Leibniz Universität Hannover

Reports of the Institute of Biostatistics

No 14 / 2007

Leibniz University of Hannover

Natural Sciences Faculty

Titel: *Confidence intervals and limits for the risk ratio revisited*

Authors: *Frank Schaarschmidt*

1 Introduction

Consider two independent binomial random variables X_i, $i = 1, 2$, with $X_i \sim Bin(\pi_i, n_i)$. Interest is to estimate $\rho = \pi_2/\pi_1$. Note, that the risk ratio is not invariant with respect to the definition of success and failure, as is the odds ratio. Therefore, properties of confidence intervals are not symmetric in $\pi_i = 0.5$.

2 Confidence intervals for ρ

The point estimate for ρ is $\hat{\rho} = p_2/p_1$, with $p_i = x_i/n_i$. Gart and Nam (1988) discuss a simple large sample interval presented in equation (1), called "Crude" in the following.

$$exp\left(log\left(\tilde{\rho}\right) \pm c\sqrt{\hat{u}}\right) \tag{1}$$

where

$$\tilde{\rho} = \frac{\left(x_2 + 0.5\right) / \left(n_2 + 0.5\right)}{\left(x_1 + 0.5\right) / \left(n_1 + 0.5\right)},$$

and

$$\hat{u} = \hat{V}\left(log\left(\tilde{\rho}\right)\right) = \frac{1}{x_2 + 0.5} + \frac{1}{x_1 + 0.5} - \frac{1}{n_2 + 0.5} - \frac{1}{n_1 + 0.5},$$

and $c = z_{1-\alpha/2}$ is the quantile of the standard normal distribution.

This method yields degenerate intervals $[1, 1]$ in case $x_2 = n_2$ and $x_1 = n_1$. It is one of the few intervals considered in Gart and Nam (1988) which can be computed in the case of the event $x_2 = 0$ and $x_1 = 0$. It has the advantage that calculated bounds for ρ are the same as the reciprocal of the bounds calculated for $1/\rho$ Gart and Nam (1988), i.e. they are invariant with respect to exchanging numerator and denominator. Dann and Koch (2005) call this method "Modified Taylor Series".

An adaptation of the ideas of Agresti and Coull (1998) is presented by Dann and Koch (2005) as Adapted Agresti method. Here, a version appropriate for more general settings that those presented by Dann and Koch is given. This method replaces $\tilde{\rho}$, and $\hat{u}$ in equation (1) by

$$\tilde{\rho} = \frac{\left(x_2 + 1\right) / \left(n_2 + 2\right)}{\left(x_1 + 1\right) / \left(n_1 + 2\right)},$$

and

$$\hat{u} = \hat{V}\left(log\left(\tilde{\rho}\right)\right) = \frac{1}{x_2 + 1} + \frac{1}{x_1 + 1} - \frac{1}{n_2 + 2} - \frac{1}{n_1 + 2}.$$

This method will be called "Add-2" in the following. Dann and Koch (2005) assign pseudo-observations in a ratio depending on parameter ρ hypothesized under a null-hypothesis of non-inferiority, which is not a valid method in general situations.

The third method considered here is the "Score" method discussed in Gart and Nam (1988), section 3.3 (Methods based on Likelihood Methods). Confidence bounds for ρ are found by iterative process, involving the solution of quadratic equations. In the limited simulation study presented by Gart and

Nam (1988), this method shows best coverage probabilities among the considered methods. For computational details, refer to Gart and Nam (1988). Gart and Nam (1988) state, that the method is not computable for the case $x_1 = x_2 = 0$. However, in my implementation, problems in the iterative process occured also for a number of other events, like $x_1 = 0$, $x_2 = 0$, $x_1 = n_1$, $x_2 = n_2$. In the simulation study, I replace $x_1 = 0.5$, $x_1 = 0.5$, $x_1 = n_1 - 0.5$, $x_1 = n_2 - 0.5$, in the iterative process if these events occured, respectively. Therefore, the method referred to as "Score" in this report is not exactly the same as that described by Gart and Nam (1988).

