

R.w. Payne

Issue No. 10

The
GENSTAT
Newsletter

EDITORS:

*R.W. PAYNE
ROTHAMSTED EXPERIMENTAL STATION
HARPENDEN
HERTFORDSHIRE
AL5 2JQ*

*M.G. RICHARDSON
NAG CENTRAL OFFICE
MAYFIELD HOUSE
256 BANBURY ROAD
OXFORD
OX2 7DE*

NAG



CONTENTS

	Page
1. Editorial	3
2. Genstat 4.03 Status Report	3
3. Genstat 4.04	7
4. Backing Store Changes in Genstat 4.04	12
5. The 'HELP' Directive	16
6. Diallel Crosses, Partially Balanced Incomplete Block Designs with Triangular Association Schemes and Rectangular Lattices	16
7. A Regression Model for Genotypical Stability	32
8. Non-hierarchical Classification - CLASSIFY (and CLASSF)	40
9. Genstat and PRIME CPL	42
10. Genstat Error Notice S3010	47
11. Notices	
11a. Genstat Conference 1983	52
11b. "An Introduction to Genstat"	53
11c. German Genstat Manual	53
11d. ITSM Announcements	53
12. Genstat Newsletter Order Form	

EDITORIAL

The Editors wish to apologise for the late appearance of this issue of the GENSTAT Newsletter; delayed by the pressure of other work. We hope to return approximately to schedule with issue 11, copy for which should be submitted as soon as possible.

Occasionally, minor subscription queries to NAG Central Office cause a disproportionate amount of work. It would be very helpful if anyone with such a query could quote the reference number on the top right hand corner of the address label (or, better still, enclose the label itself). The reference number has a very simple interpretation: for licenced package sites, it is the licence number; for other subscribers the number to the left of the solidus (/) indicates the final issue number of the subscription, that to the right is the number of copies of each issue ordered.

Finally, we regret that the price of the Newsletter has had to be increased to £5.00 per year. When the price was first set at £2.50, it approximately covered printing and postage costs. This is no longer the case, since costs have risen substantially. At the new price, NAG will still be providing most of the editorial and typing effort free of charge.

GENSTAT 4.03 STATUS REPORT

Since the last Newsletter, Genstat has been converted for the Siemens (BS 2000) by Dipl. Math. R Valder of the University of Dusseldorf. Versions for Harris (Stephen Morris, Chelsea College), Honeywell Multics (Doug Clark, University of Bath), Honeywell Sigma 6 (Gordon West, H M Government Communications Centre, Hanslope) and Perkin-Elmer (Graham Jagger and Tony Field, Life Science Research) have also appeared. An interactive version has been produced for DEC-10's by John Byrne of York University.

Ron Baxter of CSIRO is producing a version for the Vax under Unix. This is likely to be the last conversion of 4.03; when it is completed the final tally will be

- 25 separate conversions completed;
- on at least 18 different machine ranges;
- from 16 distinct manufacturers.

Orders for Genstat over the last six months have averaged 5 new copies per month. By the time this Newsletter appears, the number of licenced sites should exceed 200.

Lists of all Genstat sites, classified by country and machine range, are given overleaf. Anyone wishing to contact another site should first approach the Genstat Co-ordinator at NAG Central Office.

GENSTAT SITES LISTED BY COUNTRY
=====

AUSTRALIA

APPLIED COMPUTING NEDLANDS W AUSTRALIA
AUSTRALIAN DEP HEALTH CANBERRA
CSIRO AUSTRALIA (COMP RESEARCH)
CSR LIMITED SYDNEY AUSTRALIA
INST FAMILY STUDIES MELBOURNE AUSTRALIA
MITCHELL COLL/AGRISearch BATHURST NSW AUSTRALIA
NSW DEPT AGRIC HAYMARKET AUSTRALIA
S AUSTRALIA DEPT AGRIC ADELAIDE
S AUSTRALIA INS TECH (COM CEN) POURAKA
SEQEB BRISBANE AUSTRALIA
TASMANIAN TREASURY HOBART
U ADELAIDE (WAITE INS) S. AUSTRALIA
U ADELAIDE S AUSTRALIA
U AUSTRALIA NAT (COM SER CEN) CANBERRA
U JAMES COOK (COM CEN) N. QUEENSLAND AUS
U MACQUARIE (COMP CENT) NORTH RYDE AUSTRALIA
U NEW ENGLAND (COM CEN) ARMIDALE AUS
U QUEENSLAND (PRENTICE COM CEN) ST. LUCIA AUS
U SYDNEY NSW (FISHER LIB)
U TASMANIA (COMP CEN) HOBART AUSTRALIA
U W AUSTRALIA (RAINE MED STATS) NEDLANDS
VICTORIA DEP OF AGRIC MELBOURNE AUSTRALIA
W AUSTRALIA DEPT AGRICULTURE S. PERTH

