January 15 2003

A SIMPLE APPROACH TO POWER AND SAMPLE SIZE CALCULATIONS IN LOGISTIC REGRESSION AND COX REGRESSION MODELS

MICHAEL VÆTH[1] AND EVA SKOVLUND[2]

[1]Department of Biostatistics, University of Aarhus, Vennelyst Boulevard 6,
DK-8000 Århus C, Denmark
[2]Section of Medical Statistics, University of Oslo, P.O.Box 1122 Blindern,
N-0317 Oslo, Norway

Short title:
POWER AND SAMPLE SIZE CALCULATIONS FOR REGRESSION MODELS

Keywords:
Sample size, power, logistic regression, Cox regression

Correspondance to:

Michael Væth
Department of Biostatistics
University of Aarhus
Vennelyst Boulevard 6
DK-8000 Århus C
Denmark

Phone: +45 8942 6129
Fax: +45 8942 6140
e-mail: vaeth@biostat.au.dk

A SIMPLE APPROACH TO POWER AND SAMPLE SIZE CALCULATIONS IN LOGISTIC REGRESSION AND COX REGRESSION MODELS

SUMMARY

For a given regression problem it is possible to identify a suitably defined equivalent two-sample problem such that the power or sample size obtained for the two-sample problem also apply to the regression problem. For a standard linear regression model the equivalent two-sample problem is easily identified, but for generalised linear models and for Cox regression models the situation is more complicated. An approximately equivalent two-sample problem may, however, also be identified here. In particular, we show that for logistic regression and Cox regression models the equivalent two-sample problem is obtained by selecting two equally sized samples for which the parameters differ by a value equal to the slope times twice the standard deviation of the independent variable and further requiring that the overall expected number of events is unchanged. In a simulation study we examine the validity of this approach to power calculations in logistic regression and Cox regression models. Several different covariate distributions are considered for selected values of the overall response probability and a range of alternatives. For the Cox regression model we consider both constant and non-constant hazard rates. The results show that in general the approach is remarkably accurate even in relatively small samples. Some discrepancies are, however, found in small samples with few events and a highly skewed covariate distribution. Comparison with results based on alternative methods for logistic regression models with a single continuous covariate indicates that the proposed method is at least as good as its competitors. The method is easy to implement and therefore provides a simple way to extend the range of problems that can be covered by the usual formulas for power and sample size determination.

INTRODUCTION

Standard software for power and sample size calculations usually covers one- and two-sample problems, but less frequently offers calculations for regression problems and then typically only for standard linear regression. For a given regression problem it is, however, possible to identify a suitably defined equivalent two-sample problem for which simple formulas for power and sample size are available or for which standard software can provide the answer. The sample size or power thus obtained will then apply also to the regression problem. For a standard linear regression model this equivalence is particularly simple and the calculation will be identical to the one obtained if the correct formula was implemented. For generalised linear models, including logistic regression, and for proportional hazards regression further approximations are needed in order to obtain a simple solution. To assess the validity of these approximations for small to moderate sample sizes a simulation study must be undertaken. In this article we outline the approach and report the main results from a comprehensive simulation study of its performance when applied to logistic regression and proportional hazards regression. In the final section we compare our results with those of previously published approaches and briefly discuss extensions to regression models with several covariates.

THE BASIC IDEA

The basic idea is most easily explained in a normal linear regression setting since no approximations are involved here. Consider a continuous response y and a covariate x satisfying a standard linear regression $y = a + bx + e$ for which the error, e, has a normal distribution with mean 0 and standard deviation σ. Usually we want to determine a sample size such that a test of the hypothesis of no association, i.e. that the slope b is 0, has a specified power for an alternative value of the slope representing an important association not to be overlooked. To accomplish this we must specify a level of significance α, the power $1-\beta$ of the test for an alternative b_A, a value σ for the standard deviation of the random variation about the regression line, and the expected variation of the covariate *x* in the sample measured by the sum of squared deviations SSD_X. In the design situation the extra variability introduced by estimating the variance is usually ignored and the test statistic then becomes

$$\frac{\hat{b}}{s.e.(\hat{b})} = \hat{b}\sqrt{\frac{SSD_X}{\sigma^2}},$$

where $\hat{b}$ is the usual least squares estimate of the slope. On the hypothesis of no association the test statistic is a standard normal deviate. If the slope is b_A the test statistic follows a normal distribution with variance 1 and mean

$$b_A\sqrt{\frac{SSD_X}{\sigma^2}} = \sqrt{N}\,\frac{s_X}{\sigma}\,b_A, \tag{1}$$

where s_X is the empirical standard deviation of the covariate *x*. To derive the appropriate sample size the null distribution of the test statistic is used to identify an α-level rejection region, and the sample size is then found as the smallest *N* for which the non-null probability of a value in the rejection region is larger than $1-\beta$. Alternatively the sample size may be fixed and the power of the test for a specified alternative can be obtained.

