doi: 10.1111/j.1469-1809.2006.00316.x

Computing power in case-control association studies through the use of quadratic approximations: application to meta-statistics

M. Guedj,[1,2]* E. Della-Chiesa,[1] F. Picard[1] and G. Nuel[1]
[1]*Laboratoire Statistique et Genome, FRANCE*
[2]*Serono, FRANCE*

Summary

In the framework of case-control studies many different test statistics are available to measure the association of a marker with a given disease. Nevertheless, choosing one particular statistic can lead to very different conclusions. In the absence of a consensus for this choice, a tempting option is to evaluate the power of these different statistics *prior* to make any decision. We review the available methods dedicated to power computation and assess their respective reliability in treating a wide range of tests on a wide range of alternative models.

Considering Monte-Carlo, non-central chi-square and Delta-Method estimates, we evaluate empirical, asymptotic and numerical approaches. Additionally we introduce the use of the Delta-Method, extended to order 2, intended to provide better results than the traditional order-1 Delta-Method. Supplementary data can be found at: http://stat.genopole.cnrs.fr/software/dm2.

Keywords: Power, association tests, Delta-Method

Introduction

Case-control association studies are considered to be the simplest framework to help elucidate the genetic basis of complex diseases (Risch, 2000). Even if they have some weaknesses with ragard to potential confounding factors such as population stratification, they remain an important tool in genetical epidemiology and are often preferred to family-based studies (Zhao, 2000) due to the availability of data. Such an approach involves unrelated individuals split into cases, who are diagnosed with the disease of interest, and unaffected controls. This merely relies on the assumption that disease-related genetic determinants should accumulate among cases.

Tests of association are used as a first step in the analysis process. Various tests are proposed based on either genotypes (Table 1) such as the genotypic, Hardy-Weinberg equilibrium or Cochran-Armitage tests, or alleles such as the allelic test.

Since different single statistics can be used to test for association, another strategy is to combine them via meta-statistics with the hope of gaining power.

If they aim to establish an association between markers and disease, each test has a slightly different null hypothesis (H_0) and hence a different efficiency with respect to the underlying hypothesis. One way to compare them is to assess their power (π), defined as the ability of a test to reject the null hypothesis when the alternative hypothesis (H_1) is true. Power studies require the distribution of statistics under H_1.

This article reviews and discusses the most commonly used mathematical frameworks to approximate the H_1 distribution (and hence to compute π) in the context of genetic association studies. Our study includes two popular approaches. The first is empirical and based on Monte-Carlo simulations under the alternative hypothesis. The second is based on the

*Corresponding author: Mickaël Guedj, Laboratoire Statistique et Genome, 523 place des terrasses de l'Agora, 91000 EVRY, Tel: +33-1-6087-3800, Fax: +33-1-6087-3809, E-mail: guedj@genopole.cnrs.fr

Table 1 The genotypic contingency table

	aa	*aA*	*AA*	total
Diseased	D_0	D_1	D_2	n_D
Control	C_0	C_1	C_2	n_C
Total	n_0	n_1	n_2	n

asymptotic non-central chi-square distribution of the statistics under H_1. We compare these two approaches with the Delta-method with emphasis on its extension to order 2. As expected, non-central chi-square approximations appear to be very reliable (whenever available), while the order-1 Delta-Method is not. To treat non-explicit cases (combinations of statistics, for instance) for which computationally expensive Monte-Carlo simulations are usually considered, we show that the order-2 Delta-Method approximations are sufficiently efficient to represent a valid and cheaper alternative.

Method

Testing for association

Let us denote by x a case-control sample that is a realisation of the random variable X, and which can be represented by a genotypic contingency table (Table 1). To establish an association we consider a null hypothesis (H_0) used to test a particular distribution of the observations. To do so we consider a statistic defined as a function of these observations: $\mathcal{S} = f(X)$, and carefully chosen such that $\mathcal{S}$ grows when H_0 is less likely. Using the distribution of $\mathcal{S}$ under H_0, we can find a threshold (t_α) such as $\alpha = \mathbb{P}_{H_0}(\mathcal{S} \geq t_\alpha)$, where level α can be set to 5% for example.

Computing power

Using the distribution of $\mathcal{S}$ under an alternative association hypothesis H_1, we can compute the power $\pi(\alpha)$ of the test such that: $\pi(\alpha) = \mathbb{P}_{H_1}(\mathcal{S} \geq t_\alpha)$. To calculate $\pi(\alpha)$, the first step is to define a genetic model as well as the null and alternative hypotheses.

Genetic Model

Consider a bi-allelic disease susceptibility locus (DSL) with A being the susceptibility allele and a the other, p the frequency of allele A, and r_0, r_1 and r_2 the genotype frequencies in the general population. Assuming that the Hardy-Weinberg equilibrium (HWE) holds in the population, genotypic frequencies reduce to $r_2 = p^2$, $r_1 = 2p(1-p)$ and $r_0 = (1-p)^2$. Now we introduce the prevalence of the disease (K_p), and penetrances (f_i) associated with each genotype (i). Considering the relative risks (RR_i) such that $RR_i = \frac{f_i}{f_0}$ for $i = 1$ or 2, we define the four main modes of inheritance (MOI) corresponding to the modes of action of the DSL on the disease: recessive ($RR_1 = 1$), multiplicative ($RR_1 = \sqrt{RR_2}$), additive ($RR_1 = \frac{RR_2+1}{2}$) and dominant ($RR_1 = RR_2$).

