Nauka Przyroda Technologie

2014
Tom 8
Zeszyt 3
#43

ISSN 1897-7820 http://www.npt.up-poznan.net
Dział: Rolnictwo

MICHAŁ SIATKOWSKI[1], KATARZYNA GÓRCZAK[2], IDZI SIATKOWSKI[2]

[1]Institute of Agricultural Engineering
Poznań University of Life Sciences
[2]Department of Mathematical and Statistical Methods
Poznań University of Life Sciences

ANALYSIS OF A SINGLE EXPERIMENT IN AFFINE RESOLVABLE BLOCK DESIGN IN MULTI ENVIRONMENTS

ANALIZA POJEDYNCZEGO DOŚWIADCZENIA
W AFINICZNIE ROZKŁADALNYCH UKŁADACH BLOKOWYCH
W WIELU ŚRODOWISKACH

This paper is dedicated to Professor Tadeusz Caliński on his 85th birthday

Summary. This paper discusses the package ASE for analysing a single experiment carried out in affine resolvable proper block design in multi environments. The proposed script is written in R code and is based on the theory presented in CALIŃSKI and KAGEYAMA (2008).

Key words: affine resolvable block design, stratum, randomization model, combined analysis, GB design, R software

Introduction

Multi environment trials theory allows to perform statistical analysis for the set of objects (species, genes, genotypes) occurring in multiple environments (locations). This theory was introduced by Professor Tadeusz Caliński (CALIŃSKI and KAGEYAMA 2000, 2003, 2004, 2008). Computing package ASE (Analysis of a Single Experiment) was introduced in this paper and created in R software (R DEVELOPMENT CORE TEAM 2010), which allows to conduct statistical analysis of a single trial carried out in affine resolvable proper block design. This package is available from the corresponding author.

Methods

Let us consider a variety trials carried out in affine resolvable block design with v varieties (genotypes, species, genes, types, objects) allocated in b blocks, each of k plots, grouped into r superblocks in such a way that each superblock, composed of s blocks, contains all v varieties, each of them exactly once, and that every pair of blocks from different superblocks has the same number, k/s of varieties in common (making the design affine resolvable). Let us suppose that the randomizations of superblocks, of blocks within the superblocks and of plots within the blocks have been implemented in the trial according to the procedure described in Section 5.2.1 of CALIŃSKI and KAGEYAMA (2000).

The package ASE (i.e. algorithm and script in R) is based on the theory presented in the paper of CALIŃSKI and KAGEYAMA (2008). Labels presented in result file are the same as the ones in CALIŃSKI and KAGEYAMA (2008). The model (in matrix notation) is as follows:

$$\boldsymbol{y} = \boldsymbol{\Delta}'\boldsymbol{\tau} + \boldsymbol{G}'\boldsymbol{\alpha} + \boldsymbol{D}'\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is $(n \times 1)$ vector of data concerning variable traits (e.g. yields) observed on $n = rv$ units of the experiment, $\boldsymbol{\Delta}' = \boldsymbol{1}_r \otimes \boldsymbol{I}_v$, $\boldsymbol{G}' = \boldsymbol{I}_r \otimes \boldsymbol{1}_v$, and $\boldsymbol{D}' = \mathrm{diag}[\boldsymbol{D}'_1 : \boldsymbol{D}'_2 : \ldots : \boldsymbol{D}'_r]$ with $\boldsymbol{D}'_h = \boldsymbol{N}_h$, the incidence matrix and $h = 1, 2, \ldots, r$. The symbols $\boldsymbol{1}_a$ and $\boldsymbol{I}_a$ denote the $(a \times 1)$ vector of ones and the $(a \times a)$ identity matrix, respectively. Further, $\boldsymbol{\tau}$ represents the variety (treatment) parameters fixed effects, $\boldsymbol{\alpha}$ the superblock random effects, $\boldsymbol{\beta}$ the block random effects, whereas the $(n \times 1)$ random vectors $\boldsymbol{\eta}$ and $\boldsymbol{e}$ stand for the experimental unit and technical error, respectively. The covariance matrix of $\boldsymbol{y}$ is of the form:

