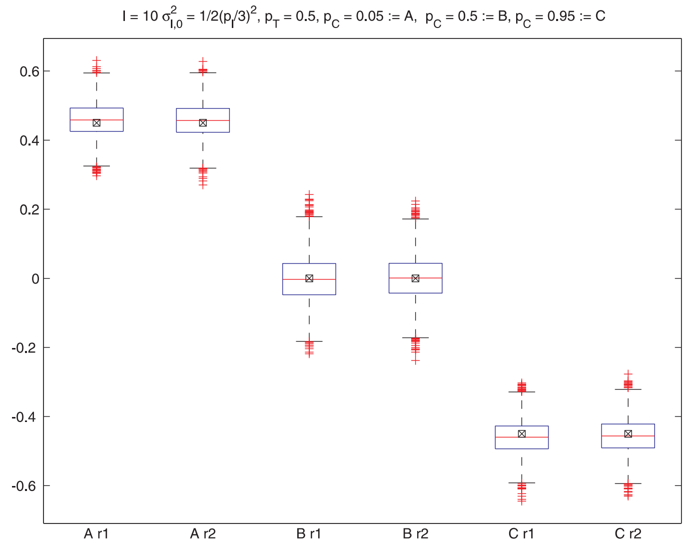

FIGURE 13. Box plots—increase only in number of subjects in control groups, i.e., $n_{T,i} = 30$ in both situations and $n_{C,i} = 50$ (labelled by r1) versus $n_{C,i} = 100$ (labelled by r2).

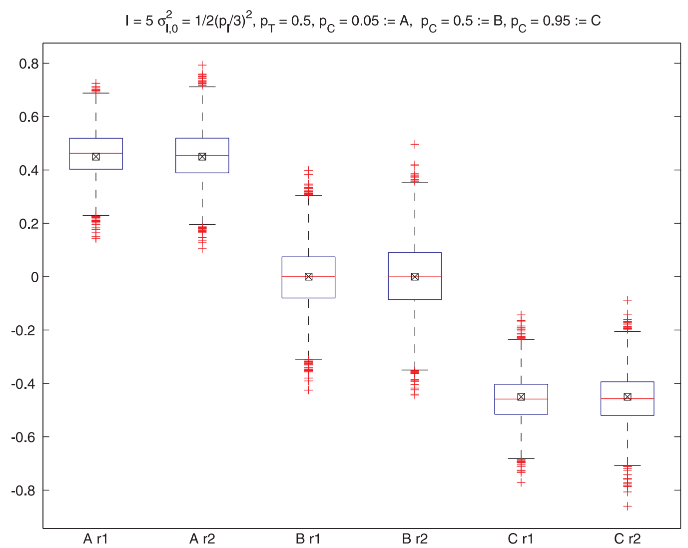

FIGURE 14. Box plots—the same sum of subjects $n = 150$, the same number of subjects across centers versus the different number of subjects across centers.

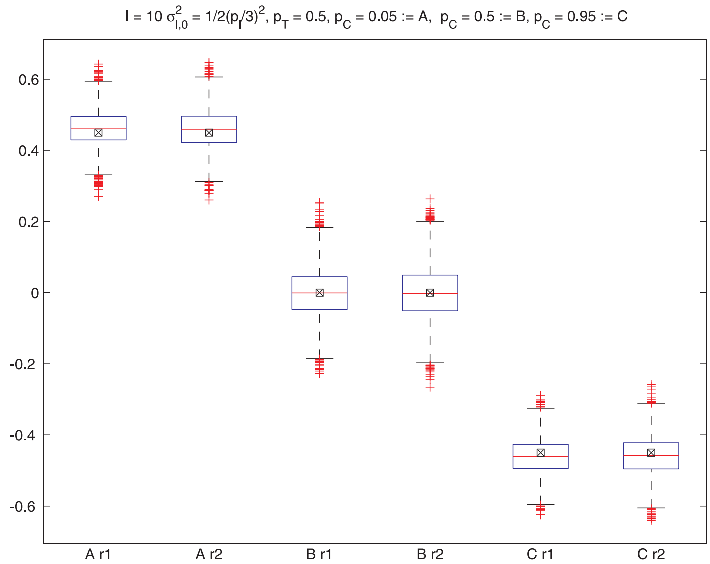

FIGURE 15. Box plots—the same sum of subjects $n = 600$, the same number of subjects across centers versus the different number of subjects across centers.

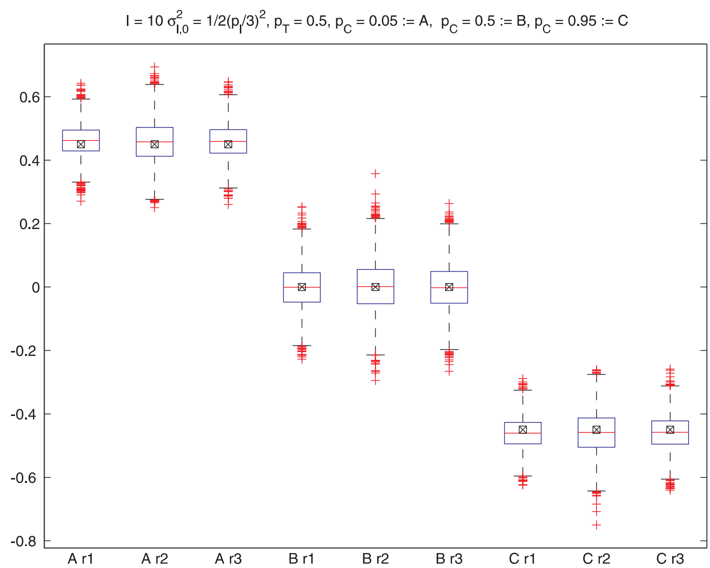

FIGURE 16. Box plots—the same sum of subjects $n = 600$.

4. Concluding remarks

As a conclusion, we can say that the point estimator (4) shows very good basic statistical properties. For example, it has relatively small interquartile range even for small numbers of subjects in a study or for a small number of subjects in groups and a large number of centers. For study design, we can recommend that the study should have similar numbers of subjects in the treated and control groups and, of course, it is better if study is also balanced across groups. This is especially important when the number of patients in groups is small. Simulations also confirmed that the estimator (5) is biased but asymptotically unbiased.

REFERENCES

[1] AGRESTI, A.—CAFFO, B.: *Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures,* Amer. Statist. **54** (2000), 280–288.

[2] DOKOUPILOVÁ, P.: *A confidence interval for the probability difference of overall treatment effect—simulation study,* Acta Univ. M. Belii Ser. Math. **18** (2011), 7–16.

[3] PAULE, R. C.—MANDEL, J.: *Consensus values and weighting factors,* J. Res. Natl. Bur. Stand. **87** (1982), 377–385.

[4] WIMMER, G.—WITKOVSKÝ, V.: *Konfidenčné intervaly pre efekt ošetrenia v klinických pokusoch,* in: Sborník prací 13. letní školy JČMF ROBUST '04, Třešť, 2004, JČMF, Praha, 2004, pp. 427–434.

Received August 11, 2011

Department of Mathematics and Statistics
Faculty of Science
Masaryk University
Kotlářská 2
CZ–611-37 Brno
CZECH REPUBLIC
E-mail: pavla@krajickova.cz