Considering these parameters, we can easily derive $f_0 = K_p/(r_0 + RR_1.r_1 + RR_2.r_2)$ and $f_i = RR_i.f_0$ for $i = 1$ or 2. With the further assumption of an infinite population, the genotype distributions (D_0, D_1, D_2) in cases and (C_0, C_1, C_2) in controls are multinomial with parameters:

$$(D_0, D_1, D_2) \sim \mathcal{M}\left(n_D; \frac{f_0 r_0}{K_p}, \frac{f_1 r_1}{K_p}, \frac{f_2 r_2}{K_p}\right),$$

$$(C_0, C_1, C_2) \sim \mathcal{M}\left(n_C; \frac{(1-f_0) r_0}{1-K_p}, \frac{(1-f_1) r_1}{1-K_p}, \frac{(1-f_2) r_2}{1-K_p}\right).$$

In such a context, H_0: $\{RR_2 = 1\}$ and H_1: $\{RR2 \neq 1\}$. Once the alternative hypothesis is explicit, power can be calculated using one of the following frameworks to approximate the distribution of $\mathcal{S}$ under H_1.

Monte-Carlo estimation

As long as it is possible to generate a case-control sample $X = \begin{cases} D_0 & D_1 & D_2 \\ C_0 & C_1 & C_2 \end{cases}$. under H_1, it is very easy to estimate the power. We first draw N samples denoting $x^{(i)}$ the i^{th} sample. From this sample we get N statistics $s^{(1)}, \ldots, s^{(N)}$ from which we get the estimation of the power:

$$\widehat{\pi}(\alpha) = \frac{\sharp\left\{s^{(i)} \geq t_\alpha\right\}}{N}.$$

This well-known method is often very easy to perform and is consequently widely used, particularly in the field of statistical genetics when alternative distributions are hard to calculate analytically (Longmate, 2001). But such an approach may generally require a lot of time to reach a given level of precision. As $\widehat{\pi}$ is distributed according to a binomial distribution, using the central limit theorem

we find that $\widehat{\pi} \sim \mathcal{N}(\pi, \pi(1-\pi)/N)$, which gives the following 95% confidence interval:

$$\left[\widehat{\pi} - 1.96\frac{\sqrt{\widehat{\pi}(1-\widehat{\pi})}}{\sqrt{N}};\quad \widehat{\pi} + 1.96\frac{\sqrt{\widehat{\pi}(1-\widehat{\pi})}}{\sqrt{N}}\right].$$

Consequently the precision of the power estimate increases with speed $1/\sqrt{N}$. One could remark that the same method can be used to estimate the threshold of tests involving statistics for which the distribution under H_0 is not easily available (*e.g.* meta-statistics).

Asymptotic non-centrality parameter

Mitra (1958) demonstrated that, under H_1, the asymptotic distribution of a chi-square frequency test applied to a $2 \times c$ contingency table follows a non-central chi-square distribution $\chi'^2(k, \lambda)$, where k is the degree of freedom and λ the non-centrality parameter, such that

$$\lambda = N_1 N_2 \times \sum_{j=1}^{c} \frac{(p_{1j} - p_{2j})^2}{N_1 p_{1j} + N_2 p_{2j}},$$

with p_{ij} the frequency of case ij and N_1, N_2 the total counts of the first and second row. Mitra derived the asymptotic power for the test:

$$\pi(\alpha) \underset{\infty}{\rightarrow} 1 - \chi'^2_{1-t_\alpha}(k, \lambda).$$

Given the expression of the non-centrality parameter, this approach can be adapted to any statistic following a chi-square distribution under H_0 (see below for the particular case of the trend test) and is appropriate when sample sizes are large enough. It has recently been presented as an appealing and fast way to approximate power in association studies (Sham et al. 2000; Gordon et al. 2002; Kang *et al*. 2004).

Delta-Method

The Delta-Method is used to approximate the distribution of $\mathcal{S}$ with X. The multinomial distribution of X (derived from the genetic model) is asymptotically distributed according to a Gaussian distribution $\mathcal{N}(M, \Sigma)$. Using an order-1 Taylor development of $\mathcal{S} = f(X)$ around M we hence approximate $\mathcal{S}$ by:

$$\mathcal{S} \simeq f(M) + {}^t(X - M) \times \nabla f(M),$$

where ^{t} is the transpose operator and ∇f is the gradient of f. This 1-order development allows us to approximate the distribution of $\mathcal{S}$ by a Gaussian distribution $\mathcal{N}(m, \sigma^2)$, with $m = f(M)$ and $\sigma^2 = {}^t\nabla f(M) \times \Sigma \times \nabla f(M)$. Then we have:

$$\pi(\alpha) \underset{\infty}{\rightarrow} 1 - \Phi\left(\frac{t_\alpha - m}{\sigma}\right)$$

where Φ is the cumulative distribution function (CDF) of a Gaussian variable with zero mean and a variance of one. Of course, the closer the distribution under H_1 is to a Gaussian distribution, the better will be this 1-order approximation.

