Testing linear forms of variance components by generalized fixed–level tests

Boris Weimann [1]

Department of Statistics
University of Dortmund
Vogelpothsweg 87
D–44221 Dortmund
Germany

Abstract

This report extends the technique of testing single variance components with generalized fixed–level tests — in situations when nuisance parameters make exact testing impossible — to the more general way of testing hypotheses on linear forms of variance components. An extension of the definition of a generalized test variable leads to a generalized fixed–level test for arbitrary linear hypotheses on variance components in balanced mixed linear models of the ANOVA–type. For point null hypotheses an alternative for the known method is given, which ist straightforward in contrast to the classic form. An example (2–way nested classification with random effects) illustrates the way how to use the results and simulation studies are carried out to prove the quality of the presented methods.

Key Words: Variance components, generalized fixed–level test, mixed linear models, nuisance parameters, linear hypotheses, approximate testing.

1 Introduction

For various statistical models there do not exist exact tests for the hypotheses of interest because of nuisance parameters. Such situations can always occur if the model includes two or more random effects. Typical representatives of this class of models are the mixed linear models.

Literature is widely available for approximative and asymptotic tests for many very special situations. A classical example is the approximative test by Satterthwaite (1946), an F–test with adapted degrees of freedom for hypotheses on single variance components. In a paper

[1]This research was supported by the Deutsche Forschungsgesellschaft (DFG); Sonderforschungsbereich 475

by Thursby (1989), a number of approximative tests is compared. All of these procedures are only of restricted usability.

The concept of testing with generalized p–values was introduced by Tsui and Weerahandi (1989). Weerahandi (1991) and Zhou and Mathew (1994) used generalized p–values for tests on variance components in their papers, where the hypotheses were usually only formulated for single variance components.

In this paper the test with generalized p–values is extended to the case of arbitrary linear hypotheses in balanced mixed linear models. In order to do this, the definition of a generalized test variable which was first introduced by Tsui and Weerahandi (1989) has to be extended, because it proves to be too restrictive. The new procedure is demonstrated on the example of the hierarchical two–way classification. Simulation studies show that the new method usually holds the nominal significance level quite well, even in the case of small data sets.

Two–sided hypotheses which are to be tested against composite alternatives are a problem mostly unregarded up to now. Weerahandi (1995) proposed a solution, but he did not formulate a concrete construction principle for the test procedure. This paper will show up a straightforward procedure which, as far as the significance level is concerned, is comparable to tests for one–sided hypotheses.

In variance component models the problem of a quite small power may occur for some parts of the alternative for any kind of test. For some constellations of parameters the empirical power functions are given for a special testing problem in the above mentioned hierarchical two–way classification. A detailed analysis of the power function will be a subject of further research.

The restriction to balanced models can be abandoned in some situations. Khuri (1990) showed that generalized p–values can be applied if the model in unbalanced on the last stage only. An application of this procedure to testing linear hypotheses and a generalization to arbitrary forms of unbalancedness is desirable.

2 General testing principle

Consider an observable random vector Y with the cumulative distribution function $F(y, \xi)$, where $\xi = (\vartheta, \delta^T)^T$ is a vector of unknown parameters, ϑ being the parameter of interest, and δ a vector of nuisance parameters. Let Υ be the sample space of possible values of Y and Θ be the parameter space of ϑ. An observation of Y is denoted by y.

Definition 2.1
A random variable of the form $T = T(Y, y, \xi)$ is said to be a *generalized test variable* if it has the following three properties:

1. $t_{obs} = T(y, y, \xi)$ does not depend on unknown parameters.
2. When ϑ is specified, T has a probability distribution that is free of nuisance parameters.
3. For fixed y and δ, $\Pr(T \leq t|\vartheta)$ is a monotone function in ϑ for any given t.

Without loss of generality the first property can be considered to be redundant, because if it is not satisfied we can cross over to the transformation $\tilde{T} := T(Y, y, \xi) - T(y, y, \xi)$ and impose properties 2 and 3 on $\tilde{T}$.

Property 2 is imposed to ensure that p–values based on generalized test variables are computable when ϑ is specified. Property 3 ensures that the sample space can be stochastically ordered on the basis of the generalized test variable. If $\Pr(T > t)$ is a nondecreasing function in ϑ, then T is said to be *stochastically increasing* in ϑ.

Consider the problem of testing one–sided hypotheses of the form

$$(1) \quad H_0 : \ \vartheta \leq \vartheta_0 \quad \text{vs.} \quad H_1 : \ \vartheta > \vartheta_0 \ ,$$

where ϑ_0 is a prespecified value of the parameter ϑ.

Definition 2.2
Let $T = T(Y, y, \xi)$ be a stochastically increasing (in the parameter of interest ϑ) test variable according to definition 2.1. Then, the subset of the sample space defined by

$$(2) \quad C_y(\xi) \ = \ \{Y \in \Upsilon | T(Y, y, \xi) \geq t_{obs}\}$$

is said to be a generalized extreme region for testing H_0 against H_1.

Definition 2.3
If $C_y(\xi)$ is a generalized extreme region according to (2), then

$$(3) \quad p(t_{obs}) \ = \ \sup_{\vartheta \leq \vartheta_0} \Pr(Y \in C_y(\xi)|\vartheta)$$

is said to be its *generalized p–value* for testing H_0.