Now consider a two-sample problem for which the response in the two groups, group 1 and group 2 say, is assumed to follow normal distributions with a common standard deviation ω and means μ and $\mu+\delta$, respectively. Samples of size *m* and *n* are to be drawn from the two groups. Let $N = m+n$ denote the total sample size and let $\pi_1 = m/N$ and $\pi_2 = 1-\pi_1$ denote the sample fractions. To determine the appropriate size of a study for comparison of the two groups, i.e. to test the hypothesis $\delta = 0$, we must specify a level of significance α, the power $1-\beta$ of the test for a value δ_A representing the expected difference, a value ω for the standard deviation, and the fraction π_1 of the total sample coming from group 1. The sample size calculation is based on the test statistic

$$\frac{\bar{y}_2 - \bar{y}_1}{\omega\sqrt{\frac{1}{n}+\frac{1}{m}}},$$

where $\bar{y}_1$ and $\bar{y}_2$ are the sample averages. On the hypothesis of no difference the test statistic is a standard normal deviate. If the difference is δ_A the statistic follows a normal distribution with standard deviation 1 and mean

$$\frac{\delta_A}{\omega\sqrt{\frac{1}{m}+\frac{1}{n}}} = \sqrt{N}\,\frac{\sqrt{\pi_1(1-\pi_1)}}{\omega}\delta_A, \qquad (2)$$

To derive the appropriate sample size the null distribution of the test statistic is used to identify an α-level rejection region, and the sample size is then found as the smallest *N* for which the non-null probability of a value in the rejection region is larger than $1-\beta$. Almost all computer programs for sample size determinations will do this calculation. Comparing formula (1) and (2) the following approach to sample size calculations for the regression problem suggests itself. Select π_1, ω and δ_A such that

$$\frac{s_X}{\sigma} b_A = \frac{\sqrt{\pi_1(1-\pi_1)}}{\omega}\delta_A.$$

The sample size derived for the two-sample problem will then also apply to the regression problem, since the non-null distributions then become identical. A simple choice would be to let $\pi_1 = 0.5$, $\omega = \sigma$ and $\delta_A = 2 s_X b_A$, i.e. to derive the sample size for the regression problem consider a two-sample problem with equal sample sizes, a standard deviation ω equal to the root mean square error σ, the value of the difference on the alternative hypothesis equal to $\delta_A = 2 s_X b_A$, and then compute the total sample size for this two-sample problem.

LOGISTIC REGRESSION

The sample size result for the linear regression problem is in principle exact since the standard deviation is assumed known. Moreover, the solution obtained by identifying a two-sample problem with the same non-centrality parameter is simple and easy to apply in practice. For other regression models the properties of statistical tests are usually derived from asymptotic theory. Unfortunately, the asymptotic results will often be complicated, since the variance is typically a function of the parameters specifying the mean. Further approximations are therefore needed in order to obtain a sufficiently simple approach to sample size calculations. We here consider a logistic regression model, but similar results may easily be derived for other generalised linear models.

In a logistic regression model the association between a binary response *y* and a continuous covariate *x* is modelled by the relation

$$\ln\left(p/(1-p)\right) = a + bx\,,$$

where $p = p(x) = P(y = 1; x)$ is the probability of a positive response. To test the hypothesis of no association one of several asymptotically equivalent large-sample tests may be used. For the developments here it is convenient to consider Wald's test statistic, which relies on the asymptotic normality of the maximum likelihood estimate. The test statistic is simply

$$\frac{\hat{b}}{s.e.(\hat{b})}$$

A straightforward, but tedious, calculation shows that

$$s.e.(\hat{b}) = \sqrt{\frac{1}{\sum p_i(1-p_i)\left\{\sum w_i x_i^2 - \left(\sum w_i x_i\right)^2\right\}}}\,,$$

where p_i is the estimated probability of a positive response for observation *i* and the weights w_i are given by

$$w_i = \frac{p_i(1-p_i)}{\sum p_j(1-p_j)}.$$

Note that on the hypothesis of no association these weights are all equal and the term in brackets becomes SSD_X. If the slope is b_A Wald's test statistic is asymptotically normal with variance 1 and mean

$$\frac{b_A}{s.e.(\hat{b})} = \sqrt{N}\sqrt{\sum w_i x_i^2 - \left(\sum w_i x_i\right)^2}\sqrt{\sum p_i(1-p_i)/N}\ b_A. \qquad (3)$$

For values of b_A close to 0 the mean is thus approximately equal to

$$\sqrt{N}\ s_X\sqrt{\tilde{p}(1-\tilde{p})}\ b_A, \qquad (4)$$

where s_X is the empirical standard deviation of the covariate *x* and $\tilde{p}$ is a typical response probability, e.g. the average value or the value for *x* equal to the average value of the covariate. In the simulation study described below the latter value was used.