For cases where the Gaussian distribution of the statistic under H_1 is not realistic, we propose to use a order-2 Taylor expansion around M. We hence get a more precise approximation based on the distribution of a quadratic form in normal variables (QFNV):

$$\mathcal{S} \simeq f(M) + {}^t(X - M) \times \nabla f(M) + \frac{1}{2}{}^t(X - M) \times \nabla^2 f(M) \times (X - M),$$

where $\nabla^2 f$ is the Hessian of f.

In the case of the first order development, the computation of power only requires evaluating the CDF of a normal distribution. With the second order development, however, the distribution of $\mathcal{S}$ is approximated by a combination of chi-squares and the CDF is not straighforward to derive. Technical details can be found in Appendix 1 and derivations of the distribution for the statistics considered are available at: http://stat.genopole.cnrs.fr/software/dm2.

Application

Statistics considered

Here we consider four statistics.

(i) The *genotypic test* compares genotypic frequencies between affected and unaffected subjects by using the Pearson's chi-square statistic:

$$\mathcal{S}_G = \sum_{i=0}^{2} \frac{\left(D_i - \frac{n_D \times n_i}{n}\right)^2}{\frac{n_D \times n_i}{n}} + \frac{\left(C_i - \frac{n_C \times n_i}{n}\right)^2}{\frac{n_C \times n_i}{n}} \underset{H_0}{\sim} \chi^2(2),$$

From the formula given in (2.2.3) and the parameters of the genetic model (2.2.1), we can derive the non-centrality parameter for this statistic:

$$\lambda_G = n_D n_C \times \sum_{i=0}^{2} \frac{\left(\frac{f_i r_i}{Kp} - \frac{(1-f_i)r_i}{1-Kp}\right)^2}{n_D \frac{f_i r_i}{Kp} + n_C \frac{(1-f_i)r_i}{1-Kp}}$$

and $\mathcal{S}_G \widetilde{H_1} \chi^2(2, \lambda_G)$

(ii) Another test based on genotypes is the *Cochran-Armitage test for trends* (Armitage, 1995). It measures a linear trend in proportions weighted by a dose effect score x_i associated to each column with x_i corresponding to the number of susceptibility allele:

$$\mathcal{S}_T = \frac{n.[n.(D_1 + 2D_2) - n_D.(n_1 + 2n_2)]^2}{n_D n_C.[n.(n_1 + 4n_2) - (n_1 + 2n_2)^2]} \widetilde{H_0} \chi^2(1).$$

For this particular case (trend test with three categories) Gordon *et al.* (2005) derived the expression of the non-centrality parameter, based on previous work of Chapman & Nam (1968). With our notation, it comes down to:

$$\lambda_T = n_D n_C \times \frac{\left[\sum x_i \left(\frac{(1-f_i)r_i}{1-Kp} - \frac{f_i r_i}{Kp}\right)\right]^2}{\sum x_i^2 \left(n_D \frac{f_i r_i}{Kp} + n_C \frac{(1-f_i)r i}{Kp}\right) - \frac{\left[\sum x_i \left(n_D \frac{f_i r_i}{Kp} + n_C \frac{(1-f_i)r i}{Kp}\right)\right]^2}{n}}$$

and $\mathcal{S}_T \widetilde{H_1} \chi^2(1, \lambda_T)$

Note that chi-square approximations are appropriate when sample sizes are in accordance with Cochran's condition (each expected cell count > 5 -, Cochran, 1952).

(iii) Another strategy is to combine simple statistics via meta-statistics with the hope of gaining power. Nevertheless, the actual null hypothesis tested is not explicit and distributions of such statistics (under H_0 or H_1) are not easy to assess out of Monte-Carlo simulations. Our aim considering $\mathcal{S}_\Sigma = \mathcal{S}_G + \mathcal{S}_T$ and $\mathcal{S}_\Pi = \mathcal{S}_G \times \mathcal{S}_T$ is to assess the efficiency of competing approaches to handle their power computation.

Simulations

Simulations are performed using the susceptibility allele frequency (p) as a factor of variation. All simulations are considered for a prevalence $K_p = 0.05$, $n_D = n_C = 500$ and the four MOIs ($RR_2 = 1.5$). Each Monte-Carlo estimate of power is carried out on the basis of $N = 10,000$ simulations, and is considered as a reference to compare with the other approaches. Using this approach to compute a power $\widehat{\pi}$, we get a 95% confidence interval of radius $0.0196\sqrt{\widehat{\pi}(1-\widehat{\pi})}$ centered on $\widehat{\pi}$. For example, this radius gives 0.588% for $\widehat{\pi} = 10\%$ (or 90%), 0.784% for $\widehat{\pi} = 20\%$ (or 80%) and is always smaller than 0.98% (case $\widehat{\pi} = 50\%$).

Results

Results concerning $\mathcal{S}_G$ and $\mathcal{S}_T$ are compiled in Figure 1. For the set of parameters considered, the additive and multiplicative models give very close results so we have displayed them only for the additive, recessive and dominant models.