3 Validation

Applying my algorithm for the Score method on the examples given in Gart and Nam (1988), Section 5, in order to compute two-sided nominal 95% confidence intervals gives the limits in Table 3.

x_2	n_2	x_1	n_1	estimate	lower limit	upper limit
8	15	4	15	2	0.815019	5.336303
6	10	6	20	2	0.8435384	4.5940787

Table 1: Calculation of the results of Example 1 and 2 in Gart and Nam (1988)

The calls to invoke these computations are:

```
RRScore(x=c(8,7), y=c(4,11))

RRScore(x=c(6,4), y=c(6,14))
```

The functions are defined in the file

```
"G:/Dell - Schaarschmidt/SIM_EU/RiskRatio/GartNamScore2.txt"
 ".../SIMRR_Rcode.txt"
```

4 Previous simulation studies

Gart and Nam (1988) show, that the Crude interval in equation (1) performs liberal or conservative, depending on the setting. According to their simulations, the score method should be chosen.
Dann and Koch (2005) review the Add-2-method among others in a comparative simulation study. Their study does not consider unbalanced designs, restrict to consideration of the size of a non-inferiority test using the confidence intervals, restrict themselves to sample size $n = 100, 140, 200$, and do not present results for a sufficient number of settings $\{\pi_2, \pi_1\}$. Moreover, they consider methods with conceptual problems, such as including parameters assumed under settings of hypothesis testing into construction of confidence intervals. For these reasons, their simulation study is not sufficient for a general recommendation of the methods.
They show, that the crude method performs liberal, while the Add-2 method keeps the nominal confidence level in most of the considered settings.

5 Simulation study

Estimated (n=10000) coverage probabilities over the parameter space of $\pi_1 = (0.01, ..., 0.99)$, $\pi_2 = (0.01, ..., 0.99)$ is displayed in contour plots in the following. Red marks areas, where coverage probability is lower than 0.94, green marks areas, where coverage probability is between 0.94 and 0.96, blue marks areas where coverage probability is larger than 0.96.

5.1 Summary

For two-sided 95% confidence intervals, the Score method clearly outperforms the Crude and Add-2 method. Especially for small sample sizes, the Crude and Add-2 method can be severly liberal, especially for situations, where π_1 and π_2 are very different. For settings, where π_1 and π_2 are close to each other (the diagonal of the plots), all methods show coverage probability close to or above the nominal level. For small sample sizes, the Score interval has coverage closest to nominal level for these situation.

The one-sided application of all discussed interval methods can not be recommended for small sample sizes. All methods show severe assymmetry of coverage probability between lower and upper confidence limits. A two-sided nominal 95% confidence interval might exclude the true value in 5% of the cases by the lower bound and in nearly no case by the upper bound, or vice versa, depending on the particular setting of π_1, π_2.

5.2 Coverage probability of two-sided 95% intervals

5.2.1 Crude

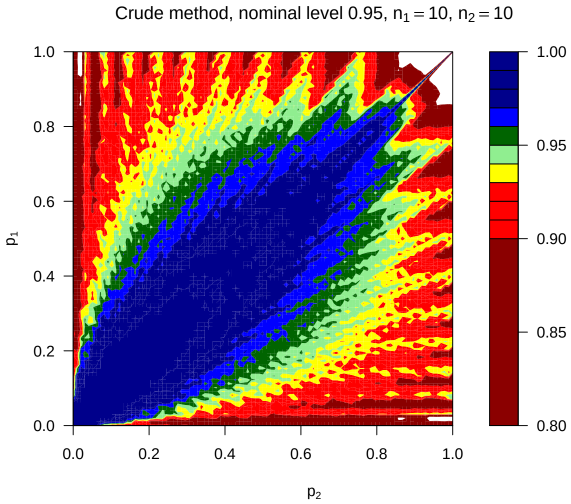

Figure 1: Coverage probability of two-sided, nominal 0.95 Crude confidence intervals for $n_1, n_2 = 10$

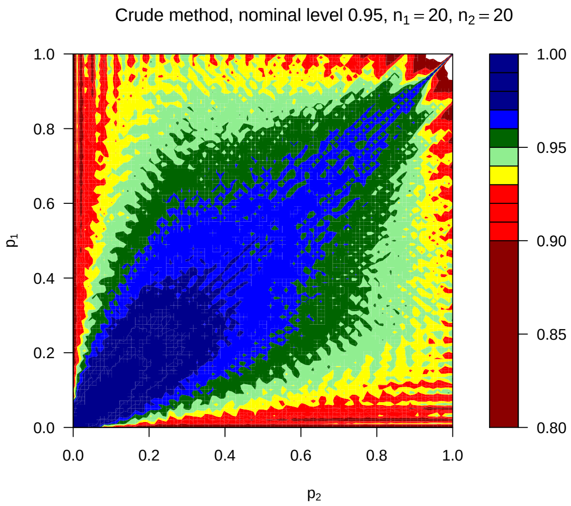

Figure 2: Coverage probability of two-sided, nominal 0.95 Crude confidence intervals for $n_1, n_2 = 20$