AUSTRIA

U SALZBURG (EDV) AUSTRIA

BELGIUM

U GHENT STATE (CEN DIG COM CEN) BELGIUM

BRAZIL

EMBRAPA BRASILIA BRAZIL

CANADA

AGRICULTURE CANADA CE FARM OTTAWA
PACIFIC BIOLOGICAL STN NANAIMO
U MCGILL (MATHS)
U TORONTO (FAC DENTISTRY) CANADA

CHILE

U CHILE (AGRON) SANTIAGO DE CHILE

COLOMBIA

CIAT CALI COLUMBIA

DENMARK

DANMARKS GEOLOGISKE UNDERSOEGELSE KOEBENHAVN
ELSAM FREDERICIA DENMARK
KOEBENHAVN SC ECON BUS ADMIN DENMARK
NEUCC LYNGBY (EDB-CENT) DENMARK
NOVO RES INST BAGSVAERD DENMARK
RECKU KOEBENHAVN DENMARK
RIGSHOSPITALET KOEBNHAVN DENMARK
U AALBORG (COMP CENT) DENMARK
U AARHUS (RECAU)
U ODENSE (DATACENTER) DENMARK

EIRE

AGRICULTURAL INST DUBLIN
GUINNESS DUBLIN EIRE
U COLL CORK (COMP BUREAU) EIRE

FRANCE

CENT AIR ST-CYR FRANCE
CENT INTERNAT RECH DERMATOL VALBONNE FRANCE
CENT NAT RES ZOOTEC JOUY-EN-JOSAS FRANCE
U PARIS SUD (MATH) FRANCE
U PAUL SABATIER TOULOUSE
UTAC LINAS-MONTLHERY FRANCE

WEST GERMANY

INST NUM STAT KOELN W. GERMANY
U BERLIN FREE W. GER.
U BERLIN TECH (INST STAT OEK, & OR)
U DUISBURG (HSRZ) W. GERMANY
U DUSSELDORF (RZ) W. GERMANY
U HOHENHEIM STUTTGART W GERMANY
U OSNABRUECK (RZ) W GERMANY

HONG KONG

U & POLY COMP CENT HONG KONG

ICELAND

U ICELAND (COMP SER) REYKJAVIK

INDIA

ICRISAT PATANCHERU AP INDIA

ITALY

CINECA BOLOGNA ITALY
ISTIT APPL CALCOLO ROMA ITALY
U FIRENZE (CENT DI CALC) ITALY
U GENOVA (INST DI MAT) ITALY
U NAPOLI (CENTRO DI CALC) ITALY
U PADOVA (CENT CALC)
U PAVIA (CENT CALC)
U ROMA (CENT DI CALC INTERFAC) ITALY

JAPAN

FUYO TOKYO JAPAN

KUWAIT

U KUWAIT (COM SER) ADELIYAH

MEXICO

U MEXICO NACIONAL AUTONOMA

THE NETHERLANDS

EUROPEAN SPACE AGENCY
GIST-BROCADES N.V. DELFT NETHERLANDS
INST TNO THE HAGUE NETHERLANDS
ITAL WAGENINGEN NL
LANDINRICHTINGSDIENST UTRECHT NETHERLANDS
PROEFSTATION TUINBOUW ONDER GLAS NETH
U UTRECHT (SOC WET CEN DAT ANAL) NLANDS
U WAGENINGEN AGRIC (COMP CEN) NL

NEW ZEALAND

DSIR (PHYS & ENG) LOWER HUTT N. ZEALAND
LEVIN HORTIC.RES.STN.LEVIN N. ZEALAND
N.Z. FOREST PRODUCTS TOKOROA NEW ZEALAND
N.Z. MIN WORKS & DEVELOP WELLINGTON NORTH N.Z.
N.Z. STATE SERVICES COMMISSION (CSD)
U AUCKLAND (COM CEN) NEW ZEALAND
U CANTERBURY (LINCOLN COLL) N ZEALAND
U MASSEY (COMP CEN) PALMERSTON NORTH N.Z.
U OTAGO (COMP CEN) DUNEDIN NZ
U WELLINGTON (COMP CEN) N ZEALAND

NIGERIA

U AHMADU BELLO (COMP CEN) NIGERIA

NORWAY

U TROMSO (EDB-SENTRET) NORWAY

PAPUA NEW GUINEA

PAPUA NEW GUINEA NAT COMP CEN WAIGANI

SENEGAL

INST SENEGALAIS RES AG (OCEANOGRAPHY) DAKAR

SOUTH AFRICA

PREMIER GROUP JO'BURG
U FORT HARE SOUTH AFRICA
U NATAL SA (COMP CEN)
U SOUTH AFRICA PRETORIA

SWEDEN

U LUND (COMP CEN) SWEDEN
U STOCKHOLM (QZ DATA CENTER) SWEDEN
U UPPSALA (DATACENTER) SWEDEN

SYRIA

C.E.R.S. SYRIA C/O SYSTEME LEEDS
ICARDA ALEPPO SYRIA

TAIWAN R.O.C.