The non-central chi-square approach (NC) is fully adapted to chi-square distributed statistics and hence gives accurate estimates of power. As non-central chi-square distributions are particular cases of QFNV, the order-2 Delta-Method (DM2) unsurprisingly also gives good results. By comparison, the order-1 Delta-Method (DM1) underestimates the power in the two cases. This underscores the fact that the Gaussian approximation under H_1 made by this approach is not realistic. However it provides better estimates for the trend test than for the genotypic test. This variation is due to the fact that the distribution of $\mathcal{S}_T$ under H_0 and H_1 is closer to a Gaussian distribution - that requires this approach - than $\mathcal{S}_G$. For instance, the expected value for the Wilk-Shapiro statistic test for normality is 0.69 for one degree-of-freedom chi-square distributed samples and 0.81 for two degrees-of-freedom chi-square distributed ones.

In the literature it has been suggested that factors such as the ratio of cases to controls, minor allele frequency and total sample size affect the accuracy of the analytic power calculations (Ji *et al.* 2005). We investigated such effects considering the genotypic and trend tests for the four MOIs and values of (0.04, 0.2, 1), (0.2,0.3,0.4,0.5) and (40,200,1000) for the ratio, the minor allele frequency and the total sample size,

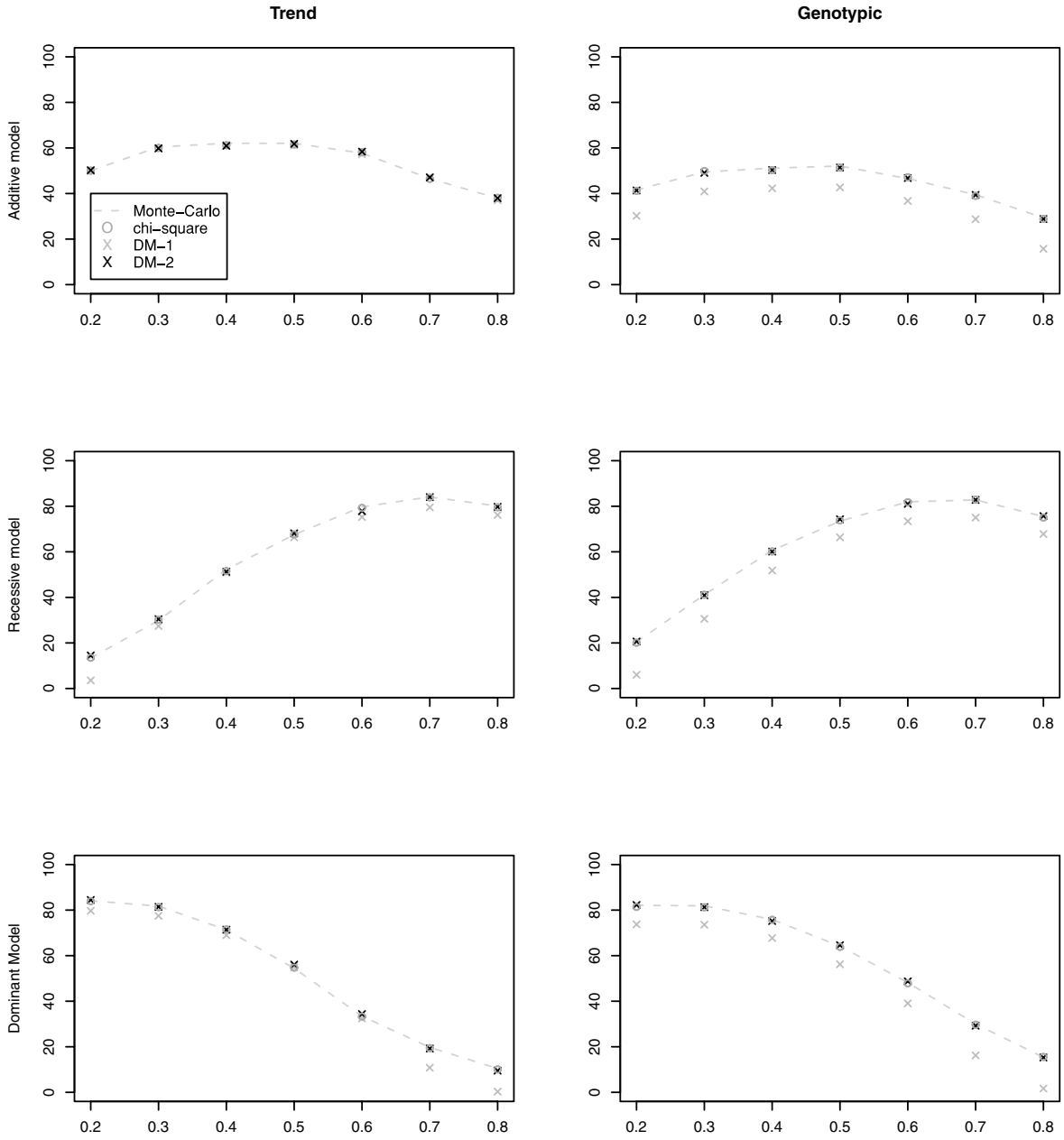

Figure 1 Power estimation (at the 5% significance level) for the trend and genotypic tests according to the allele frequency (p).

respectively. Normalized absolute differences have been computed for the NC and DM2 calculations (data not shown). However, we did not observe any clear effect of these factors on the accuracy of the analytic calculations.