Corollary 2.4
The generalized p–value according to (3) is equivalent to

$$(4) \quad p(t_{obs}) \;=\; \Pr(Y \in C_y(\xi) | \vartheta = \vartheta_0) \ ,$$

which is easy to determine.

Proof:
This follows directly from property 3 of a generalized test variable: if T is stochastically increasing in ϑ, the supremum over $\Theta_0 = \{\vartheta | \vartheta \leq \vartheta_0\}$ is obtained on the upper boundary of Θ_0. □

Definition 2.5
Let $p(t_{obs})$ be a generalized p–value on a continuous generalized test variable $T = T(Y, y, \xi)$. Let $H_0 : \ \vartheta \in \Theta_0$ be the null hypothesis being tested against the alternative $H_1 : \ \vartheta \in \Theta_1$. Then, the rule defined as

$$(5) \quad \text{reject } H_0 \text{ if } p(t_{obs}) \leq \alpha$$

is said to be a *generalized fixed–level test* of level α.

Corollary 2.6
The generalized p–value according to (3) as a function of the observed value t_{obs} resp. y is not uniformly distributed over the interval $[0, 1]$. For that reason, the generalized fixed–level test according to (5) is not an exact test of level α, but an approximate one.

Proof:
Assume a continuous generalized test variable T. The generalized p–value

$$p(t_{obs}) = \Pr(T(Y, y, \xi) \geq t_{obs} | \vartheta = \vartheta_0) = 1 - F_T(t_{obs}, \vartheta_0)$$

is a function of the observed value of T. Considering the observed $t_{obs} = T(y, y, \xi)$ as a random variable $T^* = T(Y, Y, \xi)$ leads in general to different distributions for T and T^*, because only the distribution of T depends on the observed value t_{obs}. From the probability integral transform it follows, that $F_T(T)$ has a uniform distribution over the interval $[0, 1]$. Because of

$$\begin{aligned} p(T^*) &= \Pr(T(Y, y, \xi) \geq T(Y, Y, \xi) | \vartheta = \vartheta_0) \\ &= 1 - F_T(T^*, \vartheta_0) \\ &\not\sim 1 - F_T(T) \end{aligned}$$

it follows, that $p(T^*)$ in general does not have an uniform distribution over $[0, 1]$. □

Definition 2.7
Let $\pi(y, \vartheta) := \Pr(Y \in C_y(\xi)|\vartheta)$ be the *data–based power function of* T. A test based on a generalized extreme region $C_y(\xi)$ is said to be *p–unbiased* if

$$(6) \quad \pi(y, \vartheta) \geq \pi(y, \vartheta_0) \quad \text{for all } \vartheta \in \Theta_1 \ ,$$

and *p–similar* (on the boundary) if, given any $y \in \Upsilon$,

$$(7) \quad \pi(y, \vartheta_0) = p(t_{obs})$$

does not depend on the nuisance parameters δ, where $p(t_{obs})$ ist the generalized p–value according to (3).

This concept of testing with generalized p–values was first introduced by Tsui and Weerahandi (1989) and is presented in detail in Weerahandi (1995).

3 Testing point null hypotheses

Consider point null hypotheses and composite alternative hypotheses of the form

$$(8) \quad H_0: \ \vartheta = \vartheta_0 \quad \text{vs.} \quad H_1: \ \vartheta \neq \vartheta_0$$

where ϑ_0 is a particular value of the parameter that has been specified.

In such situations Weerahandi extends definition 2.1 by substituting

4. Given any fixed t_{obs} and δ, the probability $\Pr(T \in C_y(\xi))$ is a nondecreasing function of (i) $\vartheta - \vartheta_0$ when $\vartheta \geq \vartheta_0$, and (ii) $\vartheta_0 - \vartheta$ when $\vartheta < \vartheta_0$.

for property 3 of a generalized test variable.

By this definition the data–based power function $\pi(y, \vartheta)$ increases on Θ_1 with the distance to Θ_0. Particularly the resulting generalized fixed–level test is p–unbiased. A problem occurs when the generalized extreme region is to be constructed, because the construction is not as obvious and clearly determined as in the case of one–sided null hypotheses.

A possibility to avoid the problem of constructing a generalized extreme region is to use the same generalized test variable for one–sided and point null hypotheses and define the generalized p–value for the point null hypothesis in the usual way by

$$(9) \quad \begin{aligned} p(t_{obs}) &= 2 \cdot \min\{\Pr(T(Y, y, \xi) > t_{obs}), \Pr(T(Y, y, \xi) < t_{obs})\} \\ &= 2 \cdot \min\{\Pr(Y \in C_y(\xi)), 1 - \Pr(Y \in C_y(\xi))\} \end{aligned}$$

This proceeding also guarantees the p–unbiasedness of the resulting generalized fixed–level test. Moreover, there cannot be problem–immanent reasons against the incidentally assumed symmetry of the generalized extreme region. Nevertheless by using (9) for point null hypotheses it is no longer possible to construct extreme regions of maximal length or other optimality properties.