$$\mathrm{Cov}(\boldsymbol{y}) = \boldsymbol{\phi}_1\sigma_1^2 + \boldsymbol{\phi}_2\sigma_2^2 + \boldsymbol{\phi}_3\sigma_3^2 + \boldsymbol{\phi}_4\sigma_4^2$$

where the matrices $\boldsymbol{\phi}_1 = \boldsymbol{I}_n - k^{-1}\boldsymbol{D}'\boldsymbol{D}$, $\boldsymbol{\phi}_2 = k^{-1}\boldsymbol{D}'\boldsymbol{D} - v^{-1}\boldsymbol{G}'\boldsymbol{G}$, $\boldsymbol{\phi}_3 = v^{-1}\boldsymbol{G}'\boldsymbol{G} - n^{-1}\boldsymbol{1}_n\boldsymbol{1}'_n$ and $\boldsymbol{\phi}_4 = n^{-1}\boldsymbol{1}_n\boldsymbol{1}'_n$ are symmetric, idempotent and pairwise orthogonal, the scalars σ_1^2, σ_2^2, σ_3^2 and σ_4^2 represent the relevant unknown stratum variances, i.e. σ_1^2 – the intra-block stratum variance (sigma1^2 in ASE), σ_2^2 – the inter-block-intra-superblock stratum variance (sigma2^2 in ASE), σ_3^2 – the inter-superblock stratum variance (sigma3^2 in ASE), and σ_4^2 – the total area stratum variance. Let $\boldsymbol{N} = [\boldsymbol{N}_1 : \boldsymbol{N}_2 : \ldots : \boldsymbol{N}_r]$, $\boldsymbol{C}_1 = r\boldsymbol{I}_n - k^{-1}\boldsymbol{N}\boldsymbol{N}'$, $\boldsymbol{Q}_1 = \boldsymbol{\Delta}\boldsymbol{\phi}_1\boldsymbol{y}$, and $\boldsymbol{\psi}_1 = \boldsymbol{\phi}_1 - \boldsymbol{\phi}_1\boldsymbol{\Delta}'\boldsymbol{C}_1^-\boldsymbol{\Delta}\boldsymbol{\phi}_1$. Then the intra-block analysis of variance can be obtained as in the Table 1. Now, let $\boldsymbol{C}_2 = k^{-1}\boldsymbol{N}\boldsymbol{N}' - v^{-1}r\boldsymbol{1}_v\boldsymbol{1}'_v$, $\boldsymbol{Q}_2 = \boldsymbol{\Delta}\boldsymbol{\phi}_2\boldsymbol{y}$, and $\boldsymbol{\psi}_2 = \boldsymbol{\phi}_2 - \boldsymbol{\phi}_2\boldsymbol{\Delta}'\boldsymbol{C}_2^-\boldsymbol{\Delta}\boldsymbol{\phi}_2$. Then, the inter-block-intra-superblock analysis of variance can be presented as in the Table 2.

Further, the combined analysis of variance can be in the form as in Table 3, where d, SS_0 and SS_1 are described in formulas (4.14) and (4.15) of CALIŃSKI and KAGEYAMA (2008).

It is known, that the intra-block best linear unbiased estimator (BLUE) and the inter-block-intra-superblock BLUE of the contrast can be written as:

$$\widehat{(\boldsymbol{c}'\boldsymbol{\tau})}_1 = \boldsymbol{c}'\boldsymbol{C}_1^-\boldsymbol{Q}_1 \text{ and } \widehat{(\boldsymbol{c}'\boldsymbol{\tau})}_2 = \boldsymbol{c}'\boldsymbol{C}_2^-\boldsymbol{Q}_2 \qquad (1)$$

respectively.