The results for the two-sample problem can be obtained from those above by letting the covariate *x* take one value for *m* of the observation and another value for the remaining $n = N - m$ observations. Let p_1 and p_2 denote the response probabilities for the two groups and $\delta_A = \ln\left\{p_2(1-p_1)/\left(p_1(1-p_2)\right)\right\}$, then the asymptotic mean of Wald's test statistic becomes

$$\sqrt{N}\sqrt{\left[\frac{1}{\pi_1 p_1(1-p_1)} + \frac{1}{\pi_2 p_2(1-p_2)}\right]^{-1}}\ \delta_A. \qquad (5)$$

For values of δ_A close to 0 the mean is approximately equal to

$$\sqrt{N}\sqrt{\pi_1(1-\pi_1)}\sqrt{\overline{p}(1-\overline{p})}\ \delta_A, \qquad (6)$$

where $\bar{p}$ is the average of the response probabilities.

A comparison of formula (4) and (6) suggests that a two-sample problem approximately equivalent to the regression problem is defined by $\bar{p} = \tilde{p}$, $\pi_1 = 0.5$ and $\delta_A = 2 s_X b_A$. The sample size computed for the equivalent two sample problem will then also apply for the regression problem.

COX REGRESSION

For the two sample problem with censored survival data Schoenfeld[1,2] obtained a simple formula for sample size determination under a proportional hazards model. His result coincides with the one based on an exponential survival time distribution. To implement the basic approach developed above for Cox regression models we therefore consider an exponential regression model in which the hazard rate λ depends on the covariate *x* as

$$\ln(\lambda) = \ln(\lambda(x)) = a + bx,$$

To test the hypothesis of no association we again consider Wald's test statistic. Note that this is not the test statistic used when the data is analyzed by Cox regression, but in view of Schoenfeld's result we would expect a sample size or power calculation derived from the exponential model to apply also for the Cox regression model. The distribution of Wald's test is asymptotically normal with variance 1 and mean equal to

$$\frac{b_A}{s.e.(\hat{b})} = \sqrt{N}\,\frac{\sqrt{\sum w_i x_i^2 - \left(\sum w_i x_i\right)^2}}{\sqrt{\left(\sum P_i / N\right)^{-1}}}\; b_A, \qquad (7)$$

Here P_i is the response probability, i.e. the probability of an event, for the *i*'th individual and the weights w_i are given by

$$w_i = \frac{P_i}{\sum P_j}.$$

These probabilities depend in general on the underlying survival distribution and on length of accrual and follow-up. For values of b_A close to 0 the mean is approximately equal to

$$\sqrt{N}\ s_X \sqrt{\tilde{P}}\ b_A, \tag{8}$$

where $\tilde{P}$ denotes the average response probability.

The two-sample problem can again be viewed as a special case of the regression problem. Let λ_1 and λ_2 denote the hazard rates for the two groups and introduce the log-hazard ratio $\delta_A = \ln\{\lambda_2/\lambda_1\}$, then the asymptotic mean of Wald's test statistic becomes

$$\sqrt{N}\sqrt{\left[\frac{1}{\pi_1 P_1}+\frac{1}{\pi_2 P_2}\right]^{-1}}\ \delta_A, \tag{9}$$

where π_0 and π_1, as before, denote the sample fractions and P_1 and P_2 are the response probabilities in the two groups. For values of δ_A close to 0 the asymptotic mean is approximately equal to

$$\sqrt{N}\sqrt{\pi_1(1-\pi_1)}\sqrt{\overline{P}}\ \delta_A, \tag{10}$$

where $\overline{P}$ is the average response probability.

An equivalent two-sample problem is therefore obtained by choosing $\overline{P} = \tilde{P}$, $\pi_1 = 0.5$ and $\delta_A = 2s_X b_A$. A sample size computed for this two-sample problem would be approximately correct for the exponential regression problem and consequently also for the Cox regression problem.

SIMULATIONS

The theoretical results suggest that sample size and power calculations for logistic regression and Cox regression models with a single continuous covariate can be obtained directly from the analogous calculations for a two-sample problem with equally sized samples for which the parameters differ by the slope times twice the standard deviation of the independent variable and with the further requirement that the overall expected number of events is the same as in the regression sample.