4 Linear hypotheses

Consider testing problems on linear hypotheses of the form

$$(10)\quad \begin{array}{llcll} H_0^I: & d^T\xi = c & \text{vs.} & H_1^I: & d^T\xi \neq c \quad , \\ H_0^{II}: & d^T\xi \leq c & \text{vs.} & H_1^{II}: & d^T\xi > c \quad , \\ H_0^{III}: & d^T\xi \geq c & \text{vs.} & H_1^{III}: & d^T\xi < c \quad , \end{array}$$

where $d \in \mathbb{R}^n$, $c \in \mathbb{R}$ and n is the dimension of the parameter space Θ.

In the case of testing linear hypotheses, the classification of the parameter vector ξ into the parameter of interest ϑ and the vector of nuisance parameters δ has to be modified. In general, all parameters now are of interest, but all parameters also function as nuisance parameters.

So we transform the hypotheses (10), leaving an arbitrary single parameter on the left side of the special null hypothesis:

$$(11)\quad \begin{array}{llcll} H_0^I: & \xi_i = \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) & \text{vs.} & H_1^I: & \xi_i \neq \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) \quad , \\ H_0^{II}: & \xi_i \leq \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) & \text{vs.} & H_1^{II}: & \xi_i > \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) \quad , \\ H_0^{III}: & \xi_i \geq \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) & \text{vs.} & H_1^{III}: & \xi_i < \frac{1}{d_i}\left(c - \sum\limits_{j\neq i} d_j\xi_j\right) \quad , \end{array}$$

Now by definition ξ_i takes the role of the parameter of interest (ϑ) and all other ξ_j ($j \neq i$), collected in the vector

$$\xi_\delta := (\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_n)^T \ ,$$

function as nuisance parameters.

It will be necessary to modify the definition of a generalized test variable, because property 2 in the case of testing linear hypotheses will prove to be too restrictive. So we come to an adjustment of definition 2.1:

Definition 4.1
A random variable of the form $T = T(Y, y, \xi)$ is said to be a *generalized test variable* if it has the following three properties:

1. $t_{obs} = T(y, y, \xi)$ does not depend on unknown parameters.
2. When ξ_i is specified, and under the assumption of H_0^I (according to (11)), the random variable T has a probability distribution that is independent of the vector of nuisance parameters ξ_δ.
3. For fixed y and ξ_δ, $\Pr(T \leq t|\xi_i)$ is a monotonic function of ξ_i for any given t.

The other definitions in section 2, related to the new definition of a generalized test variable, are no further affected and can be kept in the original form.

Without loss of generality let a generalized test variable be defined as stochastically increasing rather than stochastically monotonic in the parameter of interest. In the case of a stochastically decreasing random variable T it is either possible to invert the inequalities in (12) or, for example, to cross over to the transformation $T^* := 1/T$ which then once again is stochastically increasing in the parameter of interest.

So, the generalized p–values for the three testing problems (10) resp. (11) are given for

$$\begin{aligned} (12) \quad & H_0^I : && p(t_{obs}) = 2 \cdot \min\Big(\Pr(T(Y, y, \xi) \geq t_{obs}|H_0^I)\ ,\ \ \Pr(T(Y, y, \xi) \leq t_{obs}|H_0^I)\Big) \\ & H_0^{II} : && p(t_{obs}) = \Pr(T(Y, y, \xi) \geq t_{obs}|d^T\xi = c) \\ & H_0^{III} : && p(t_{obs}) = \Pr(T(Y, y, \xi) \leq t_{obs}|d^T\xi = c)\ . \end{aligned}$$

Calculating $p(t_{obs})$ under the assumption $d^T\xi = c$ is equivalent to determining the special supremum over H_0. Because of the monotonic property of T, the supremum in all cases is placed on the boundary.

5 Mixed linear models

Consider mixed linear models of the form

$$(13)\quad z \;\sim\; \left(X\omega, \sum_{i=1}^{m} \sigma_i^2 U_i\right) \qquad \text{, i.e. } \mathrm{E}\,(z) = X\omega \text{ , } \mathrm{Cov}\,(z) = \sum_{i=1}^{m} \sigma_i^2 U_i \text{ ,}$$

with

$$X\omega = \sum_{i=1}^{q} X_i\omega_i = 1_n\mu + \sum_{i=2}^{q} X_i\omega_i \quad \text{and} \quad U_m = I_n \text{ .}$$

If we cross over to a reduced model that is invariant with respect to mean value transformations, we get

$$y \;=\; \mathrm{Proj}_{\,R(X)^\perp} z = Mz \quad \text{with} \quad M = I_n - XX^+ \text{ ,}$$

where $R(X)$ is the range of the matrix X and X^+ is the Moore–Penrose inverse of X. So, y is the projection of z onto the complement of $R(X)$ and it follows that

$$(14)\quad y \;\sim\; \left(0, \sum_{i=1}^{m} \sigma_i^2 V_i\right) \quad \text{with} \quad V_i = MU_iM \text{ .}$$