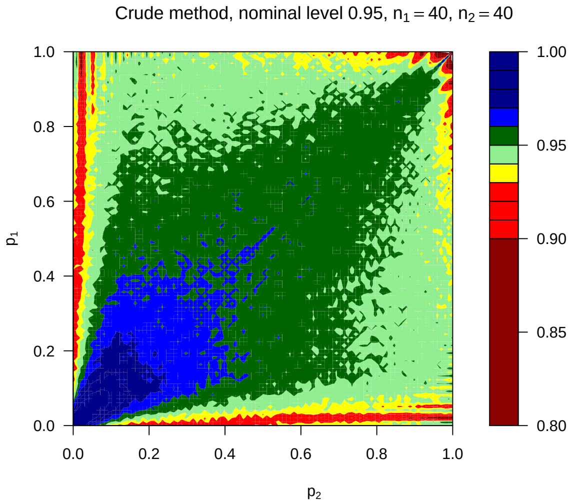

Figure 3: Coverage probability of two-sided, nominal 0.95 Crude confidence intervals for $n_1, n_2 = 40$

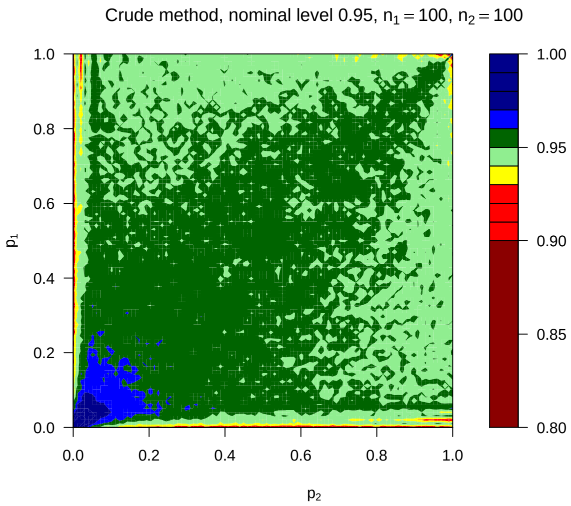

Figure 4: Coverage probability of two-sided, nominal 0.95 Crude confidence intervals for $n_1, n_2 = 100$

5.2.2 Add-2

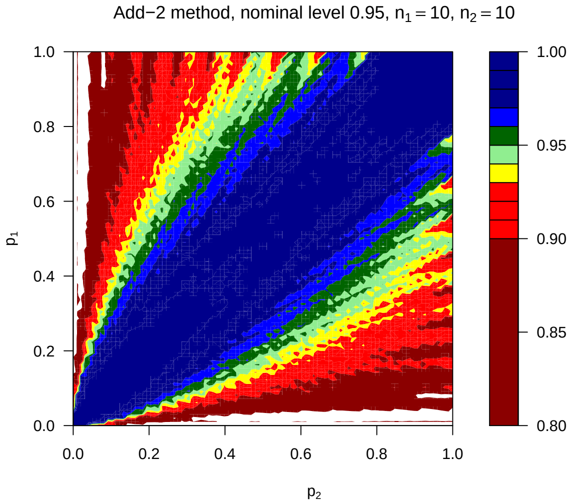

Figure 5: Coverage probability of two-sided, nominal 0.95 Add-2 confidence intervals for $n_1, n_2 = 10$

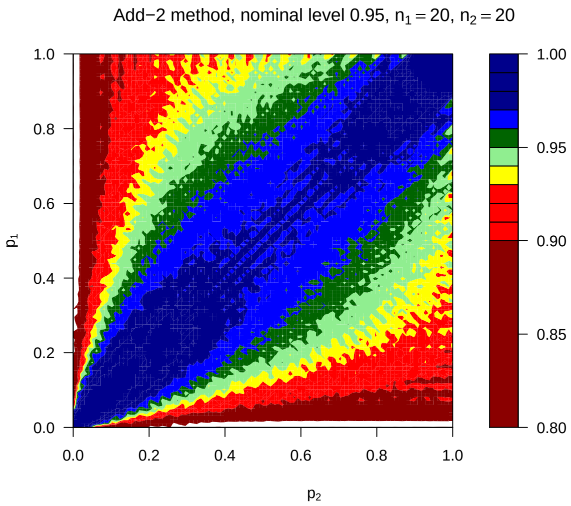

Figure 6: Coverage probability of two-sided, nominal 0.95 Add-2 confidence intervals for $n_1, n_2 = 20$