In a simulation study we examined the validity of this approach to power calculations using Stata[3] both for data generation and analysis. For each regression model a number of scenarios was considered. In each scenario we specified a set of values of the covariate, an overall response probability, and for the simulations of survival data also an underlying survival time distribution and a censoring scheme. In a given scenario data sets were generated for 9 to 14 different values of the slope b_A. We considered samples of size 100, 200 and 500 and the proportion of rejections in 10,000 simulated replications was used to estimate the power of Wald's test of the hypothesis of no association. A rejection occurred if

$$\frac{|\hat{b}|}{s.e.(\hat{b})} > 1.96,$$

where *b* is the estimate obtained from the logistic regression or the Cox regression analysis.

Without loss of generality covariate values were generated as a sample from a distribution standardized to have mean zero and variance 1. The following covariate distributions were used: normal, uniform, double exponential, and gamma with shape parameter 3. We also simulated the two-sample problem to check that the standard formulas were appropriate. Since the covariate had mean zero the overall response probability could be determined from the constant term *a* in the regression model. For logistic regression we used the values -1.5, -1, -0.5, 0, 0.5, 1, 1.5 for *a* giving, respectively, a response probability of 0.18, 0.27, 0.38, 0.50, 0.62, 0.73, and 0.82 for $x = 0$.

The survival data were generated as censored survival times from an exponential distribution or from Weibull distributions with shape parameter 0.2, 0.5, 2, or 5. For the exponential distribution the values -1, 0, and 1 were used as the constant term *a* in the regression. For the Weibull simulations these values were modified such that the overall response probability was approximately unchanged. To mimic a clinical trial censoring schemes were selected to reflect a uniform patient arrival in an accrual period of length A with an additional follow-up interval of length $1-A$, see e.g. Schoenfeld & Ritcher[1] for use of similar censoring schemes. Three values of A, 0.2, 0.5 and 0.8, were used. The overall proportion of responses, i.e. one minus the proportion of censored values, is a complex function of the parameters defining the scenario, but is in most cases essentially equal to the response probability at the average value of the covariate. In the survival data scenarios the response probability for $x = 0$ varied from 0.20 to 0.91.

The simulated power was compared with the power of the equivalent two-sample test. For the logistic regression model this is just a test of no association in a 2×2 table and we used the standard formula for the power of this test, see e.g. Machin & Campbell[4], with response probabilities p_1 and p_2 selected such that $\ln\left(p_1/(1-p_1)\right) = a - s_X b_A$ and $\ln\left(p_2/(1-p_2)\right) = a + s_X b_A$. For a we used the value chosen for the simulation. In real-life applications this option is obviously not available, but we wanted the comparison to focus on the validity of the approximation and not on the effect of misspecification of the parameters. For Cox regression the simulated power was compared to the formula for the power derived by Schoenfeld[2] for the two-sample problem with the logarithm of the hazard ratio given as $2s_X b_A$ and the expected proportion of events equal to the probability of response for the average value of the covariate computed for the parameters specified in the simulation scenario. Again, this option would not be available in practice; indeed the choice of an appropriate value for the expected proportion of events is a major problem in applications and is discussed extensively by several authors[2,5,6,7].

Since the covariates were generated as pseudo-random numbers from a distribution the simulations also allowed us to assess the impact of sampling variation in the covariates. The simulated power could be compared both to the two-sample formula based on the population values of the mean and variance of the covariate (0 and 1, respectively) and the formula based on the mean and variance in the sample, which may differ slightly from their population values. A comparison with the latter curve focuses on the quality of the approximation while deviations from the former curve also include sampling variation in the covariate.

RESULTS

The full set of simulation results are found in an appendix, which can be downloaded from http://www.biostat.au.dk/~vaeth. Here we show results from some representative scenarios and describe the general aspects of the remaining results.