PIG RESEARCH INST TAIWAN

TURKEY

U EGE (E.H.B.ENS.) IZMIR TURKEY

UNITED KINGDOM

A. & A.E.E. BOSCOMBE DOWN
AERE (COMP SCI) HARWELL
BP INTERNATIONAL LTD LONDON
BRITISH - AMERICAN TOBACCO SOUTHAMPTON
BRITISH SHIP RES ASSN WALLSEND
BUSINESS STATISTICS OFFICE NEWPORT (VIA CCA)
CADBURY SCHWEPPE'S LTD BIRMINGHAM
CAMBRIDGE COLL ARTS & TECH (COM CEN)
CEGB NW REGION STOCKPORT
CLINICAL RES CENT HARROW
DES MOWDEN HALL DARLINGTON
DOE BUILDING RES STN GARSTON
DOE (TRRL) CROWTHORNE
FORESTRY COMMISSION FARNHAM
GKN TECHNOLOGY LTD WOLVERHAMPTON
GLAXO OPERATIONS UK GREENFORD
GLAXO OPERATIONS UK ULVERSTON
GREATER LONDON COUNCIL (DCTR-GNL'S DEPT)
HMGC MILTON KEYNES
HUNTINGDON RESEARCH CENTRE
IMPERIAL CANCER RESEARCH FUND LONDON
IMPERIAL TOBACCO (ITL) BRISTOL
INST NAVAL MEDICINE GOSPORT
INST OCCUPATIONAL MEDICINE EDINBURGH
LIFE SCIENCE RESEARCH OCCOLD SUFFOLK
MET OFFICE BRACKNELL
MIN AG FISH FOOD (ADAS) WEYBRIDGE
MOD(A&AEE) BOSCOMBE DOWN
MRC NAT INST MED RES MILL HILL
NAPIER COLL COMM & TECH (COM CEN) EDINBURGH
NAT FOUNDATION ED RESEARCH SLOUGH
NEGAS (COMP SERV) LEEDS
NERC COMP SERV BIDSTON BIRKENHEAD
NORTHERN REG HEALTH AUTH NEWCASTLE
NORTH STAFFS HOSPS (COMP DEP) HARTSHILL
NORTH WEST WATER (COMP CENT) WARRINGTON
PFIZER CENTRAL RESEARCH SANDWICH
POLY BRIGHTON (COMP CENT)
POLY CENT LONDON (COMP UNIT)
POLY CITY LONDON (COM CEN)
POLY COVENTRY (LANCHESTER) (COM CEN)
POLY HUDDERSFIELD
POLY KINGSTON (COMP UNIT)

POLY LIVERPOOL (COMP SER)
POLY NORTH EAST LONDON (COMP CEN)
POLY PLYMOUTH (COMP CENT)
POLY SOUTH BANK (COMP CENTRE) LONDON
POLY THAMES (COMP SERV) LONDON
POLY TRENT (DEP COM SER) NOTTS
RANK XEROX ENG GROUP WELWYN
RECKITT & COLMAN HULL
R.I.B.A. LONDON
ROTHAMSTED ARC COMP CENT HARPENDEN
R.S.P.B. SANDY BEDFORDSHIRE
RUTHERFORD LAB (IBM) COMP & AUTO DIV
SCHLUMBERGER CAMBRIDGE RESEARCH
SCOTTISH OFFICE (COMP SERV) EDINBURGH
SEVERN-TRENT WATER AUTHORITY
SHELL RESEARCH CENTRE SITTINGBOURNE
SHELL THORNTON RESEARCH CENTRE CHESTER
SHELL UK LTD LONDON
SMITH KLINE & FRENCH WELWYN GARDEN CITY
THAMES WATER D.O.P. (COMPUTING) READING
UNILEVER RES LAB PORT SUNLIGHT
U.K.A.E.A. WINFRITH DORSET
U ABERDEEN (COMP CEN)
U ASTON (COMP SER)
U BATH (SWRCC)
U BELFAST QUEEN'S (COMP CEN)
U BRADFORD (COMP LAB)
U BRISTOL (SWUCN)
U BRISTOL (COMP CEN)
U CAMBRIDGE (COMP LAB)
U COLL CARDIFF (COMP CEN)
U COLL N WALES (COMP LAB) BANGOR
U EDINBURGH (PLU)
U ESSEX (COMP CEN)
U HULL (COMP CEN)
U LANCASTER (COMP LAB)
ULCC (ADVISORY SERV)
U LEEDS (COMP STUDIES)
U LIVERPOOL (COMP LAB)
U LONDON CHELSEA COLL (COMP CEN)
U LONDON IMP COLL (COMP CEN)
U LONDON QMC (COMP CEN)
U MANCHESTER (RCC)
U NEWCASTLE (NUMAC)
U NOTTINGHAM (COMP CEN)
U OXFORD (OUCS)
U READING (COMP UNIT)
U SALFORD (COMP LAB)
U SHEFFIELD (COMP SER)
U SOUTHAMPTON (COMP SERVICES)
U STIRLING (COMP UNIT)
U ST ANDREWS (COMP LAB)
U SURREY (COMP UNIT)
U SUSSEX (COMP CEN)
U WARWICK (COMP UNIT)
U YORK (COMP SCI DEPT)
W SUSSEX C.C. (COUNTY SURVEYOR) CHICHESTER
WATER RESEARCH CENTRE MEDMENHAM
WELSH WATER AUTHORITY BRECON POWYS
WESSEX REGIONAL HEALTH AUTH (COMP CENT) WINCHESTER