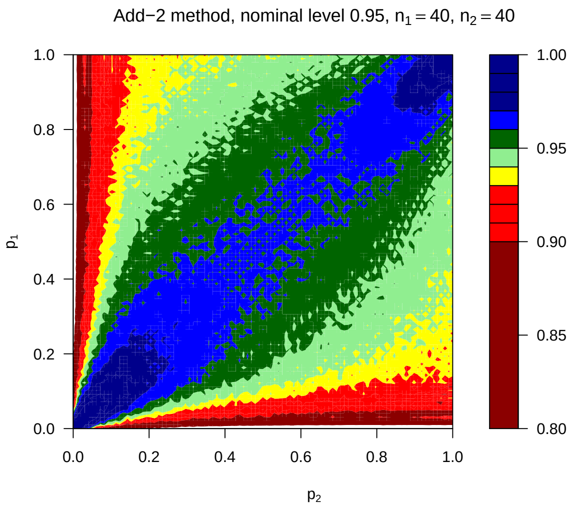

Figure 7: Coverage probability of two-sided, nominal 0.95 Add-2 confidence intervals for $n_1, n_2 = 40$

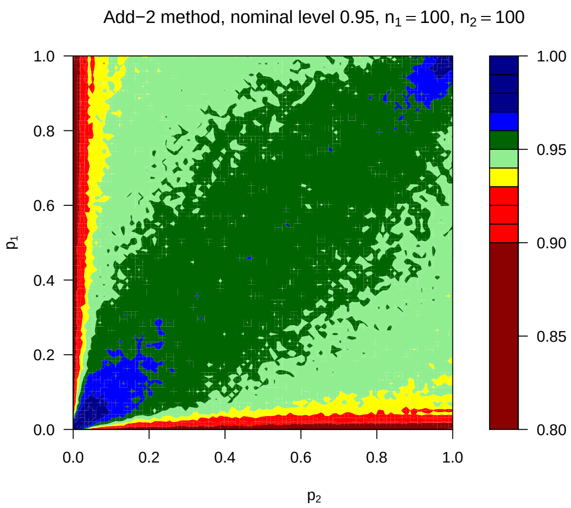

Figure 8: Coverage probability of two-sided, nominal 0.95 Add-2 confidence intervals for $n_1, n_2 = 100$

5.2.3 Score

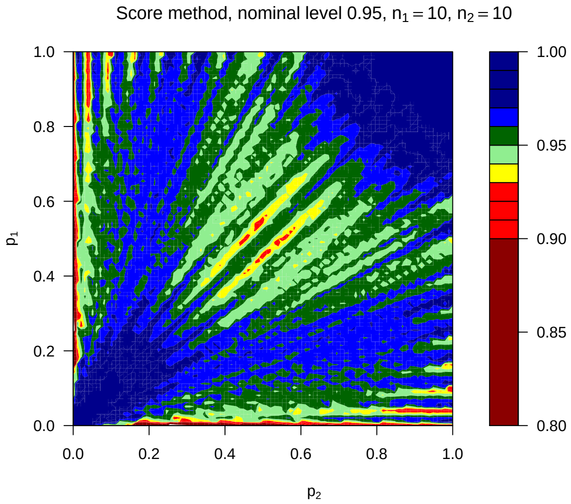

Figure 9: Coverage probability of two-sided, nominal 0.95 Score confidence intervals for $n_1, n_2 = 10$

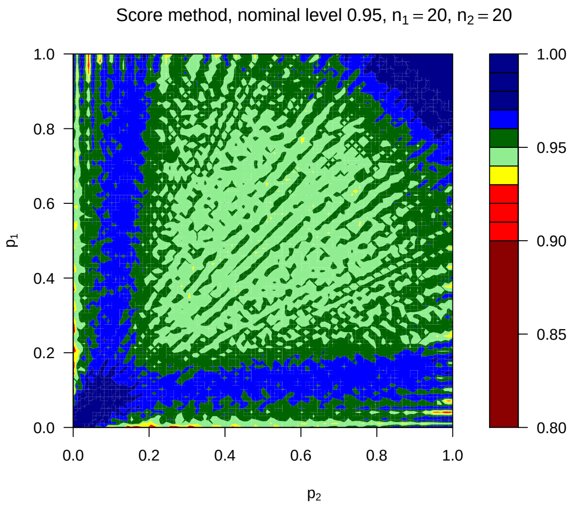

Figure 10: Coverage probability of two-sided, nominal 0.95 Score confidence intervals for $n_1, n_2 = 20$