In ANOVA–models $V_1, \ldots, V_m$ are linearly independent and there always exists a basis of pairwise orthogonal projectors $P_1, \ldots, P_m$ which span the same vector space as $V_1, \ldots, V_m$. So, the basis transformation matrix $\Phi = (\varphi_{ij})_{i,j=1,\ldots,m}$ is determined by

$$(15)\quad V_i \;=\; \sum_{i=1}^{m} \varphi_{ij} P_j \text{ ,} \quad i = 1, \ldots, m \text{ .}$$

The sum of squares S_i and mean squares M_i are given by

$$(16)\quad \begin{aligned} S_i \;&=\; z^T P_i z \qquad\quad , \; i = 1, \ldots, m \text{ ,} \\ M_i \;&=\; \frac{1}{\mathrm{tr}\, P_i} z^T P_i z \quad , \; i = 1, \ldots, m \text{ .} \end{aligned}$$

Under the assumption of normality of the random vector z is follows, that the mean squares M_i are stochastically independent with expectation

$$(17)\quad \mathrm{E}\,(M_i) \;=\; \sum_{\nu=1}^{m} \sigma_\nu^2 \varphi_{\nu i} \text{ ,} \quad i = 1, \ldots, m \text{ ,}$$

and the following terms have central χ^2–distributions:

$$(18)\quad \operatorname{tr} P_i \cdot \frac{M_i}{\mathrm{E}\,(M_i)} \quad\sim\quad \chi^2_{\operatorname{tr} P_i}\ .$$

For some special null hypotheses — if two mean squares have the same expectation under H_0 — (18) can be used to construct exact F–tests. In general a construction of exact F–tests is impossible.

For more detailed information about balanced mixed linear models see Hartung *et al.* (1997) or Khuri and Sinha (1998) for the unbalanced case.

For the problem of testing an arbitrary linear hypothesis of variance components (cf. (11)) consider the following random variable

$$(19)\quad T(Y,y,\sigma^2) \quad=\quad \frac{\sum\limits_{l\in L} \beta_l \cdot \mathrm{E}\,(M_l)\dfrac{s_l}{S_l} + \beta_0 c}{\alpha_0 A \dfrac{s_i}{S_i} + \sum\limits_{k\in K} \alpha_k \cdot \mathrm{E}\,(M_k)\dfrac{s_k}{S_k}}\ ,$$

with $\sigma^2 = (\sigma_1^2, \ldots, \sigma_m^2)^T$, s_i the observed value of S_i, $K, L \subseteq \{1, \ldots, i-1, i+1, \ldots, m\}$, constants $\alpha_k, \beta_l \in \mathbb{R}$ and

$$(20)\quad A \quad=\quad \mathrm{E}\,(M_i) - \sigma_i^2 \varphi_{ii} + \varphi_{ii}\left[\frac{1}{d_i}\left(c - \sum_{j\neq i} d_j \sigma_j^2\right)\right]\ ,$$

so that

$$(21)\quad \alpha_0 A + \sum_{k\in K} \alpha_k \cdot \mathrm{E}\,(M_k) \quad=\quad \sum_{l\in L} \beta_l \cdot \mathrm{E}\,(M_l) + \beta_0 c\ ,$$

and all added terms shall be nonnegative:

$$(22)\quad \begin{array}{llll} \alpha_k\,\mathrm{E}\,(M_k) & \geq 0 \quad \forall \quad k \in K & , & \alpha_0 A \geq 0\ , \\ \beta_l\,\mathrm{E}\,(M_l) & \geq 0 \quad \forall \quad l \in L & , & \beta_0 c \geq 0\ . \end{array}$$

Theorem 5.1
The random variable $T(Y, y, \sigma^2)$ from (19) with assumptions (21) and (22) possesses the three properties of a generalized test variable according to definition 4.1.

Proof:

1. The observed value of T

$$t_{obs} \;=\; T(y, y, \sigma^2) \;\overset{(19)}{=}\; \frac{\sum\limits_{l \in L} \beta_l \cdot \mathrm{E}\,(M_l) + \beta_0 c}{\alpha_0 A + \sum\limits_{k \in K} \alpha_k \cdot \mathrm{E}\,(M_k)} \;\overset{(21)}{=}\; 1$$

is constant and therefore especially independent of any parameters.

2. Since α_k, β_l and s_i are constant and due to (18)

$$\sum_{k \in K} \alpha_k \cdot \mathrm{E}\,(M_k) \frac{s_k}{S_k} \qquad \text{and} \qquad \sum_{l \in L} \beta_l \cdot \mathrm{E}\,(M_l) \frac{s_l}{S_l}$$

are linear combinations of independent $1/\chi^2$–expressions, free of any unknown parameter. β_0 and c are constant. Finally, for the left term in the denominator of T in (19) we get

$$(23) \quad \alpha_0 A \frac{s_i}{S_i} \;=\; \alpha_0 s_i \frac{A}{\mathrm{E}\,(M_i)} \frac{\mathrm{E}\,(M_i)}{S_i} \;\overset{H_0^I}{=}\; \alpha_0 s_i \frac{\mathrm{E}\,(M_i)}{S_i}\;,$$

also a $1/\chi^2$–expression, which at least under the assumption of H_0^I is free of nuisance parameters.