Figure 2 presents the results for the two meta-statistics ($\mathcal{S}_\Sigma$ and $\mathcal{S}_\Pi$). In these cases, the non-central chi-square approach is not applicable. Even if $\mathcal{S}_\Sigma$ is the sum of two chi-square distributed statistics, $\mathcal{S}_G$ and $\mathcal{S}_T$ are not independent and hence $\mathcal{S}_\Sigma$ is not merely distributed according to a three degrees-of-freedom chi-square distribution under H_0. DM1 still badly estimates power. DM2 is really efficient to treat $\mathcal{S}_\Sigma$. As previously underlined, a linear combination of (dependent or not) chi-square distributed statistics is a QFNV which explains that DM2 works well on $\mathcal{S}_\Sigma$. Nevertheless DM2 does not manage to assess the power of $\mathcal{S}_\Pi$. We can imagine that such a product of chi-square distributed statistics would have required the use of the Delta-Method to a higher

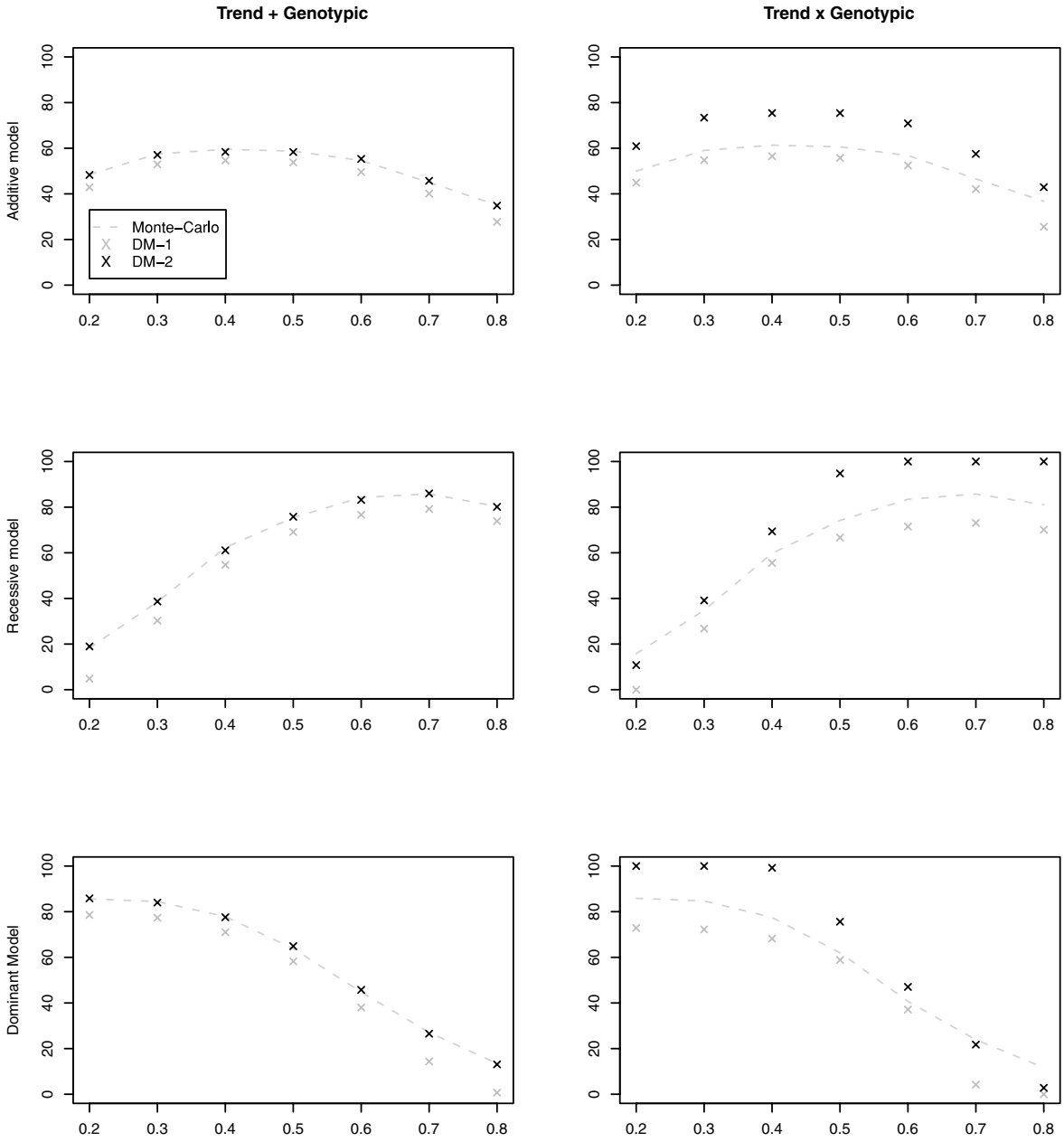

Figure 2 Power estimation (at the 5% significance level) for the meta-statistics according to the allele frequency (p).

order, for which the determination of the CDF would have been numerically very expensive and unrealistic in practice.