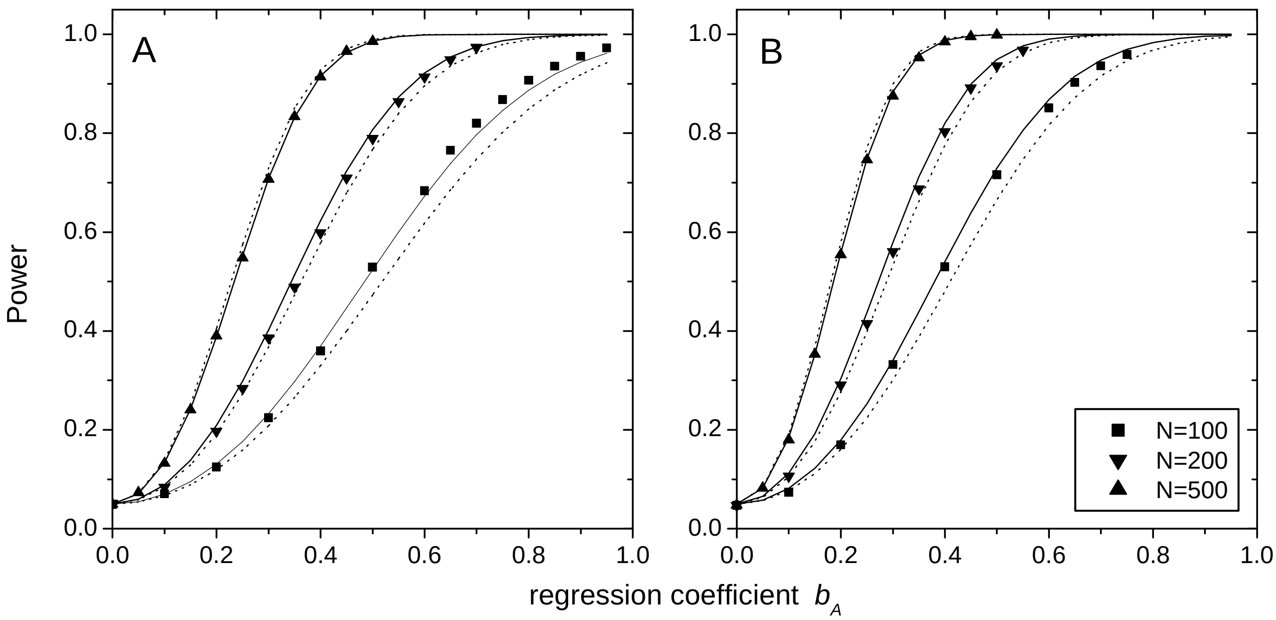

Figure 1. Simulated and estimated power of Wald's test in a logistic regression model with a covariate sampled from a standard normal distribution. Sample sizes are 100, 200 and 500. In panel A $p(0)$ is 0.18 and in panel B $p(0)$ is 0.38. Solid curves show estimated power using the sample mean and standard deviation of the covariate. Dotted curves show estimated power using the corresponding population values.

Figure 1 includes the results for a normal covariate with response probabilities of 18% (left panel) and 38% (right panel). Negative values of the slope b_A are not considered, since the symmetry of the standard normal distribution ensures that the power depends only on the absolute value of the slope. Moreover, when the covariate distribution is symmetric results from a scenario with response probability p also apply to a scenario with response probability $1-p$.

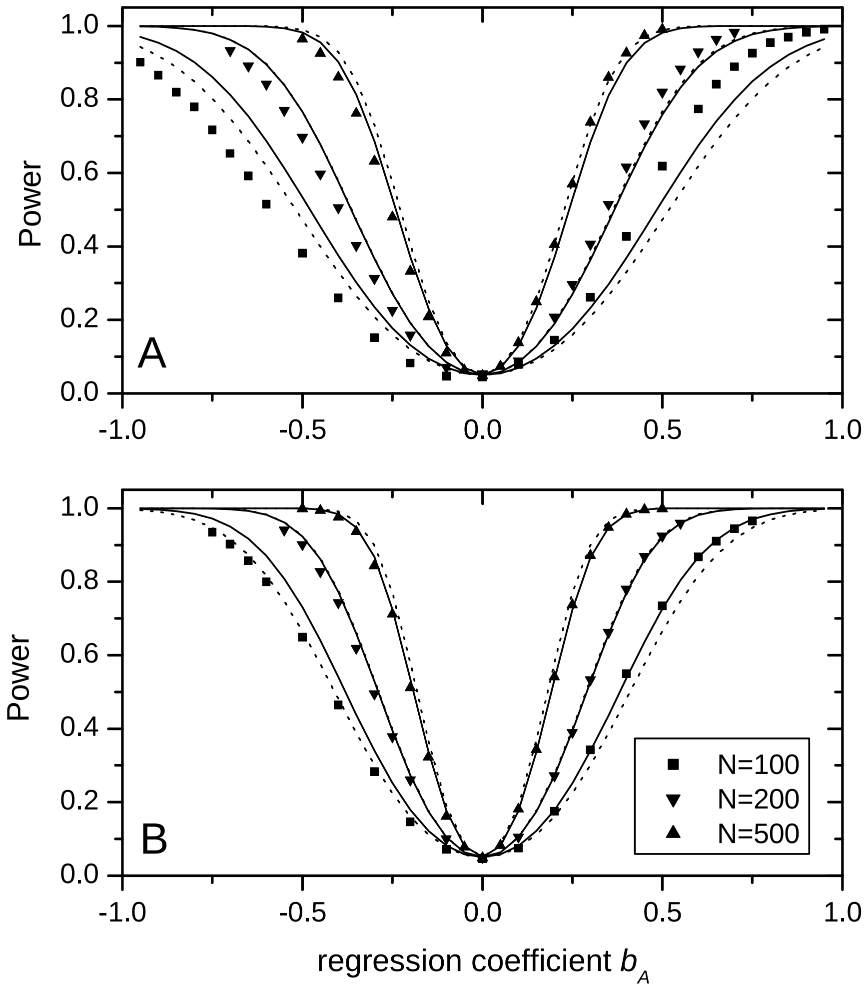

Figure 2. Simulated and estimated power of Wald's test in a logistic regression model with a covariate sampled from a standardized gamma distribution with shape parameter 3. Sample sizes are 100, 200 and 500. In panel A $p(0)$ is 0.18 and in panel B $p(0)$ is 0.38. Solid curves show estimated power using the sample mean and standard deviation of the covariate. Dotted curves show estimated power using the corresponding population values.