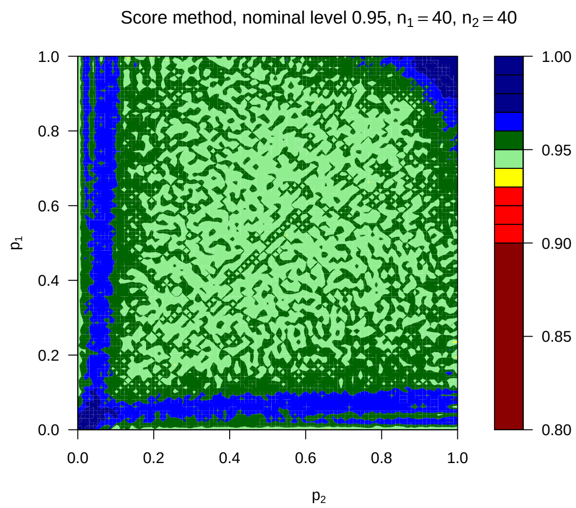

Figure 11: Coverage probability of two-sided, nominal 0.95 Score confidence intervals for $n_1, n_2 = 40$

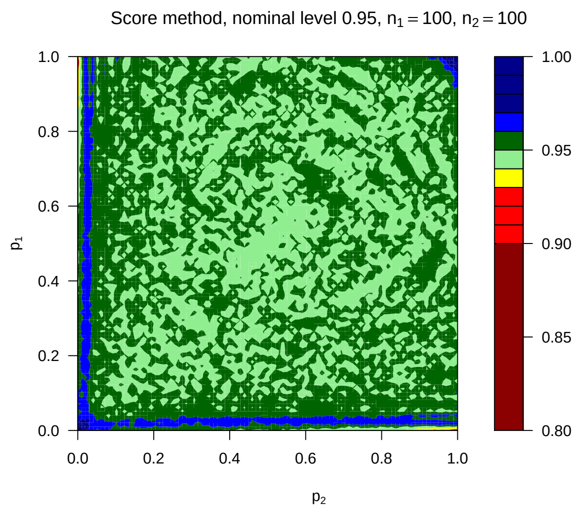

Figure 12: Coverage probability of two-sided, nominal 0.95 Score confidence intervals for $n_1, n_2 = 100$

5.3 Lower 97.5% confidence limits

Here, lower 97.5% lower bounds are considered. In this section the results for the Add-2 method are not shown because of its bad two-sided coverage. Coverage probabilities below 0.97 are marked in red, coverage probabilities between 0.97 and 0.98 are marked in green, coverage probabilities higher than 0.98 are marked in blue.

5.3.1 Crude

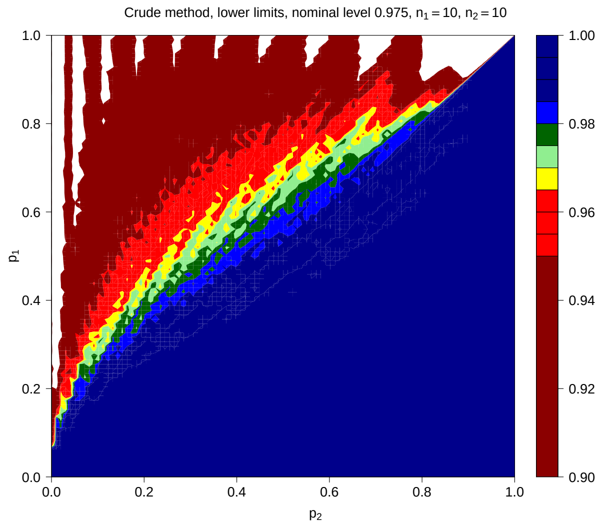

Figure 13: Coverage probability of lower, nominal 0.975 Crude confidence limits for $n_1, n_2 = 10$

