Biometrics 000, 000–000
September 2020

DOI: 10.1111/biom.13184

Robustness of ANCOVA in randomized trials with unequal randomization

JONATHAN W. BARTLETT

Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK
email: j.w.bartlett@bath.ac.uk

Summary: Randomized trials with continuous outcomes are often analyzed using ANCOVA, with adjustment for prognostic baseline covariates. The ANCOVA estimator of the treatment effect is consistent under arbitrary model misspecification. In an article recently published in the journal, Wang *et al* proved the model based variance estimator for the treatment effect is also consistent under outcome model misspecification, assuming the probability of randomization to each treatment is 1/2. In this reader reaction we derive explicit expressions which show that when randomization is unequal, the model based variance estimator can be biased upwards or downwards. In contrast, robust sandwich variance estimators can provide asymptotically valid inferences under arbitrary misspecification, even when randomization probabilities are not equal.

Key words: ANCOVA, baseline adjustment, randomized trials

1. Introduction

In randomized trials with continuous outcomes the baseline covariate adjusted treatment effect estimator (ANCOVA) is consistent under arbitrary misspecification of the assumed linear regression model (Yang and Tsiatis, 2001). Recently in the journal Wang et al. (2019) proved that the model based variance estimator from an ANCOVA analysis of a randomized trial is valid under arbitrary misspecification, and therefore advocated its use for analysis of trials with continuous outcomes. Concurrently, the US Food and Drug Administration (2019) have recently issued draft guidance on the topic of baseline covariate adjustment in randomized trials with continuous outcomes. This draft guidance also advocates use of ANCOVA, and states that the type I error rate is controlled even when the model is misspecified.

Wang et al. (2019) assumed that the probability of randomization to each arm is 1/2, and stated that if this is not the case, a robust variance estimator should be used. While a randomization probability of 1/2 is most common, many trials are conducted with unequal randomization probabilities. In particular often the probability of randomization to the experimental arm is greater than 1/2 in light of a hoped for improved outcome on the experimental treatment compared to control. In this reader reaction we provide explicit expressions for the variance of the ANCOVA treatment effect estimator and the probability limit of the model based variance estimator, thereby shedding more light on the impact of model misspecification on the validity of model based inferences in this setting.

2. Model based ANCOVA variance estimation with unequal randomization

Following the notation of Wang et al. (2019), we assume we observe n i.i.d. copies of $(\mathbf{W}, A, Y)$, where $\mathbf{W}$ is a $k \times 1$ column vector of bounded baseline covariates, A is the binary treatment group indicator ($A = 1$ for experimental treatment, $A = 0$ for control) and Y is the continuous outcome. Like Wang et al. (2019), we assume $A \perp\!\!\!\perp W$, but we let $P(A = 1) = \pi$, where π may differ from 1/2.

The target of inference is the average treatment effect $\Delta = E(Y|A = 1) - E(Y|A = 0)$. The unadjusted estimator of Δ is the difference in treatment group sample means: $\hat{\Delta}^{unadj} = \sum_{i=1}^{n} Y_i A_i / \sum_{i=1}^{n} A_i - \sum_{i=1}^{n} Y_i(1 - A_i) / \sum_{i=1}^{n}(1 - A_i)$. The ANCOVA estimator adjusts for the baseline covariates $\mathbf{W}$ by fitting the following linear regression model:

$$E(Y|A, \mathbf{W}) = \beta_0 + \beta_A A + \beta_{\mathbf{W}}^T \mathbf{W}, \tag{1}$$

where the regression coefficients are estimated by the ordinary least square estimators $\hat{\beta}_0$, $\hat{\beta}_A$, and $\hat{\beta}_{\mathbf{W}}$. The ANCOVA estimator $\hat{\Delta}^{ancova}$ of Δ is $\hat{\Delta}^{ancova} = \hat{\beta}_A$. We let $\underline{\beta}_0$, $\underline{\beta}_A$ and $\underline{\beta}_W$ denote the probability limits of these estimators.

As noted by Wang et al. (2019), Yang and Tsiatis (2001) and Tsiatis et al. (2008) proved, under the stated assumptions, that $\hat{\Delta}^{ancova}$ is a consistent estimator of Δ under arbitrary misspecification of the linear model in equation (1), so that $\underline{\beta}_A = \Delta$. Following Wang et al. (2019), we let $Var^*(\hat{\Delta}^{ancova})$ denote the asymptotic variance of $\hat{\Delta}^{ancova}$, in the sense that $n^{1/2}(\hat{\Delta}^{ancova} - \Delta)$ converges in distribution to a mean zero normal with variance $Var^*(\hat{\Delta}^{ancova})$.

Inferences from ANCOVA are by default in statistical software packages based on the so called model based variance

estimator for $\hat{\Delta}^{ancova}$, which is given by

$$\widehat{Var}_M(\hat{\Delta}^{ancova}) = \frac{\widehat{Var}(Y - \hat{\beta}_0 - \hat{\beta}_A A - \hat{\beta}_{\mathbf{W}}^T \mathbf{W})}{(n-1)\left\{\widehat{Var}(A) - \widehat{Cov}(\mathbf{W}, A)^T \widehat{Var}(\mathbf{W})^{-1} \widehat{Cov}(\mathbf{W}, A)\right\}},$$

where following Wang et al. (2019) the estimated variances and covariances on the right hand side are sample variance and sample covariances, with degrees of freedom taken into account (see the Supporting Information of Wang et al. (2019) for precise definitions). Wang et al. (2019) prove that when $\pi = 1/2$, $n\widehat{Var}_M(\hat{\Delta}^{ancova})$ converges in probability to the true asymptotic variance $Var^*(\hat{\Delta}^{ancova})$. As a consequence, under these assumptions, asymptotically Wald-type hypothesis tests have the correct type I error under the null $\Delta = 0$ and the corresponding confidence intervals attain their nominal coverage levels.