UNITED STATES OF AMERICA

N.Y. STATE AGRIC EXP STN GENEVA NY USA
N.Y. STATE DEPT. HEALTH ALBANY N.Y. USA
U CALIFORNIA-DAVIS (COMP CENT) U.S.A.
U CALIFORNIA-RIVERSIDE U.S.A.
U CINCINNATI (STATS) USA
U COLORADO STATE FORT COLLINS U.S.A.
U CORNELL NEW YORK
U MICHIGAN STATE (COMP CEN) E. LANSING USA

YUGOSLAVIA

U ZAGREB (COMP CENT) YUGOSLAVIA

GENSTAT SITES LISTED BY MACHINE RANGE
=====

BURROUGHS

GIST-BROCADES N.V. DELFT NETHERLANDS
NSW DEPT AGRIC HAYMARKET AUSTRALIA
U CALIFORNIA-DAVIS (COMP CENT) U.S.A.
U DUISBURG (HSRZ) W. GERMANY
U GENOVA (INST DI MAT) ITALY
U MEXICO NACIONAL AUTONOMA
U OTAGO (COMP CEN) DUNEDIN NZ
U SOUTH AFRICA PRETORIA
U TASMANIA (COMP CEN) HOBART AUSTRALIA
U WARWICK (COMP UNIT)
VICTORIA DEP OF AGRIC MELBOURNE AUSTRALIA

CDC

BRITISH - AMERICAN TOBACCO SOUTHAMPTON
CINECA BOLOGNA ITALY
CSIRO AUSTRALIA (COMP RESEARCH)
DOE (TRRL) CROWTHORNE
INST TWO THE HAGUE NETHERLANDS
LANDINRICHTINGSDIENST UTRECHT NETHERLANDS
RIGSHOSPITALET KOEBNHAVN DENMARK
S AUSTRALIA DEPT AGRIC ADELAIDE
S AUSTRALIA INS TECH (COM CEN) POURAKA
U AALBORG (COMP CENT) DENMARK
U AARHUS (RECAU)
U ADELAIDE (WAITE INS) S. AUSTRALIA
U AHMADU BELLO (COMP CEN) NIGERIA.
U BERLIN FREE W. GER.
U BOLOGNA ITALY
U BRADFORD (COMP LAB)
U COLORADO STATE FORT COLLINS U.S.A.
ULCC (ADVISORY SERV)
U LONDON IMP COLL (COMP CEN)
U MANCHESTER (RCC)
U MICHIGAN STATE (COMP CEN) E. LANSING USA
U SYDNEY NSW (FISHER LIB)
U TROMSO (EDB-SENTRET) NORWAY
U UTRECHT (SOC WET CEN DAT ANAL) NLANDS
U W AUSTRALIA (RAINE MED STATS) NEDLANDS

CII IRIS

CENT NAT RES ZOOTECH JOUY-EN-JOSAS FRANCE
U FIRENZE (CENT DI CALC) ITALY

DEC SYSTEMS 10 & 20

CLINICAL RES CENT HARROW
IMPERIAL CANCER RESEARCH FUND LONDON
MRC NAT INST MED RES MILL HILL
POLY CENT LONDON (COMP UNIT)
POLY CITY LONDON (COM CEN)
POLY KINGSTON (COMP UNIT)
POLY LIVERPOOL (COMP SER)
POLY SOUTH BANK (COMP CENTRE) LONDON
POLY TRENT (DEP COM SER) NOTTS
U AUSTRALIA NAT (COM SER CEN) CANBERRA
U COLL N WALES (COMP LAB) BANGOR
U ESSEX (COMP CEN)
U JAMES COOK (COM CEN) N. QUEENSLAND AUS
U NEW ENGLAND (COM CEN) ARMIDALE AUS
U QUEENSLAND (PRENTICE COM CEN) ST. LUCIA
U WAGENINGEN AGRIC (COMP CEN) NL
U YORK (COMP SCI DEPT)

DEC VAX-11

AGRICULTURAL INST DUBLIN
AGRICULTURE CANADA CE FARM OTTAWA
APPLIED COMPUTING NEDLANDS W AUSTRALIA
CAMBRIDGE COLL ARTS & TECH (COM CEN)
C.E.R.S. SYRIA C/O SYSTIME LEEDS
CSR LIMITED SYDNEY AUSTRALIA
DANMARKS GEOLOGISKE UNDERSOEGELSE KOEBENHAVN
DSIR (PHYS & ENG) LOWER HUTT N. ZEALAND
EUROPEAN SPACE AGENCY
GKN TECHNOLOGY LTD WOLVERHAMPTON
ICARDA ALEPPO SYRIA
ICRISAT PATANCHERU AP INDIA