3. By construction the parameter of interest σ_i^2 (the former ξ_i in section 4) in T only appears in $\mathrm{E}\,(M_i)$ in the denominator of (23), which again only appears in the denominator of T in (19). With respect to the vector of variance components σ^2 we have

$$T(Y, y, \sigma^2) \;\propto\; \frac{q_1}{q_2 \dfrac{A}{\mathrm{E}\,(M_i)} + q_3}\;.$$

Because of (22) it follows that $q_1, q_2, q_3 \in \mathrm{I\!R}_0^+$, and for that reason T is stochastically increasing in σ_i^2.

With 1., 2. and 3. T indeed is a generalized test variable in the sense of definition 4.1. □

The question that occurs is how to get the constants α_k and β_l. This can be done by an iterative proceeding:

Construction principle for generating generalized test variables for testing linear hypotheses in balanced mixed linear models.

1. Formulate and transform the linear hypothesis of interest, so that a single parameter σ_i^2 is isolated on one side of the hypothesis as it is done in (11).

2. For generating T iteratively start with the expression α_0/A which yields all variance components except for σ_i^2 and A is given by (20).

3. The aim now is to set t_{obs} equal to 1. Therefore we have to add $1/\chi^2$–expressions in the numerator and denominator of T.

4. To find an admissible set of α_k and β_l, the easiest way is to eliminate the variance components according to their appearance in the model (14) from left to right. For example: If the leftmost variance component in the numerator (that is not yet egalized in the denominator) is to be egalized, take that $1/\chi^2$–expression according to (18) with an expectation (17) whose leftmost variance component is the one to be egalized.

5. The last step is to set β_0 which is clearly determined as $\beta_0 := \alpha_0\varphi_{ii}/d_i$ by the former proceeding.

6 An illustrative example

The balanced 2–way nested classification model with random effects is given by

$$
\begin{aligned}
y_{ijk} &= \mu + a_i + b_{ij} + e_{ijk} \qquad i = 1,\ldots,r \quad j = 1,\ldots,s \quad k = 1,\ldots,t \\
& \text{with} \quad E[y_{ijk}] = \mu \ , \ a_i \sim (0,\sigma_a^2) \ , \ b_{ij} \sim (0,\sigma_b^2) \ , \ e_{ijk} \sim (0,\sigma_e^2) \\
& a_i, b_{ij} \text{ and } e_{ijk} \text{ stochastically independent .}
\end{aligned} \tag{24}
$$

Under the assumption of normally distributed random effects a_i, b_{ij} and e_{ijk}, we have the following distribution statements for the sums of squares (cf. (18)):

$$
\begin{aligned}
S_a &\sim (st\sigma_a^2 + t\sigma_b^2 + \sigma_e^2) \cdot \chi^2_{r-1} \\
S_b &\sim (t\sigma_b^2 + \sigma_e^2) \cdot \chi^2_{r(s-1)} \\
S_e &\sim \sigma_e^2 \cdot \chi^2_{rs(t-1)} \ .
\end{aligned}
$$

A generalized test variable for arbitrary linear hypotheses (under the restriction of $d_1 \neq 0$) is

$$
\begin{aligned}
(25)\quad T(Y,y,\sigma^2) &= \frac{\left(1-\frac{s\cdot d_2}{d_1}\right)(t\sigma_b^2+\sigma_e^2)\frac{s_b}{S_b}+\frac{s\cdot d_2}{d_1}\sigma_e^2\frac{s_e}{S_e}+\frac{st}{d_1}c}{\left(st\frac{1}{d_1}(c-d_2\sigma_b^2-d_3\sigma_e^2)+t\sigma_b^2+\sigma_e^2\right)\frac{s_a}{S_a}+\frac{st\cdot d_3}{d_1}\sigma_e^2\frac{s_e}{S_e}} \\
&= \frac{\left(1-\frac{s\cdot d_2}{d_1}\right)\frac{s_b}{\chi^2_{r(s-1)}}+\frac{s\cdot d_2}{d_1}\frac{s_e}{\chi^2_{rs(t-1)}}+\frac{st}{d_1}c}{\frac{\left(st\frac{1}{d_1}(c-d_2\sigma_b^2-d_3\sigma_e^2)+t\sigma_b^2+\sigma_e^2\right)}{st\sigma_a^2+t\sigma_b^2+\sigma_e^2}\frac{s_a}{\chi^2_{r-1}}+\frac{st\cdot d_3}{d_1}\frac{s_e}{\chi^2_{rs(t-1)}}} \\
&\overset{H_0^I}{=} \frac{\left(1-\frac{s\cdot d_2}{d_1}\right)\frac{s_b}{\chi^2_{r(s-1)}}+\frac{s\cdot d_2}{d_1}\frac{s_e}{\chi^2_{rs(t-1)}}+\frac{st}{d_1}c}{\frac{s_a}{\chi^2_{r-1}}+\frac{st\cdot d_3}{d_1}\frac{s_e}{\chi^2_{rs(t-1)}}}
\end{aligned}
$$

Because no assumptions about $d \in IR^3$ except for $d_1 \neq 0$ have been made, negative terms can occur in T. Should this be the case, these negative terms have to be added to the numerator and the denominator of T, which neither influences $t_{obs} = 1$ nor leads to a dependence of the generalized test variable on nuisance parameters.

