The work was supported by the Minister of Agriculture of the Czech Republic.

### Type-I-Risk of Additivity Tests

The actual type-I-risk of all the five tests of additivity were verified for the mixed ANOVA model by simulation. Only the most common normal type-I-risk 5% was assumed.

For verifying the type-I-risk, the number of levels of the fixed factor was assumed $a = 3, 4, 5, 10$, numbers of levels of the random factor $b$ was chosen between 4 and 50 (by 2 between 4 and 20, by 5 between 20 and 50), the variance of the random factor $\sigma^2 = 2, 5, 10$ and the variance of the random error $\sigma^2 = 1$.

In one step of the simulation a set of data was generated based on the model without interaction. Then the test of no interaction was performed. The percentage of significant test after several steps is assumed to be the actual level of the test.

The 10 000 replications were repeated 10 times and the standard error of the estimation of the mean actual level was computed based on these 10 replications. Then the one-sided one-sample $t$-test of the hypothesis ‘the actual $p$-value is equal to or less than 0.05’ was performed (on the 5% level) for each of the two tests.

The results of these tests are summarized in Table 1.

<table>
<thead>
<tr>
<th>Test</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tukey test</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>Mandel test</td>
<td>0.012</td>
<td>0.055</td>
</tr>
<tr>
<td>Graybill test</td>
<td>0.017</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Table 1: Number and percentage of the simulated cases actual test level is equal to or lower than nominal level 5%.

For the Tukey and Mandel tests in the vast majority (> 95%) of cases the actual level is not significantly above the 0.05 level. In less than 4% of the cases the $t$-test is higher than the nominal 0.05. For the other tests the nominal level is higher than 0.05 in slightly more cases. However, this may also be false positives caused by multiple testing.

In the ANOVA models with both effects fixed there is an important assumption about summing of both effects to zero, $i.e.$$\sum_{i=1}^{a} \gamma_i = 0$. In the case of random model that is not valid. It is assumed that the expected value of random term $\gamma_i(b_j)$ equals zero, but in one particular case the sum is not zero (almost surely). It can cause inaccuracy of the results. However, for high numbers of levels of the random factors $b$, the sum converges to zero (law of large numbers) and this problem disappears.

For %5 type-I-risk all these tests hold the level of type-I-risk and therefore the tests designed for fixed models can be used for the mixed models as well.

### A Modification of the Tukey Test

As mentioned above the Tukey, Mandel, Johnson – Graybill, LII and Tussel test hold the level of type-I-risk even when one factor is considered as random. In this part the power of these tests is compared by means of simulation. While Tukey test has relatively good power when the interaction is a product of the main effects, $i.e.$, when $\gamma_{ij} = a \cdot b \cdot \text{interaction type A}$ its very poor.

To increase the power of Tukey test the test propose the following modification. In Tukey test a model $y_{i.} = \beta + a_i + b_i + (a_i \cdot b_i)$ is tested against a submodel $y_{i.} = \beta + a_i + b_i + \epsilon_i$. The estimators of row effects $a_i = \bar{y}_{i.} - \bar{y}$ and column effects $b_i = \bar{y}_{i.} - \bar{y}$ are calculated in the Tukey test in both models although the dependency of $y_{ij}$ on these parameters is not linear for the full model.

The main idea behind the presented modification is that a full model $y_{i.} = \beta + a_i + b_i + (a_i \cdot b_i) + \epsilon_i$ is fitted by a modified test against a submodel $y_{i.} = \beta + a_i + b_i + \epsilon_i$ by a likelihood ratio test. The estimates of row and column effect therefore differ for each model. See the simulation results for modified test.

If the sample size is small bootstraping without replacement or sampling from fitted model is used to control type-I-risk level.

Acknowledgement. The work was supported by the Minister of Agriculture of the Czech Republic. MEB/002760/03.

References.