ITAL WAGENINGEN NL
N.Y.STATE DEPT. HEALTH ALBANY N.Y. USA
PACIFIC BIOLOGICAL STN NANAIMO
PFIZER CENTRAL RESEARCH SANDWICH
POLY BRIGHTON (COMP CENT)
RANK XEROX ENG GROUP WELWYN
SCHLUMBERGER CAMBRIDGE RESEARCH
SHELL RESEARCH CENTRE SITTINGBOURNE
U ADELAIDE S AUSTRALIA
U BELFAST QUEEN'S (COMP CEN)
U CALIFORNIA-DAVIS (COMP CENT) U.S.A.
U CALIFORNIA-RIVERSIDE U.S.A.
U CANTERBURY (LINCOLN COLL) N ZEALAND
U COLL CORK (COMP BUREAU) EIRE
U ICELAND (COMP SER) REYKJAVIK
U LANCASTER (COMP LAB)
U MACQUARIE (COMP CENT) NORTH RYDE AUSTRALIA
U OXFORD (OUCS)
U STIRLING (COMP UNIT)
U ST ANDREWS (COMP LAB)
U SUSSEX (COMP CEN)
UTAC LINAS-MONTLHERY FRANCE
W SUSSEX C.C. (COUNTY SURVEYOR) CHICHESTER
WATER RESEARCH CENTRE MEDMENHAM

HARRIS

POLY COVENTRY (LANCHESTER) (COM CEN)
UNILEVER RES LAB PORT SUNLIGHT
U ASTON (COMP SER)
U LONDON CHELSEA COLL (COMP CEN)

HEWLETT PACKARD 3000

GLAXO OPERATIONS UK ULVERSTON

HONEYWELL SERIES 60

NERC COMP SERV BIDSTON BIRKENHEAD
U ABERDEEN (COMP CEN)
U BRISTOL (COMP CEN)
U COLL CARDIFF (COMP CEN)
U PAUL SABATIER TOULOUSE
U PAVIA (CENT CALC)

HONEYWELL SIGMA 6

A. & A.E.E. BOSCOMBE DOWN
HMGCC MILTON KEYNES
MOD(A&AEE) BOE MBE DOWN

IBM 360 & SIMILAR

AREE (COMP SCI) HARWELL
AUSTRALIAN DEP HEALTH CANBERRA
BRITISH SHIP RES ASSN WALLSEND
CADBURY SCHWEPPE LTD BIRMINGHAM
CENT AIR ST-CYR FRANCE
CENT INTERNAT RECH DERMATOL VALBONNE FRANCE
CIAT CALI COLUMBIA
EMBRAPA BRASILIA BRAZIL
FUYO TOKYO JAPAN
GREATER LONDON COUNCIL (DCTR-GNL'S DEPT)
IMPERIAL TOBACCO (ITL) BRISTOL
INST SENEGALAIS RES AG (OCEANOLOG) DAKAR
ISTIT APPL CALCOLO ROMA ITALY
MET OFFICE BRACKNELL
MITCHELL COLL/AGRISEARCH BATHURST NSW AUSTRALIA
NAT FOUNDATION ED RESEARCH SLOUGH
NEUCC LINGBY (EDB-CENT) DENMARK
N.Z. FOREST PRODUCTS TOKOROA NEW ZEALAND
N.Z. MIN WORKS & DEVELOP WELLINGTON NORTH N.Z.
N.Z. STATE SERVICES COMMISSION (CSD)
PIG RESEARCH INST TAIWAN
PREMIER GROUP JO'BURG
RECKITT & COLMAN HULL
RUTHERFORD LAB (IBM) COMP & AUTO DIV
SCOTTISH OFFICE (COMP SER) EDINBURGH
SEQEB BRISBANE AUSTRALIA
SHELL RESEARCH CENTRE SITTINGBOURNE
SHELL THORNTON RESEARCH CENTRE CHESTER
SMITH KLINE & FRENCH WELWYN GARDEN CITY
U AUCKLAND (COM CEN) NEW ZEALAND
U BERLIN TECH (INST STAT OER, & OR)

U CAMBRIDGE (COMP LAB)
U CHILE (AGRON) SANTIAGO DE CHILE
U CINCINNATI (STATS) USA
U CORNELL NEW YORK
U EGE (E.H.B.ENS.) IZMIR TURKEY
ULCC (ADVISORY SERV)
U LEEDS (COMP STUDIES)
U LIVERPOOL (COMP LAB)
U MCGILL (MATHS)
U NEWCASTLE (NUMAC)
U PADOVA (CENT CALC)
U PARIS SUD (MATH) FRANCE
U & POLY COMP CENT HONG KONG
U SALZBURG (EDV) AUSTRIA
U STOCKHOLM (QZ DATA CENTER) SWEDEN
U TORONTO (FAC DENTISTRY) CANADA
U UPPSALA (DATACENTER) SWEDEN
U WELLINGTON (COMP CEN) N ZEALAND

ICL SYSTEM 4

ROTHAMSTED ARC COMP CENT HARPENDEN
U BRISTOL (SWUC)