In terms of comparison between the four strategies considered, $\mathcal{S}_T$, $\mathcal{S}_G$, $\mathcal{S}_\Sigma$ and $\mathcal{S}_\Pi$ (Figure 3), meta-statistics power estimates mainly lie between trend and genotypic ones and hence do not clearly represent a better alternative to single-statistics. However, they do more than merely averaging power estimates of single statistics, and hence can appear as a clever alternative to combine efficiency according to the model.

Discussion

Studying power is an important tool in statistics to compare the efficiency of different tests or to help design a study. With the accumulation of new analysis methods that have recently arisen from the accumulation of

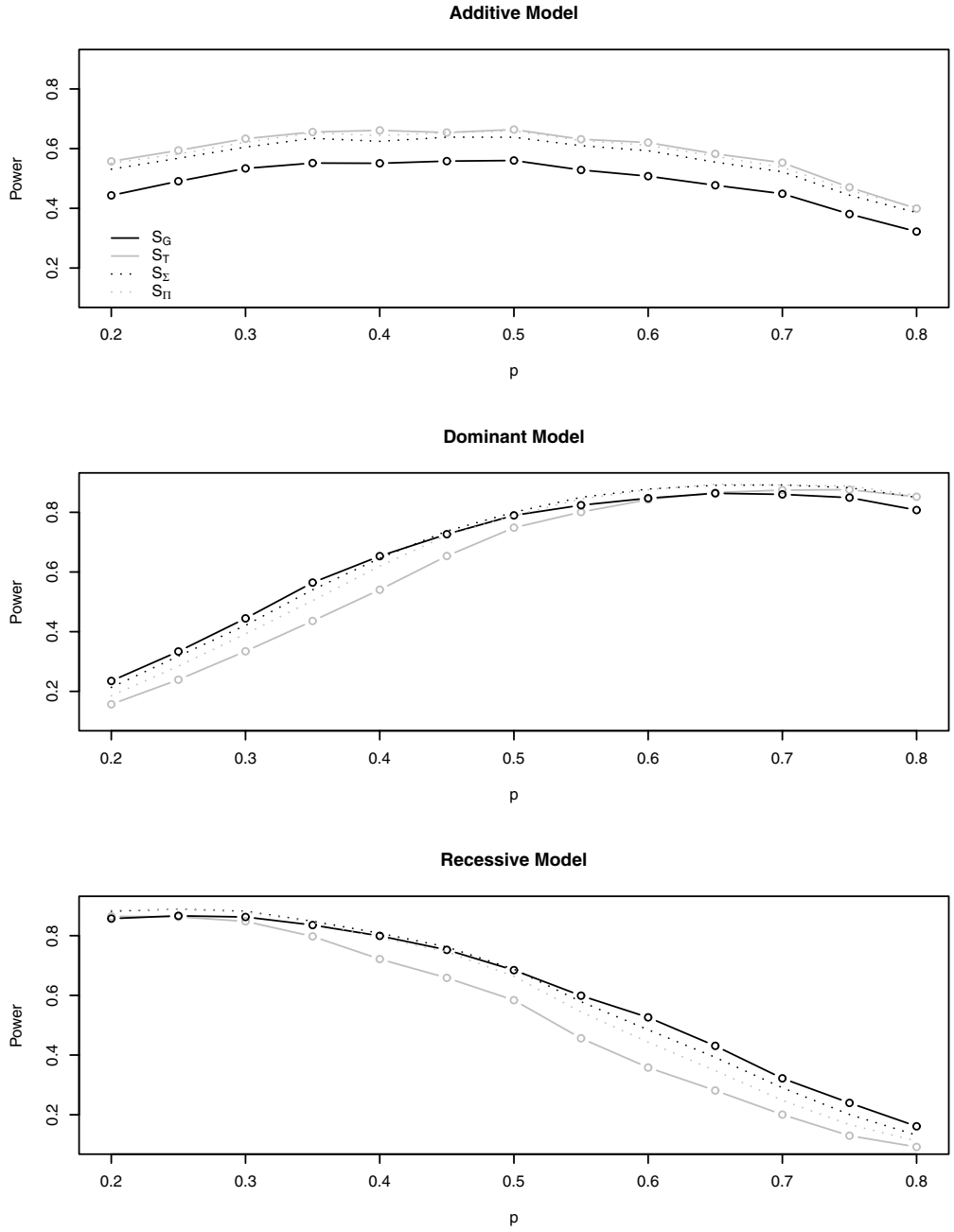

Figure 3 Power comparison: this figure compares the power of the four statistics $\mathcal{S}_G$, $\mathcal{S}_T$, $\mathcal{S}_\Sigma$ and $\mathcal{S}_\Pi$ according to the allele frequency (p). It is done for the additive, recessive and dominant models. Power is computed by Monte-Carlo at the 5% significance level.

large-scale data, statistical genetics does not escape this rule. In this article we focus on the computation of power in the context of simple-marker analyzes *via* the genotypic and trend statistics, as well as simple combinations of them.