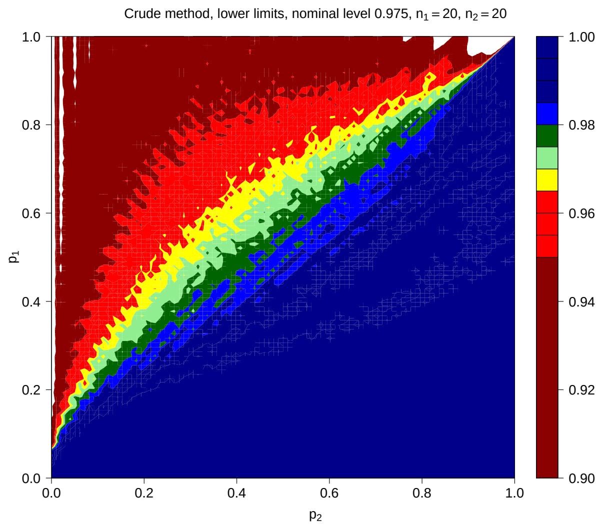

Figure 14: Coverage probability of lower, nominal 0.975 Crude confidence limits for $n_1, n_2 = 20$

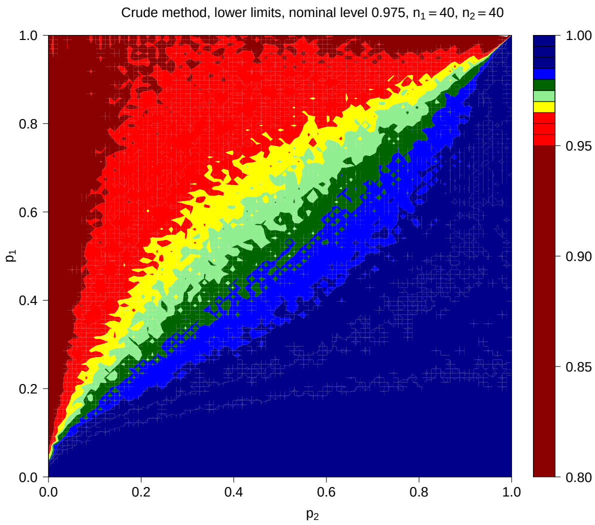

Figure 15: Coverage probability of lower, nominal 0.975 Crude confidence limits for $n_1, n_2 = 40$

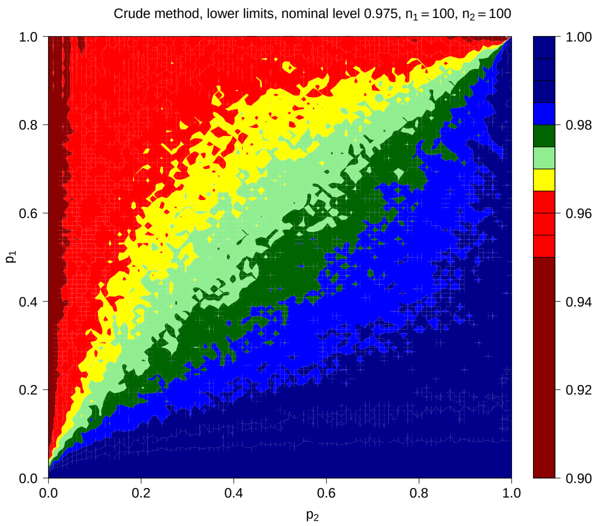

Figure 16: Coverage probability of lower, nominal 0.975 Crude confidence limits for $n_1, n_2 = 100$

5.3.2 Score

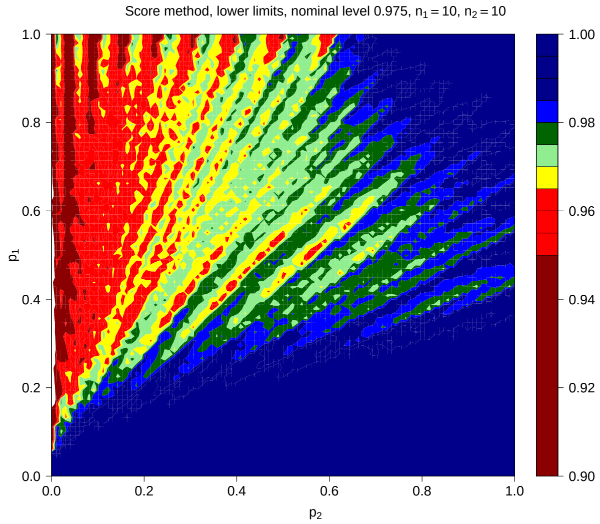

Figure 17: Coverage probability of lower, nominal 0.975 Score confidence limits for intervals for $n_1, n_2 = 10$