ICL 1900 & SIMILAR

DES MOWDEN HALL DARLINGTON
DOE BUILDING RES STN GARSTON
HUNTINGDON RESEARCH CENTRE
INST NAVAL MEDICINE GOSPORT
MIN AG FISH FOOD (ADAS) WEYBRIDGE
NORTHERN REG HEALTH AUTH NEWCASTLE
PAPUA NEW GUINEA NAT COMP CEN WAIGANI
THAMES WATER D.O.P. (COMPUTING) READING
U HULL (COMP CEN)
U READING (COMP UNIT)
U SALFORD (COMP LAB)
WELSH WATER AUTHORITY BRECON POWYS
WESSEX REGIONAL HEALTH AUTH (COMP CENT)

ICL UPPER 2900

BUSINESS STATISTICS OFFICE NEWPORT (VIA CCA)
CEGB NW REGION STOCKPORT
GUINNESS DUBLIN EIRE
NEGAS (COMP SERV) LEEDS
NORTH STAFFS HOSPS (COMP DEP) HARTSHILL
NORTH WEST WATER (COMP CENT) WARRINGTON
NOVO RES INST BAGSVAERD DENMARK
N.Z. STATE SERVICES COMMISSION (CSD)
POLY HUDDERSFIELD
SEVERN-TRENT WATER AUTHORITY
U.K.A.E.A. WINFRITH DORSET
U BATH (SWRCC)
U EDINBURGH (PLU)
U HOHENHEIM STUTTGART W GERMANY
U LANCASTER (COMP LAB)
U LONDON QMC (COMP CEN)
U NOTTINGHAM (COMP CEN)
U OXFORD (OUCS)
U SOUTHAMPTON (COMP SERVICES)

MODCOMP

INST NUM STAT KOELN W. GERMANY

PERKIN ELMER

LIFE SCIENCE RESEARCH OCCOLD SUPPLK

GENSTAT 4.04

The new release of Genstat is now in service at Rothamsted. The first versions produced will probably be for IBM and PRIME computers, incorporating amendments to correct errors discovered in the initial test period. The source will then be sent out for conversion to other machines.

Manual updates are available from NAG. We discuss below the main differences between 4.03 and 4.04.

PRIME

CSR LIMITED SYDNEY AUSTRALIA
DOE BUILDING RES STN GARSTON
FORESTRY COMMISSION FARNHAM
GLAXO OPERATIONS UK GREENFORD
GLOUCESTERSHIRE C.A.T. CHELTENHAM
INST FAMILY STUDIES MELBOURNE AUSTRALIA
INST OCCUPATIONAL MEDICINE EDINBURGH
KOBENHAVN SC ECON BUS ADMIN DENMARK
LEVIN HORTIC.RES.STN.LEVIN N. ZEALAND
NAPIER COLL COMM & TECH (COM CEN) EDINBURGH
N.Y. STATE AGRIC EXP STN GENEVA NY USA
POLY NORTH EAST LONDON (COMP CEN)
POLY PLYMOUTH (COMP CENT)
POLY THAMES (COMP SERV) LONDON
R.I.B.A. LONDON
ROTHAMSTED ARC COMP CENT HARPENDEN
R.S.P.B. SANDY BEDFORDSHIRE
SHELL RESEARCH CENTRE SITTINGBOURNE
U FORT HARE SOUTH AFRICA
U MASSEY (COMP CEN) PALMERSTON NORTH N.Z.
U SHEFFIELD (COMP SER)
U SURREY (COMP UNIT)
W AUSTRALIA DEPT AGRICULTURE S. PERTH

SEL

SIEMENS BS2000

PROEFSTATION TUINBOUW ONDER GLAS NETH
U DUSSELDORF (RZ) W. GERMANY
U GHENT STATE (CEN DIG COM CEN) BELGIUM

TELEFUNKEN TR400

U OSNABRUECK (RZ) W GERMANY

UNIVAC 1100

BP INTERNATIONAL LTD LONDON
ELSAM FREDERICIA DENMARK
RECKU KOBENHAVN DENMARK
SHELL UK LTD LONDON
TASMANIAN TREASURY HOBART
U AUSTRALIA NAT (COM SER CEN) CANBERRA
U KUWAIT (COM SER) ADELIYAH
U LUND (COMP CEN) SWEDEN
U NAPOLI (CENTRO DI CALC) ITALY
U NATAL SA (COMP CEN)
U ODENSE (DATACENTER) DENMARK
U PARIS SUD (MATH) FRANCE
U ROMA (CENT DI CALC INTERFAC) ITALY
U ZAGREB (COMP CENT) YUGOSLAVIA

*M.G. Richardson
NAG Central Office
Mayfield House
256 Banbury Road
Oxford OX2 7DE
U.K.*

*H.R. Simpson
Rothamsted Experimental Station
Harpenden
Hertfordshire AL5 2JQ
U.K.*

CHANGES.....

In a few respects Genstat 4.04 will not be compatible with earlier releases.