The results in Figure 1 shows in general an excellent agreement between the simulations and the calculations based on the empirical moments, but a slight underestimation of the power is found for the lowest response probability when the sample size is only 100. Results from simulations with uniform or double exponential covariates were almost identical to those with a normal covariate suggesting that the approach works fine for symmetrical covariate distributions unless the sample size is small and the response probability is close to 0 or 1.

Figure 2 shows the results from the same scenarios but with a covariate taken as a sample from a standardized gamma distribution with three degrees of freedom. By changing the sign of the slope results from a scenario with response probability p apply to a scenario with response probability $1-p$. Especially in the upper panel the agreement is less satisfactory unless the sample size is large. The calculations overestimate the power for negative values of the slope and, to a somewhat lesser degree, underestimate the power for positive values of the slope. In the lower panel the deviations are much less pronounced and the agreement seems satisfactory for sample size equal to 200 and 500. The gamma distribution was chosen to represent a distribution with considerable skewness to the right. For distributions with a long left tail the deviations are reversed, i.e. underestimation of the power for negative values and overestimation for positive values.

Figures 3 and 4 show selected scenarios from simulations with Cox regression models. The overall impression is very similar to the one for the logistic regression: For a symmetric covariate distribution the agreement is fine unless the expected number of events is small, i.e. small sample size and/or heavy censoring and for a gamma distributed covariate the deviations are in the same direction as for logistic regression. The disagreement is largest for heavy censoring and gradually disappears with increasing sample size.

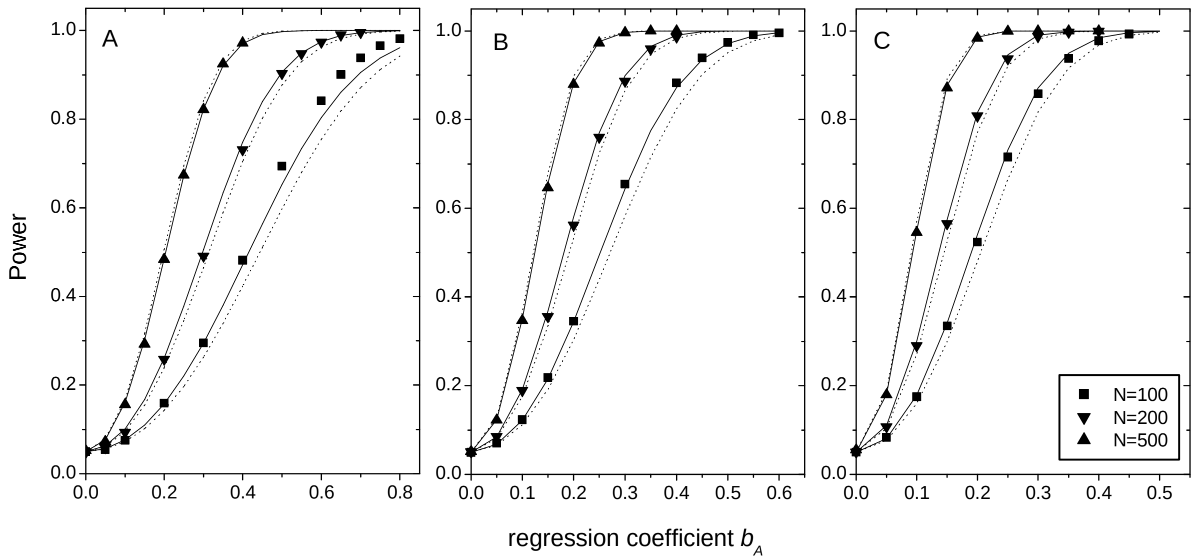

Figure 3. Simulated and estimated power of Wald's test in a Cox regression model with a covariate sampled from a standard normal distribution. Sample sizes are 100, 200 and 500. The proportion of censored values is 81% in panel A, 48% in panel B, and 9% in panel C. Solid curves show estimated power using the sample mean and standard deviation of the covariate. Dotted curve shows estimated power using the corresponding population values.