Easy to implement, Monte-Carlo simulations are often the preferred approach to compute power estimations. Nevertheless they are computationally expensive since the precision of the estimates is directly dependent on the number of simulations performed. In particular the length of the confidence interval decreases with $1/\sqrt{N}$, and hence evolves quite slowly with N. Computing power through the non-centrality parameter is logically well adapted for statistics distributed

according to a chi-square distribution under H_0. The order-1 Delta-Method is based on a Gaussian distribution of the statistic. As a result it is not efficient in the situation considered here. In the literature approaches based on the order-1 Delta-Method have been successfully developed (Slager & Schaid, 2001; Jackson et al. 2002) to compute accurate power approximations for allelic and trend tests (Slager & Schaid, 2001; Jackson et al. 2002). The required Gaussian distributions of the statistics were obtained by the authors considering $\mathcal{Z} \sim \mathcal{N}(0, 1)$ such that $(\mathcal{Z})^2 = \mathcal{S}$ instead of $\mathcal{S}$ directly (as we have done here). This approach provides very good power approximations. However, its application is restrained to z-scores or by extension to 1 degree-of-freedom chi-square distributed statistics. It is hence less general than the non-central chi-square approach. To go further, we introduce the use of the order-2 Delta-Method. This approach provides good estimates and can be used to treat simple statistics and linear combinations of them, which is an advantage over other approaches. Besides a less straightforward CDF evaluation, it represents much less computionally expensive alternative to Monte-Carlo simulations, more general than the non-central chi-square framework and more accurate than the order-1 Delta-Method.

This work has been restricted to the study of the trend and genotypic tests under alternatives that differ in the susceptibility allele frequency and the MOI only. However, our conclusions can easily be extended to other simple-marker tests (Hardy-Weinberg and allelic tests, for instance) based on more complicated meta-statistics, and applied to more elaborate alternative models taking, for instance, the coefficient of consanguinity, linkage disequilibrium and genotyping errors into account.

Even if they fail to provide greater power estimates than single statistics, meta-statistics do not suffer from substantial power loss when compared with the best single statistic for each of the situations considered in this work. We thus suggest that meta-statistics may provide a useful means for combining such tests.

If they fail to provide better results than single-statistics, meta-statistics do not present a sensible loss of power compared to the best simple statistic in each situation considered, and hence appear to be a clever possible way to combine such tests.

Acknowledgements

The authors would like to thank Bernard Prum's team as well as members of the Serono Pharmaceutical Institute, in particular Jérôme Wojcik and Hiroaki Tanaka for encouraging this work.

References

Armitage, P. (1995). Tests for linear trends in proportions and frequencies. *Biometrics* **11**, 375–386.

Chapman, D. and Nam, J. (1968). Asymptotic power of chi-square tests for linear trends in proportions. *Biometrics* **24**, 315–327.

Cochran, W. (1952). The chi-square test of goodness of fit. *Annals of Human Genetics* **23**, 315–345.

Davies, R. B. (1973). Numerical inversion of characteristic function. *Biometrika* **60**, 481–482.

Davies, R. B. (1980). The distribution of a linear combination of chi-square random variables. *Appl. Stat* **29**, 323–333.

Gordon, D., Finch, S. J., Nothnagel, M. and Ott, J. (2002). Power and sample size calculations for case-control genetic association tests when errors are present: application to single nucleotid polymorphisms. *Human Heredity* **54**, 22–33.

Gordon, D., Haynes, C., Blumenfeld, J. and Finch, S. (2005). Pawe-3d: visualizing power for association with error in case-control genetic studies of complex traits. *Bioinformatics* **21**, 3935–3937.

Jackson, M., Genin, E., Knapp, M. and Ji. E. (2002). Accurate power approximation for X-tests in case-control association studies of complex disease genes. *Annals of Human Genetics* **66**, 307–321.

Ji, F., Yang, Y., Haynes, C., Finch, S. and Gordon, D. (2005). Computing asymptotic power and sample size for case-control genetic association studies in the presence of phenotype and/or genotype misclassification errors. *Statistics Applied to Genetic and Molecular Biology* 4.

Kang, S. J., Gordon, D. and Finch, S. J. (2004). What snp genotyping errors are most costly for genetic association studies. *Genetic epidemiology* **26**, 132–141.

Longmate, J. A. (2001). Complexity and power in case-control association studies. *American Journal Human Genetics* **68**, 1229–1237.

Lu, Z.-H. and King, M. L. (2002). Improving the numerical techniqie for computing the accumulated distribution of a quadratic form in normal variables. *Econometric review* **21**, 149–165.

Mitra, S. (1958). On the limiting power function of the frequency chi-square test. *Annals of Mathematical Statistics* **29**, 1221–1233.

Risch, N. (2000). Searching for genes in complex diseases: lessons from systemic lupus erythematosus. *American society for clinical investigation* **105**, 1503–1506.

Sham, P., Cherny, S., Purcell, S. and Hewitt, J. (2000). Power of linkage versus association analysis of quatitative traits, by use of variance components models, for sibship data. *American Journal Human Genetics* **66**, 1616–1630.