1. The options of the 'REFERENCE' statement have been changed. Only NID, NUNN and RNG remain as before; the LINE option has been removed (and there is now no check on the size of the output file) and the other options have been subsumed into a single PRINT option, with additional facilities for interactive working and debugging macros.
2. In the interests of consistency, all options which in one way or another refer to marginal totals now have the setting TOTAL. In particular:

'JOIN/ and 'PRINT/MARGIN = TOTAL' (not TOT)
'MARGIN/MODE = TOTAL or MEAN' (not T or M)
'TABULATE/MARGIN = TOTAL' (not TOT)

3. The backing store directives have been drastically revised. The 'FILE' command has been replaced by FILE options which should be set to an integer (1,2,...). The old concept of a userfile identifier has been discarded (it had no existence outside the job anyway). The ordinals 1,2,... are machine independent (so your Genstat job will run without change on another machine) but the channel numbers corresponding to these integers are machine dependent.

So we may now write

'PUT' or 'GET' subfile \$ structure list

which by default will access the workfile. If FILE = f>0 is set, userfile f will be accessed directly.

'FETCH' or 'SAVE' subfile list

will, by default, access userfile 1.

The effects of these changes on programs which use backing store are illustrated in the following article in this Newsletter.

4. New directives 'BIN' (binary input) and 'BOUT' (binary output) have been added to transfer unformatted data. The directive name 'SYMAP' has been retained as a synonym of 'BOUT' but to reproduce its previous effect

'SYMAP/FORM = p'

must be used. By default this will write to userfile 1; this may be overridden by a FILE option.

..... AND ENHANCEMENTS

1. Interactive facilities

PRINT options of 'REFERENCE' (or 'INPUT') may be used

- a) to produce a copy of the input. If PRINT letter C is set and a secondary output file nominated, all input instructions are copied to this file. It will need editing (before it can be used as an input file for a subsequent job) if incorrect commands have been given or the secondary output channel has been used for other purposes.
- b) to control extended fault messages. On all machines except ICL System 4, diagnostic reports will be followed by an extra line of, hopefully, helpful text. These may be suppressed by setting the PRINT option without the letter F. Our thanks are due to John Byrne, of the University of York, for adding this feature to the DEC-10 version of 4.03 and spurring us to make it generally available.

A new directive called 'HELP' has been introduced to provide access to information about Genstat syntax. The directive is described on page 16.

Finally, the contraction R is accepted as a synonym of RUN.

2. Macros

Chasing elusive errors in macros can be difficult and time consuming. PRINT options of 'REFERENCE', 'INPUT' or 'USE' can now be set to force macro instructions to be printed as they are compiled, together with the statement number of the first statement on each line. It should therefore be possible to determine precisely which instruction has failed, even when a nest of macros is used at run time within a loop.

The 'ENDMACRO' statement now has an option LOCALS which, if set to DESTROY, will recover not only the data space but also the directory space occupied by the current instances of the local identifiers of a macro used at run-time.

3. 'INPUT' and 'OUTPUT' have acquired new options. RECL may be used to change the assumed maximum length of a record, set on entry to the program. For output, RECL will allow interactive users to override the assumption that all output is to a terminal of restricted width.

'OUTPUT' also has an option NLP to reset the number of lines per page assumed by 'GRAPH'.

'BIN' and 'BOUT', which transfer unformatted data records, have been introduced primarily for communication with other Fortran programs. They can also be used to save space and to input data efficiently.

A new option SQUASH for 'PRINT' allows the user to prevent any blank lines being output by 'PRINT' unless explicitly requested in the format of a 'PRINT/C' command.

4. Graphics

'GRAPH' can now output to graphics devices. However, for a particular device at a particular site, four special routines are needed. We have written (and tested) the routines needed for the Benson plotter connected to the ICL System 4 and the GOC terminal connected to a Prime 550 at Rothamsted. Specifications of the routines, and the source for all versions in use here, will be made available to converters and we will try to coordinate work outside Rothamsted to produce suitable versions for individual sites.

5. Dynamic formats

After a format identifier has been declared by 'SET/F' it may be used in a format list; its contents will be substituted by the compiler.

```
'SET/F' G1=2,6X : G2=3.6,/  
'READ' X $ 2(G1), 3X,G2
```

has the same effect as

```
'READ' X $ 2(2,6X), 3X,3.6,/'
```

Although not explicitly proscribed by the manual, this could not be done in 4.03.

'READ' and 'PRINT' have an extra option FMT and 'EQUATE' has two new options LFMT and RFMT, all of which can be set to identifiers declared by 'SET/F'. When one of these statements is executed, if a format option has been set, then any explicit format given with the instruction is ignored and the format identifier used to determine the required format. Be careful: if the identifier is not of the correct type a fatal error will ensue. If it is of the right type, but without values, a null format will be assumed.

These new facilities will allow considerably more flexible use of formats.

6. Calculate and table operations

'CALCULATE' can now operate on integer and factor structures directly. The values of any structure of integer mode on the right hand side of an equation are converted to real values before they are used; if the (real) results of an expression are to be stored as integers they are converted to the nearest integer. Thus $5/3$ will be treated as $5.0/3.0 = 1.6$ until it has to be stored as an integer when it will be rounded to 2.