Suppose the hypotheses of interest are for example

$$(26)\qquad H_0^I : \sigma_a^2 = \sigma_b^2 \quad \text{vs.} \quad H_1^I : \sigma_a^2 \neq \sigma_b^2$$

and

$$(27)\qquad H_0^{II} : \sigma_a^2 \leq \sigma_b^2 \quad \text{vs.} \quad H_1^{II} : \sigma_a^2 > \sigma_b^2 \; .$$

This means $d = (1,-1,0)^T$ and $c = 0$, and the generalized test variable T is

$$T(Y,y,\sigma^2) = \frac{(1+s)\frac{s_b}{\chi^2_{r(s-1)}} - s\frac{s_e}{\chi^2_{rs(t-1)}}}{\frac{st\sigma_b^2+t\sigma_b^2+\sigma_e^2}{st\sigma_a^2+t\sigma_b^2+\sigma_e^2}\cdot\frac{s_a}{\chi^2_{r-1}}} \; .$$

Since it is not obvious whether T is a monotone function in σ_a^2, the function

$$T(Y,y,\sigma^2) = \frac{(1+s)\frac{s_b}{\chi^2_{r(s-1)}}}{\frac{st\sigma_b^2+t\sigma_b^2+\sigma_e^2}{st\sigma_a^2+t\sigma_b^2+\sigma_e^2}\cdot\frac{s_a}{\chi^2_{r-1}} + s\frac{s_e}{\chi^2_{rs(t-1)}}}$$

is regarded and then T obviously is a stochastically increasing function in σ_a^2. This necessary transformation is not a consequence of the general construction principle (cf. section 5), but caused by using (25) with an arbitrary $d \in \mathbb{R}^3$. In the case of starting with a certain hypothesis and a fixed $d \in \mathbb{R}^3$ the problem of negative term in T does not occur.

Provided H_0^I is true, then

$$T(Y, y, \sigma^2) = \frac{(1+s)\frac{s_b}{\chi^2_{r(s-1)}}}{\frac{s_a}{\chi^2_{r-1}} + s\frac{s_e}{\chi^2_{rs(t-1)}}} \; .$$

The generalized fixed–level test is given by the rule

$$\text{Reject} \begin{array}{c} H_0^I \\ H_0^{II} \end{array} \text{ at the nominal level } \alpha, \text{ if } \left\{ \begin{array}{c} 2 \cdot \min\{\Pr(T>1), \Pr(T<1)\} \\ \Pr(T>1) \end{array} \right\} < \alpha \; .$$

The probabilities $P(T > 1)$ and $P(T < 1)$ are determined by simulation.

For various constellations of the parameters r, s, t and $\sigma_a^2 = \sigma_b^2$, with $\sigma_e^2 = 1$ and the nominal significance level of $\alpha = 0.05$ the following generalized p–values resulted from simulation studies (1000 runs in each simulation):

				$p(t_{obs})$						$p(t_{obs})$	
r	s	t	$\sigma_a^2 = \sigma_b^2$	H_0^I	H_0^{II}	r	s	t	$\sigma_a^2 = \sigma_b^2$	H_0^I	H_0^{II}
3	4	2	0.2	0.040	0.059	6	2	2	0.2	0.040	0.058
3	4	2	1	0.055	0.056	6	2	2	1	0.044	0.052
3	4	2	5	0.062	0.054	6	2	2	5	0.050	0.053
3	4	2	10	0.061	0.052	6	2	2	10	0.050	0.053
3	4	8	0.2	0.041	0.068	2	5	3	0.2	0.036	0.067
3	4	8	1	0.042	0.057	2	5	3	1	0.038	0.053
3	4	8	5	0.043	0.054	2	5	3	5	0.046	0.044
3	4	8	10	0.042	0.054	2	5	3	10	0.047	0.044

The simulations show, that even with small sample sizes the approximative tests have an estimated significance level near to the nominal one. With rising r the approximation becomes even better. In general the one–sided test for problem (27) tends to be more conservative than the two–sided test (26).

For $s = 3$, $t = 2$, $\sigma_b = 2$, $\sigma_e = 1$ and a nominal significance level of $\alpha = 0.05$ the following data–based power–functions $\pi(t_{obs}, \sigma_a^2)$ are computed by simulation in dependence on r:

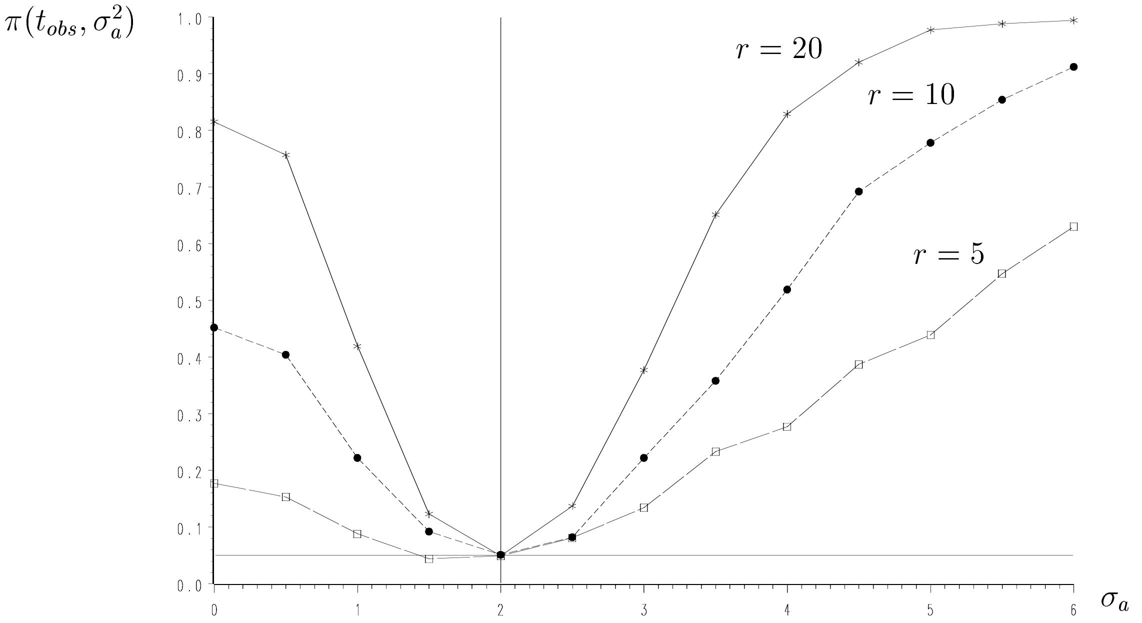

Figure 1: Estimated power of the two–sided test (cf. 26) as a function of σ_a

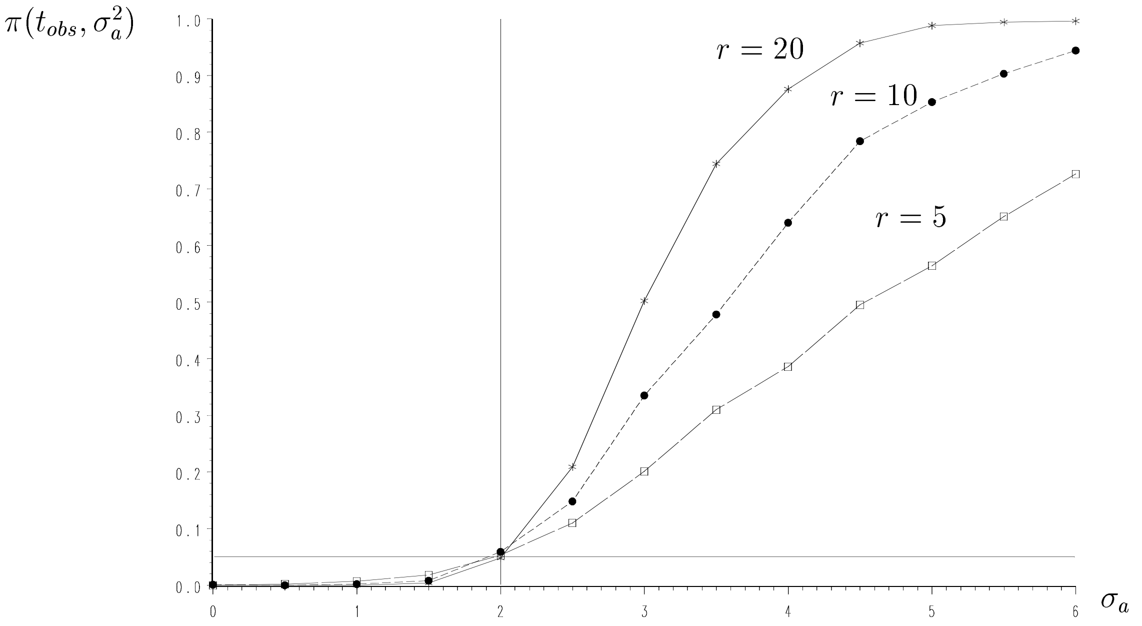

Figure 2: Estimated power of the one–sided test (cf. 27) as a function of σ_a

For the estimated data–based power–functions $\pi(t_{obs}, \sigma_b^2)$ in dependency of σ_b^2, with $\sigma_a = 2$ and under the same parameter constellation we get the following result:

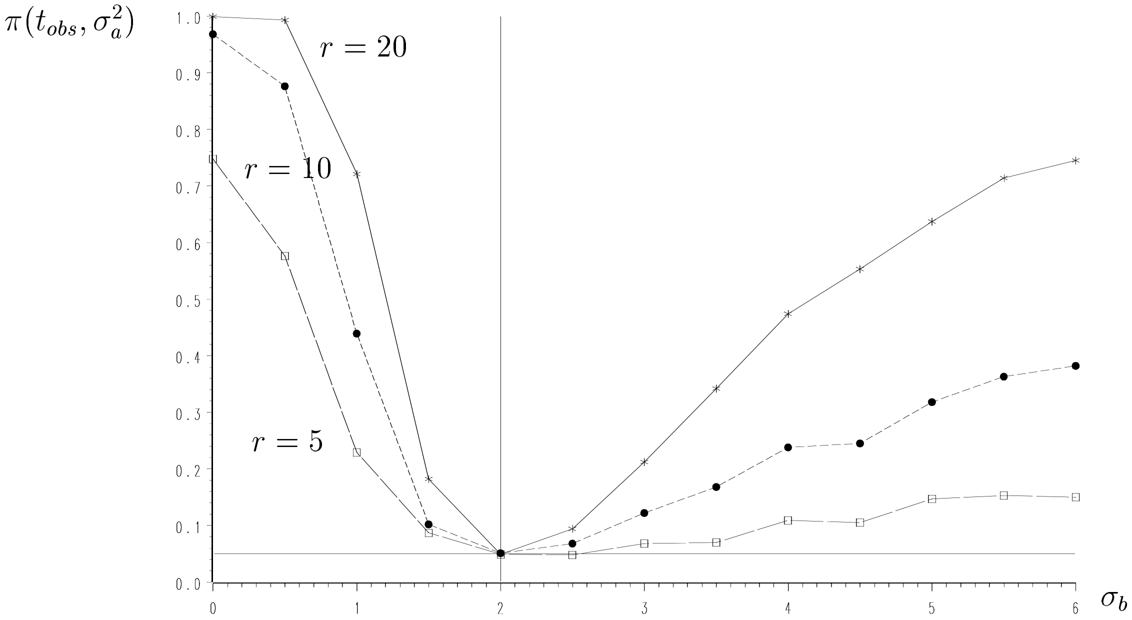

Figure 3: Estimated power of the two–sided test (cf. 26) as a function of σ_b

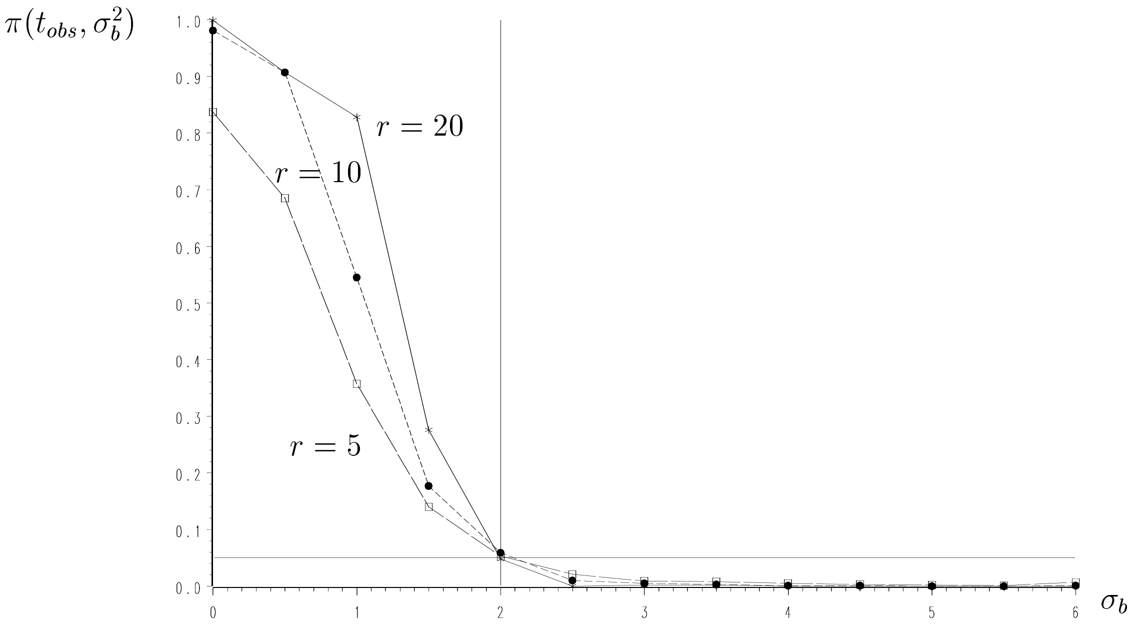

Figure 4: Estimated power of the one–sided test (cf. 27) as a function of σ_b

References

Hartung, J., Elpelt, B., and Voet, B. (1997). *Modellkatalog Varianzanalyse*. R. Oldenbourg, Muenchen.

Khuri, A. I. and Sinha, B. K. (1998). *Statistical Tests for Mixed Linear Models*. John Wiley & Sons, New York.

Khuri, I. A. (1990). Exact tests for random models with unequal cell frequencies in the last stage. *Journal of Statistical Planning and Inference*, **24**, 177–193.

Satterthwaite, F. E. (1946). An approximative distribution of estimates of variance components. *Biomet. Bulletin*, **2**, 110–114.

Thursby, J. G. (1989). A comparison of several exact and approximative tests for structural shift under heteroscedascity. *Journal of Econometrics*, **53**, 363–386.

Tsui, K. and Weerahandi, S. (1989). Generalized p–values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association*, **84**, 602–607.

Weerahandi, S. (1991). Testing variance components in mixed models with generalized p-values. *Journal of the American Statistical Association*, **86**, 151–153.

Weerahandi, S. (1995). *Exact Statistical Methods for Data Analysis*. Springer, New York.

Zhou, L. and Mathew, T. (1994). Some tests for variance components using generalized p–values. *Technometrics*, **36**, 394–402.