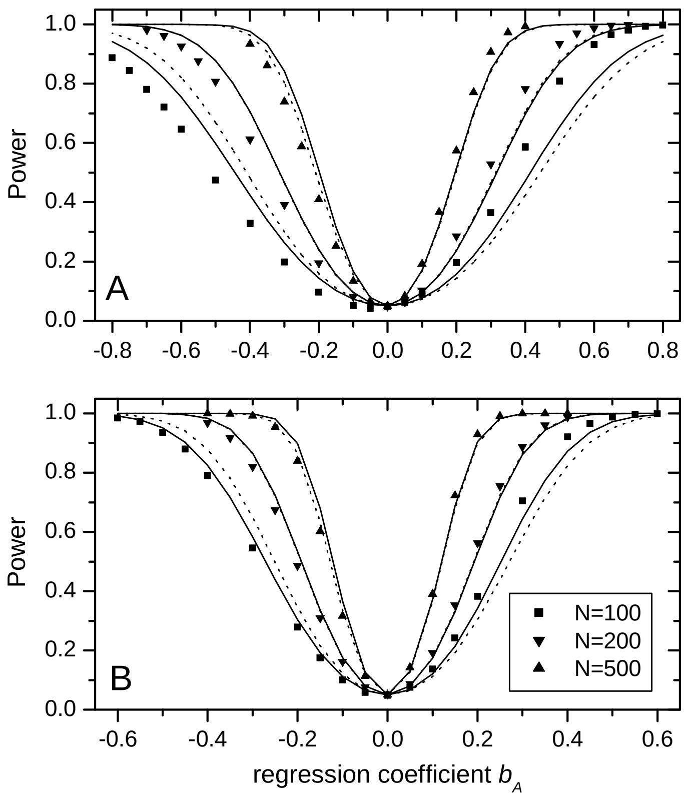

Figure 4. Simulated and estimated power of Wald's test in a Cox regression model with a covariate sampled from a standardized gamma distribution with shape parameter 3. Sample sizes are 100, 200 and 500. The proportion of censored values is 81% in panel A and 48% in panel B. Solid curves show estimated power using the sample mean and standard deviation of the covariate. Dotted curve shows estimated power using the corresponding population values.

The simulations with Weibull survival times were essentially identical to the one for exponential distributions with approximately the same probability of an event. This was to be expected as the asymptotic properties of the test statistic are a function of the number of events only[2].

The calculations based on the population values are also shown in Figures 1 to 4. Especially for small sample sizes the simulations agree less well with these curves indicating that sample size calculation is sensitive to random variation in the value of the standard deviation of the covariate. Although the correct population value is used in the calculation the actual power may differ as it depends on the covariate values in the sample.

Recent versions of a few statistical software packages[8,9] for sample size and power calculations include modules for logistic regression with a normal covariate implementing some of the alternative methods that have been developed during the last 15 years[10,11,12,13]. For selected scenarios with a standard normal covariate we compared the power reported by nQuery[8] and Power & Precision[9] with those based on the equivalent two-sample test and the simulated power. The procedure ROT0 used in nQuery's menu 'logistic regression with one covariate' uses the method developed by Hsieh[11]. The manual also explains how to implement the more accurate approach of Hsieh et al.[13] using the menu for unpaired t-test, MTT0U. The results shown in Figure 5 suggest that the present approach is at least as good as its competitors. It is well-known[8,13] that the approach used in nQuery's ROT0 procedure underestimates the power especially for distant alternatives. The results from Power & Precision show better agreement with the simulated results, but some underestimation is seen when the power is larger than 0.5. Our method essentially gives the same results as those based on Hsieh et al.[13] and in general these approaches are the most accurate.