Table 1. Intra-block analysis of variance
Tabela 1. Wewnątrzblokowa analiza wariancji

Source of variation Źródło zmienności	Degrees of freedom Stopnie swobody	Sum of squares Sumy kwadratów	Name in ASE Nazwa w ASE
Total – Ogółem	$n - b$	$y'\phi_1 y$	SSintraC
Treatments – Obiekty	$rank(\boldsymbol{C}_1)$	$\phi'_1 C_1^- \phi_1$	SSintraT
Residuals – Reszty	$n - b - v + 1$	$y'\psi_1 y$	SSintraRes

Table 2. Inter-block-intra-superblock analysis of variance
Tabela 2. Międzyblokowo-wewnątrzsuperblokowa analiza wariancji

Source of variation Źródło zmienności	Degrees of freedom Stopnie swobody	Sum of squares Sumy kwadratów	Name in ASE Nazwa w ASE
Total – Ogółem	$b - r$	$y'\phi_2 y$	SSinterC
Treatments – Obiekty	$rank(\boldsymbol{C}_2)$	$\phi'_2 C_2^- \phi_2$	SSinterT
Residuals – Reszty	$n - r - rank(\boldsymbol{C}_2)$	$y'\psi_2 y$	SSinterRes

Table 3. Inter-block-intra-superblock of combined analysis of variance
Tabela 3. Międzyblokowo-wewnątrzsuperblokowa kombinowana analiza wariancji

Source of variation Źródło zmienności	Degrees of freedom Stopnie swobody	Sum of squares Sumy kwadratów	Name in ASE Nazwa w ASE
Total – Ogółem	$n - b - v + d + 1$	$SS_0 + SS_1 + y'\psi_1 y$	SScombC
Treatments – Obiekty	d	$SS_0 + SS_1$	SScombT
Residuals – Reszty	$n - b - v + 1$	$y'\psi_1 y$	SScombRes

Package description

The ASE package has been created in R version 2.15.2 (R DEVELOPMENT CORE TEAM 2010). First, text files “data-met.txt” and “location-names.txt” are created containing input data formatted accordingly to the structure described below and names of experimental stations (locations), respectively. Then, the files are placed in the folder “D:\MET-R”. Next, the ASE script is open in R and running. The results are inserted into a text file “results.txt” and placed in the folder “D:\MET-R”.

Data structure

Data that describes an experiment that occurs in multiple locations, supposed in set of blocks and superblocks, carried out in affine resolvable design (CALIŃSKI and KAGEYAMA 2000) are presented in a text file which is a data array which columns are named as follows: locations, superblocks, blocks, objects, yields. In the analysed example 15 locations (environments) are presented – data fragment is shown below (example of winter rye, see CALIŃSKI et AL. 2009, Section 5). Each trial is arranged in affine resolvable design with $v = 18$ varieties (objects) allocated in $b = 12$ blocks grouped into $r = 4$ superblocks, each containing $s = 3$ blocks of size $k = 6$. The varieties are assigned to blocks in such a way that every pair of blocks from different superblocks has the same number of varieties, $k/s = 2$, in common. In ASE package it is named: v – objects, b – blocks, r – superblocks, s – blocksinsuperblock, k – plots.

Data fragment:

locations	superblocks	blocks	objects	yields
1	1	1	1	70.0824
1	1	2	2	74.3686
1	1	3	3	79.5294
1	1	4	4	66.6667
1	1	1	5	73.5059
1	1	2	6	72.6667
1	1	3	7	79.4353
1	1	4	8	70.3294
1	1	1	9	73.6784

....

Moreover, in second text file "location-names.txt" the names of the locations (environments) are contained: Maslowice, Przeclaw, ...