Slager, S. and Schaid, D. (2001). Case-control studies of genetic markers: power and sample size approximations for armitage's test for trend. *Human Heredity* **52**, 149–153.

Zhao, H. (2000). Family-based association studies. *Statistical Methods in Medical Research* **9**, 563–587.

Appendix A: Quadratic Form in Normal Variables

In this appendix we propose to recall the definition of this distribution and explain how it is possible to compute its cumulative distribution function (CDF).

Definition 1 If $X \sim \mathcal{N}(\mu, \Sigma)$ *is a dimension* $d \geq 1$ *(column) vector of normal variables we call*

$$Q = A + BX + X'CX$$

a quadratic form in normal variables *(QFNV) of dimension d* with *parameters* $A \in \mathbb{R}$, $B \in \mathbb{R}^{(1,d)}$ *and* $C \in \mathbb{R}^{(d,d)}$ and with $\mu \in \mathbb{R}^{(d,1)}$ *and* $\Sigma \in \mathbb{R}^{(d,d)}$ *are mean and covariance matrix of the normal variables.*

In particular, a linear combination of (central or not, independent or not) chi-square is a QFNV.

For non-degenerate parameters, it is possible to express a QFNV as a linear combination of independent non-central chi-square distribution (Lu and King, 2002). Namely:

Proposition 2 *For any QFNV Q with non-singular covariance matrix* Σ *and matrix* C, $1 \leq n \leq d$, $\lambda_j \in \mathbb{R}$, $d_j \in \mathbb{N}^*$ *and* $\nu_j > 0$ *(for all* $1 \leq j \leq n$*) such as*

$$Q = K + \sum_{j=1}^{n} \lambda_j \chi_j$$

where $K \in \mathbb{R}$ *and* $\chi_j \sim \chi^2(d_j, \nu_j)$ *are independent chi-square variables with* d_j *degrees of freedom and* ν_j *non centrality* parameters.

Proof. We first factorize Q in

$$Q = \underbrace{\left(A - \frac{BC^{-1}B'}{4}\right)}_{K} + \underbrace{\left(X + \frac{C^{-1}B'}{2}\right)' C \left(X + \frac{C^{-1}B'}{2}\right)}_{Y'CY}$$

where $Y \sim \mathcal{N}(\tilde{\mu}, \Sigma)$ with $\tilde{\mu} = \mu + C^{-1}B'/2$. We consider then the linear transformation $Z = \Sigma^{-1/2} Y$ so $Z \sim \mathcal{N}(\Sigma^{-1/2}\tilde{\mu}, I)$ and

$$Q = K + Z' \underbrace{\Sigma^{-1/2} C \left(\Sigma^{1/2}\right)'}_{\tilde{C}} Z$$

We then consider the orthogonal matrix P of the eigenvector of $\tilde{C}$ and denote by $D = P'\tilde{C}P$ the diagonal matrix of the corresponding eigenvalues. With $W = P^{-1} Y$ we get

$$Q = K + W'DW$$

with $W \sim \mathcal{N}(P^{-1}\Sigma^{-1/2}\tilde{\mu}, I)$. For $1 \leq j \leq n$, we denote by λ^j the (distinct) eigenvalue and by d_j their orders of multiplicity (one should note that they are also those of $C\Sigma$ or ΣC). Finally, we consider $\nu_j = \sum_{q=1}^{d_j} \gamma_{j,q}^2$ where $\gamma_{j,q}$ are the elements of $P^{-1}\Sigma^{-1/2}\tilde{\mu}$ corresponding to the same eigenvalue λ_j and the result is established.

From now, we hence focus of the numerical CDF evaluation of

$$Q = \sum_{j=1}^{n} \lambda_j \chi^2(d_j, \nu_j)$$

a linear combination of independent and non central chi-square distributions.

A numerical inversion of the characteristic function is then possible, resulting through truncation and trapezoidal integration (Davies, 1973; Davies, 1980) in the following formula

$$\mathbb{P}(Q < c) = \frac{1}{2} - \sum_{m=0}^{M} \left(\frac{\sin\{\theta_c[(m + 0.5)\Delta]\}}{\pi(m + 0.5)\gamma[(m + 0.5)\Delta]} \right)$$

where Δ is the (small) step interval, M the (large) number of step intervals, and $U = (M + 0.5)\delta$ the truncation value. The functions θ_c and γ are given by

$$\theta_c(u) = \sum_{j=1}^{n} \left[\frac{d_j}{2} \tan^{-1}(2u\lambda_j) + \nu_j u \lambda_j (1 + 4u^2\lambda_j^2)^{-1} \right] - cu$$

and

$$\gamma(u) = \prod_{j=1}^{n} (1 + 4u^2\lambda_j^2)^{d_j/4} \exp\left(2u^2 \sum_{j=1}^{n} \frac{\nu_j \lambda_j^2}{1 + 4u^2\lambda_j^2} \right)$$

The numerical evaluation of the CDF using this formula leads to an error of truncation ε_T depending on the truncation bound U, and to an error of integration ε_I depending on the step interval Δ. There exist many concurrent ways to choose both these values and Lu & King (2002) provide a complete review of them.