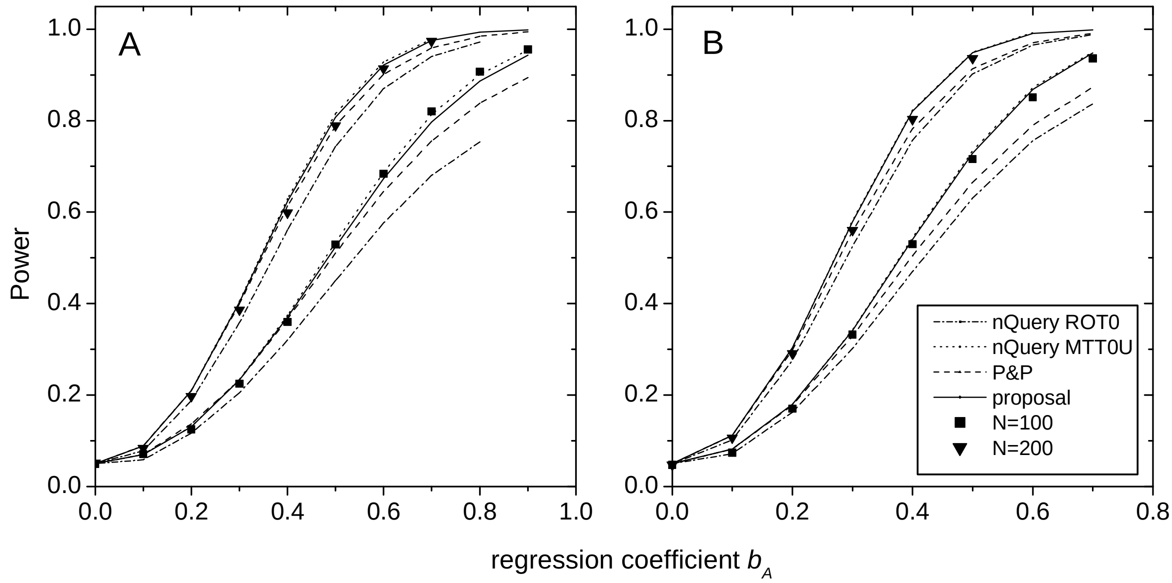

Figure 5. Simulated power, power estimated by nQuery, Power and Precision, and the proposed procedure for logistic regression with a covariate sampled from a standard normal distribution. Sample sizes are 100 and 200. In panel A $p(0)$ is 0.18 an in panel B $p(0)$ is 0.38. The sample mean and standard deviation are used for all estimated power curves.

DISCUSSION

The simulations indicate that a sample size derived from an equivalent two sample test provides a useful estimate of the sample size required for the regression problem for covariate distributions with at most a moderate skewness. For covariates with highly skewed distributions the results are less reliable, but such covariates would in practice often be transformed before they are entered into the regression model. Results for small sample sizes are also less accurate, but this is probably a general characteristic of calculations based on asymptotic theory.

We are not aware of any alternative method for easy sample size calculations in a Cox regression model with a continuous covariate, but for logistic regression Hsieh et al.[13] have described a relatively simple approach to sample size and power calculations based on a comparison of the covariate distributions in the two response categories. The present approach is different since the equivalent two-sample problem can be applied both to logistic regression and Cox regression models and may in fact be extended to any response model forming a generalised linear model.

Approaches to sample size or power calculations for logistic regression and Cox regression models with more than one independent variable are also needed. In special cases, e.g. the comparison of two slopes, it seems likely that an approach based on an equivalent two-sample test is feasible, but in general other solutions are required. For models with several covariates Hsieh et al.[13] recommends that the result of the univariate sample size calculation is modified by a so-called variance inflation factor such that

$$N_m = N_1 \big/ \left(1-\rho^2\right),$$

where N_1 and N_m are the required sample sizes with 1 and *m* covariates, respectively, and ρ is the multiple correlation coefficient between the covariate of interest and the remaining $m-1$ covariates. We have evaluated this approach, which was originally introduced by Whittemore[14], in a small simulation study of a logistic regression with two covariates taken as a sample from a two-dimensional normal distribution. The results suggest that this provides a simple and reasonably accurate method to expand the present approach to more complex problems.

REFERENCES

1. Schoenfeld DA, Richter JR. Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint. *Biometrics* 1982; **38**:163-170.

2. Schoenfeld DA. Sample size formula for the proportional-hazards regression model. *Biometrics* 1983; **39**:499-503.

3. *Stata Statistical Software: Release 6.0.* Stata Corporation: College Station TX, 1999.

4. Machin D, Campbell MJ. Statistical tables for the design of clinical trials. Blackwell Scientific Publications: Oxford, 1987, pp 10-12.

5. Morgan TM (1987). Nonparametric estimation of duration of accrual and total study length for clinical trials. *Biometrics* 1987; **43**:903-912.

6. Lachin JM, Foulkes MA. Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics* 1986; **42**:507-519.

7. Yateman NA, Skene AM. Sample sizes for proportional hazards survival studies with arbitrary patient entry and loss to follow-up distributions. *Statistics in Medicine* 1992; **11**:1103-1113.

8. Elashoff JD. nQuery Adivisor version 3.0 User's Guide. Statistical Solutions Ltd: Cork, 1999.

9. Borenstein M, Rothstein H, Cohen J. Power and precision. Biostat Inc.: Englewood NJ, 2001.

10. Self SG, Mauritsen RH. Power/sample size calculations for generalized linear models. *Biometrics* 1988; **44**:79-86.

11. Hsieh FY. Sample size tables for logistic regression. *Statistics in Medicine* 1989; **8**:795-802.

12. Self SG, Mauritsen RH, Ohara J (1992). Power calculation for likelihood ratio tests in generalized linear models. *Biometrics* 1992; **48**:31-39.

13. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine* 1998; **17**:1623-1634.

14. Whittemore AS. Sample size for logistic regression with small response probability. *Journal of American Statistical Association* 1981; **76**:27-32.