Results

As the result of the calculations conducted by the ASE package the text file "results.txt" is being created automatically inside "D:\MET-R" folder and it contains the following information for each location (all descriptions of the results are referenced in CALIŃSKI and KAGEYAMA 2008):

----
1. Maslowice
objects = 18
blocks = 12
superblocks = 4
blocksinsuperblock = 3
plots = 6
units = 72

Incidence matrix N

1 :	1 0 0	0 0 1	1 0 0	0 1 0
2 :	0 0 1	1 0 0	0 1 0	0 1 0
3 :	1 0 0	1 0 0	1 0 0	0 0 1
4 :	0 0 1	1 0 0	0 0 1	0 0 1
5 :	0 0 1	0 1 0	0 0 1	1 0 0

....

$\varepsilon_0 = 1$ and $\varepsilon_1 = (r-1)/r$ (the efficiency factors of design), $\rho_0 = v - 1 - r(s-1)$ and $\rho_1 = r(s-1)$ (of multiplicities, respectively); see CALIŃSKI and KAGEYAMA (2008), p. 3351

$\varepsilon_0 = 1$; $\varepsilon_1 = 0.75$; $\rho_0 = 9$; $\rho_1 = 8$

intra-block analysis of variance (see Table 1 or CALIŃSKI and KAGEYAMA 2008, p. 3352)

	df	SS	MS	F	F0.05	F0.01	F0.001
SSintraC	60	2913.0262					
SSintraT	17	2829.4980	166.4411	85.68	1.87	2.41	3.21
SSintraRes	43	83.5282	1.9425				

sums squares for inter-block-intra-superblock analysis of variance (see Table 2 or CALIŃSKI and KAGEYAMA 2008, p. 3353)

SSinterC	354.2867
SSinterT	354.2867
SSinterRes	0.0000

sigma1^2, sigma2^2, sigma3^2 representing the relevant unknown stratum variances; see CALIŃSKI and KAGEYAMA (2008), p. 3352

sigma1^2 = 1.942517; sigma2^2 = 6.705973; sigma3^2 = 1.343790

c'tau1 and c'tau2 are the best linear unbiased estimator for each analysis; see formula (1) and CALIŃSKI and KAGEYAMA (2008), p. 3352

Values for objects

objects	mean	Q1	c'tau1	var	F	c'tau2
4	71.49	38.8177	10.7543	0.5306	217.9737	7.65593
15	68.28	36.7199	8.9415	0.5306	150.6815	–3.06037
17	67.79	29.8356	8.0497	0.5306	122.1225	1.84119
12	67.67	29.1904	7.6915	0.5306	111.4965	2.02748
14	67.36	24.4584	6.7874	0.5306	86.8238	5.50769

....

Total mean 59.87

F0.05 = 4.07; F0.01 = 7.26; F0.001 = 12.47

comparing with the average of standard varieties; see CALIŃSKI and KAGEYAMA (2008), formula (4.16)
Comparing with the average of standard varieties

objects	c'tau1	var	F	c'tau2
15	9.3314	0.7554	115.26662	–2.2658
17	8.4396	0.7015	101.53965	2.6358
12	8.0814	0.6880	94.92943	2.8221
14	7.1772	0.7150	72.05052	6.3023
8	5.3682	0.7015	41.08216	2.7448

$\varepsilon_{11} = (r - 1)/r$ and $\varepsilon_{21} = 1 - \varepsilon_1$ (the efficiency factors of stratums, respectively); see CALIŃSKI and KAGEYAMA (2008), p. 3353
$\varepsilon_{11} = 0.75$; $\varepsilon_{21} = 0.25$

dzeta is in form presented in CALIŃSKI and KAGEYAMA (2008), formula (4.12)
dzeta(4.12) = 0.02863014

combined analysis; see CALIŃSKI and KAGEYAMA (2008), p. 3353, formula (4.1)
Combined analysis

objects	mean	tau~	tau~-t.	var	F
4	71.49	70.93	11.0587	0.5128	238.4981
15	68.28	68.63	8.7560	0.5128	149.5178
17	67.79	67.87	8.0037	0.5128	124.9292
12	67.67	67.60	7.7313	0.5128	116.5693
14	67.36	66.90	7.0354	0.5128	96.5281

....