The following theorem, proved in Web Appendix A, gives the asymptotic variance of $\hat{\Delta}^{ancova}$ for arbitrary $0 < \pi < 1$, generalising the results of Wang et al. (2019).

THEOREM 1: *Given the previously stated assumptions with $0 < \pi < 1$, the true asymptotic variance $Var^*(\hat{\Delta}^{ancova})$ of the ANCOVA estimator $\hat{\Delta}^{ancova}$ is given by*

$$Var^*(\hat{\Delta}^{ancova}) = \frac{Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 1)}{\pi} + \frac{Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 0)}{1 - \pi}.$$

The next theorem, proved in Web Appendix B, gives the probability limit of $n\widehat{Var}_M(\hat{\Delta}^{ancova})$ under arbitrary $0 < \pi < 1$.

THEOREM 2: *For the model based variance estimator $\widehat{Var}_M(\hat{\Delta}^{ancova})$ we have*

$$n\widehat{Var}_M(\hat{\Delta}^{ancova}) \xrightarrow{P} Var_M^*(\hat{\Delta}^{ancova}) = \frac{Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 1)}{1 - \pi} + \frac{Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 0)}{\pi}.$$

Together the two theorems imply that the model based variance estimator of $\hat{\Delta}^{ancova}$ is only asymptotically valid (and hence hypothesis tests and confidence intervals have correct asymptotic size and coverage) if $\pi = 1/2$, as assumed by Wang et al. (2019), or if $Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 1) = Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 0)$. The latter conditional variances are not in general equal under misspecification of the outcome model, such that when $\pi \neq 1/2$, the model based ANCOVA variance estimator is generally biased. A special case where they are equal is if $E(Y|A, \mathbf{W})$ is a linear function of A and some function of $\mathbf{W}$ and $Var(Y|A, \mathbf{W})$ does not depend on A. Otherwise bias is expected. For example, even if the conditional mean function $E(Y|A, \mathbf{W})$ is correctly specified, if $Var(Y|A, \mathbf{W})$ depends on A, the model based ANCOVA variance estimator is biased. Alternatively, even if $Var(Y|A, \mathbf{W})$ does not depend on A, if $E(Y|A, \mathbf{W})$ involves interaction terms between A and $\mathbf{W}$ again bias is expected.

We note that a special case of our result occurs when W is empty, such that $\hat{\Delta}^{ancova} = \hat{\Delta}^{unadj}$. In this case our result corresponds to the well known fact that the two sample t-test does not control the type I error rate in general if the outcome variable has different variance in the two groups, which leads to Welch's adaptation of the t-test allowing for unequal variances.

Our results imply that when $\pi \neq 1/2$, the model based ANCOVA variance estimator could be biased downwards or upwards, depending on the configuration, leading to a type I error rate either below or above the nominal level. Suppose for example that $\pi > 1/2$, such that a greater proportion of patients are randomized to the experimental treatment. Then if $Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 1) > Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 0)$ the model based ANCOVA variance is too large, leading to type I error rates lower than the nominal level and confidence intervals which over-cover. If $Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 1) < Var(Y - \underline{\beta}_{\mathbf{W}}^T \mathbf{W} | A = 0)$ the model based ANCOVA variance is too small, leading to inflated type I error rates and confidence intervals which under-cover. When $\pi \neq 1/2$, asymptotically valid inferences can be obtained by using a heteroskedastic robust sandwich variance estimator (Long and Ervin, 2000). Web Appendix C of the Supporting Information contains a simulation study demonstrating these results empirically.

3. Discussion

We have shown that the model based ANCOVA variance estimator of the average treatment effect is under general misspecification of the outcome model inconsistent when $\pi \neq 1/2$. Since model misspecification is always a concern, at least for moderate to large trials with unequal randomization we recommend using a robust sandwich variance estimator for inference, rather than the default model based variance estimator. Sandwich variance estimators are incorporated into all mainstream statistical packages, such that this approach is easily implementable in practice.

Like Wang et al. (2019), we have assumed that randomization is simple. In practice other randomization schemes, such as stratified randomization, are sometimes used. As noted by Wang et al. (2019), under such randomization schemes, obtaining asymptotically valid variances when covariates not used in the randomization are adjusted for in the outcome model, under general misspecification of the outcome model, remains an open problem.

Acknowledgments

The author thanks David Wright & Daniel Jackson for useful discussions on the topic and thanks Chris Jennison for his review & comments on the manuscript.

ORCID

Jonathan Bartlett ORCID ID: 0000-0001-7117-0195

References

Long, J. S. and Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician* **54,** 217–224.

Tsiatis, A. A., Davidian, M., Zhang, M., and Lu, X. (2008). Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. *Statistics in Medicine* **27,** 4658–4677.

US Food and Drug Administration (2019). Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biologics with Continuous Outcomes.

Wang, B., Ogburn, E. L., and Rosenblum, M. (2019). Analysis of covariance (ANCOVA) in randomized trials: More precision and valid confidence intervals, without model assumptions. *Biometrics*, https://doi.org/10.1111/biom.13062.

Yang, L. and Tsiatis, A. A. (2001). Efficiency study of estimators for a treatment effect in a pretest–posttest trial. *The American Statistician* **55,** 314–321.

Supporting information

Web Appendices A, B and C referenced in Section 2 are available with this paper at the Biometrics website on Wiley Online Library. R code for the simulation study described and reported in Web Appendix C is available at `https://github.com/jwb133/ancovaRCT`.