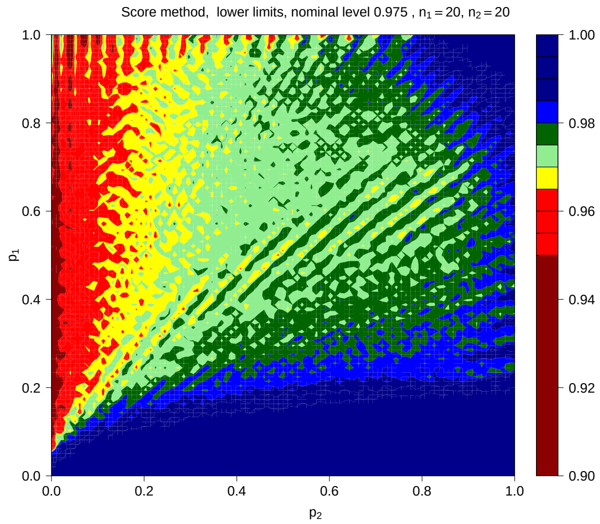

Figure 18: Coverage probability of lower, nominal 0.975 Score confidence limits for $n_1, n_2 = 20$

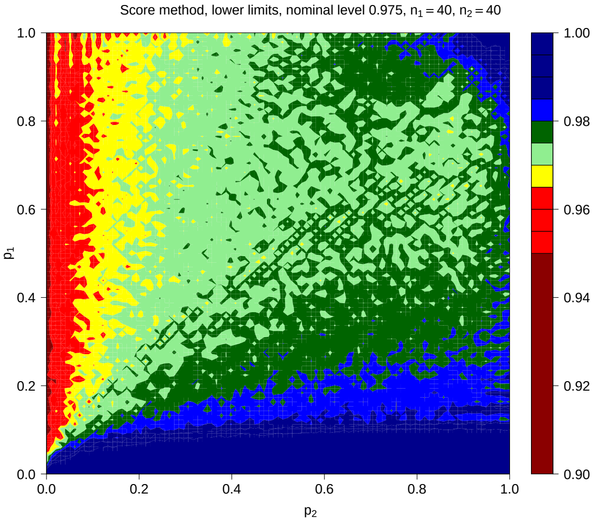

Figure 19: Coverage probability of lower, nominal 0.975 Score confidence limits for intervals for $n_1, n_2 = 40$

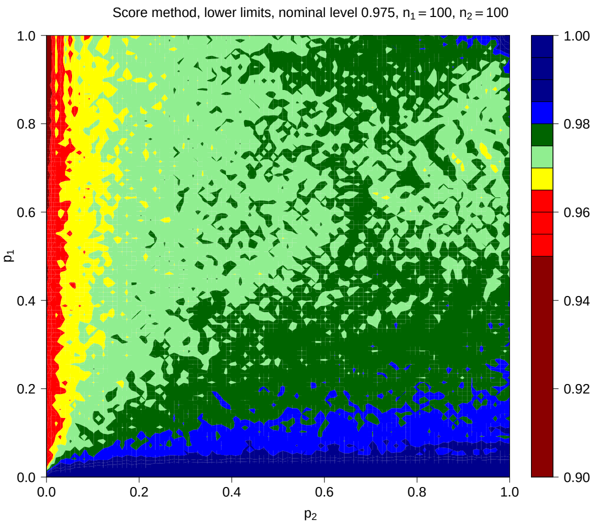

Figure 20: Coverage probability of lower, nominal 0.975 Score confidence limits for intervals for $n_1, n_2 = 100$

References

Agresti, A. and Coull, A. (1998): Approximate is better than ”exact” for interval estimation of binomial proportions. American Statistician 52: 119-126.

Gart, J.J. and Nam, J. (1988): Approximate Interval Estimation of the Ratio of Binomial Parameters: A Review and Corrections for Skewness. Biometrics 44, 323-338.

Dann, R.S., and Koch, G.G. (2005): Review and Evaluation of Methods for Computing Confidence Intervals for the Ratio of Two Proportions and Considerations for Non-inferirority Clinical Trials. Journal of Biopharmaceutical Statistics, 15:85-107.