combined analysis – comparing with the average of standard varieties; see CALIŃSKI and KAGEYAMA (2008), p. 3355, formula (4.13)
Comparing with the average of standard varieties

objects	c'tau	var	F
15	9.1251	0.7187	115.86430
17	8.3728	0.6781	103.38660
12	8.1003	0.6679	98.23832
14	7.4044	0.6882	79.66298
8	5.4329	0.6781	43.52972

....

w1(4.7), w2(4.8), sigma1^2(N), sigma2^2(N), w1(4.11), w2(4.11) are described in CALIŃSKI and KAGEYAMA (2008) in formulas (4.7), (4.8), (4.9), (4.10) and (4.11)
w1(4.7) = 0.9119457; w2(4.8) = 0.08805432
sigma1^2(N) = 1.942517; sigma2^2(N) = 6.705973
w1(4.11) = 0.9119457; w2(4.11) = 0.08805432

combined analysis of variance (see Table 3 or CALIŃSKI and KAGEYAMA 2008, p. 3355, formula (4.14) and (4.15))

	df	SS	MS	F	F0.05	F0.01	F0.001
SScombC	60	2981.36668					
SScombT	17	2897.83848	170.461087	87.75	1.87	2.41	3.21
SScombRes	43	83.5282	1.9425				

----

2. Przeclaw
....

Conclusions

This paper presents computing package created in R which allows for analysing an experiment carried out in affine resolvable proper block design conducted in the form of a series of experiments. Application performs all calculation automatically along with opening an input data set and saving the file with the results.

This package shows results for the intra-block stratum variance, the inter-block--intra-superblock stratum variance, and the combined analysis.

Acknowledgement

The authors would like to thank the anonymous referees for their useful comments and suggestions.

References

CALIŃSKI T., CZAJKA S., KACZMAREK Z., KRAJEWSKI P., PILARCZYK W., 2009. A mixed model analysis of variance for multi-environment variety trials. Stat. Pap. 50: 735-759.

CALIŃSKI T., KAGEYAMA S., 2000. Block designs: a randomization approach, I: Analysis. Lect. Notes Stat. 150.

CALIŃSKI T., KAGEYAMA S., 2003. Block designs: a randomization approach, II: Design. Lect. Notes Stat. 170.

CALIŃSKI T., KAGEYAMA S., 2004. A unified terminology in block designs: an informative classification. Discuss. Math. – Probab. Stat. 24: 127-145.

CALIŃSKI T., KAGEYAMA S., 2008. On the analysis of experiments in affine resolvable designs. J. Stat. Plan. Infer. 138: 3350-3356.

R DEVELOPMENT CORE TEAM, 2010. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. [http://www.r-project.org].

ANALIZA POJEDYNCZEGO DOŚWIADCZENIA W AFINICZNIE ROZKŁADALNYCH UKŁADACH BLOKOWYCH W WIELU ŚRODOWISKACH

Streszczenie. W pracy przedstawiono pakiet statystyczny ASE do analizy pojedynczego doświadczenia założonego w afinicznie rozkładalnym układzie blokowym w wielu środowiskach. Proponowany skrypt (program oraz procedury) napisano w kodzie R, bazując na teorii przedstawionej w pracy CALIŃSKIEGO i KAGEYAMY (2008).

Słowa kluczowe: układ blokowy afinicznie rozkładalny, warstwa, model randomizowany, analiza kombinowana, układ ogólnie zrównoważony, platforma obliczeniowa R

Corresponding address – Adres do korespondencji:
Idzi Siatkowski, Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu, ul. Wojska Polskiego 28, 60-637 Poznań, Poland, e-mail: idzi@up.poznan.pl

Accepted for publication – Zaakceptowano do opublikowania:
16.06.2014

For citation – Do cytowania:
Siatkowski M., Górczak K., Siatkowski I., 2014. Analysis of a single experiment in affine resolvable block design in multi environments. Nauka Przyr. Technol. 8, 3, #43.