The VARFAC function has an optional second argument list to allow factors with non-numeric level names to be handled.

New functions NROW, NCOL, NLEV and TYPE will allow macro writers, in particular, to access attributes of structures.

'TABULATE' can now form tables of within-cell minima, maxima and variances. 'MARGIN' can form minima, maxima, variances and medians from the body of a table (i.e. between-cells).

7. Linear regression

A new directive called 'PREDICT', which provides summaries of a fitted regression model, was described in an article in Newsletter 9. It calculates and displays tables of mean values, optionally with standard errors, classified by factors in the model. The means are adjusted for any variates also in the model and several methods are available for adjusting for effects of some of the factors.

The directives for stagewise regression, 'BEST', 'WORST' and 'MINIMISE', have new options to control their operation. When considering a term for addition to (or deletion from) the current model, the change in the residual mean square may be compared to a partial F statistic, supplied in option FIN (or FOUT). The letter R may now be used in the PRINT option to control printing of the list of residual mean squares - this list is printed by default.

New option settings and keywords have been added to improve control over output from the regression directives. In the 'Y' directive, ERROR=INVERSE specifies the likelihood of the inverse Gaussian distribution. In the ANDEV option of directives like 'FIT' the new setting ANDEV=S prints the summary analysis table using the mean square from the smallest residual sum of squares as the divisor of ratios, rather than the smallest residual mean square, as at present. The keyword DPM is available in these directives to save the diagonal of the projection matrix: this may be used to form standard errors of either fitted values or residuals.

8. Analysis of variance

'ANOVA' has three new options

a) SEED = scalar

The ANOVA dummy analysis currently uses a dummy working variate containing random numbers. The scalar sets a starting value for the random numbers (default 12345). This can be used to check a suspect analysis e.g. if a model term is thought to be incorrectly listed as aliased.

b) TOL = variate

Sets various constants involved in the analysis.

c) PROB = N, Y

If PROB = Y, probabilities are printed for the variance ratios in the AOV table (assuming that they follow F distributions).

Also, a saved 'ANOVA' output segment can be used to print further output from an analysis, without requiring the variate to be analysed again.

E.g. 'ANOVA/PR=10' Y ; OUT=OY

can be followed later by, for example, 'ANOVA/PR=2' OUT=OY

'EXTRACT' has a new nameable list COEF.

For a block model term (only) this gives covariate regression coefficients estimated in the stratum for that term, output as a variate of length equal to the number of covariates.

9. Optimisation

A new algorithm has been added, to remove the present constraint on numbers of parameters in a non-linear model. The choice of algorithm is controlled by the option METHOD; the setting GN specifies the new Gauss-Newton algorithm, which has been modified to deal with all the likelihood functions available. (This algorithm is not available for general optimisation with the setting LIK=1.) If there are more than 6 non-linear parameters, the new algorithm is automatically used; there is no longer any limit on the number of linear parameters with the setting LIK=3.

10. Time series

A new directive called 'FOURIER' has been introduced to provide the basic tool for methods of spectral analysis. It uses a fast algorithm to calculate the Fourier transform of a real or complex sequence. The cosine transform of a real sequence is also available.

A PRINT option has been added to the directive 'DERIVE'. The setting PRINT=G will produce a graph of the calculated function, that is, a "correlogram" if the function used is one of the correlation functions ACF, PACF, CCF or THACF.

The LIK option of 'ESTIMATE' has now been implemented; it was described in the manual for release 4.03 but was not available then. The default setting is LIK=EXACT, which gives a more accurate (though more time-consuming) calculation of the likelihood than the only setting available previously, LIK=LS. The setting LIK=MARG may be used to reduce bias in parameter estimates derived for small samples.

*P.W. Lane
R.W. Payne
H.R. Simpson
Rothamsted Experimental Station
Harpenden
Hertfordshire AL5 2JQ
U.K.*

BACKING STORE CHANGES IN GENSTAT 4.04

Facilities for storage and retrieval are being improved in Release 4.04 and the syntax is being revised at the same time. These changes will result in many existing Genstat programs having to be edited. This article explains why this change was made, gives details of the amended syntax (for users not having manual amendments at hand) and gives examples of the old and new syntax.

The major changes are to the 'GET' and 'PUT' directives, where userfiles can now be accessed directly. Most Genstat programs using backing store read structures from one subfile of a userfile (e.g. the macro

library) into core. In the past the subfile had to be copied into the workfile and then the requested structures read into core. The stage of copying to or from the workfile is no longer necessary, so there will be savings in disc space and computer time. The ability to create directly a userfile containing one subfile is also provided. This facility will be especially useful to people using interactive Genstat, when they want to stop a job and save the compiled code and structures for future use (using the COMP=DUMP option).

It has always been possible to specify a userfile in the 'PUT' and 'GET' directives but unexpected problems could arise. For example, the 'GET' directive could be used to copy structures from a userfile. Formerly, the 'GET' directive copied all the subfiles in the userfile into the workfile and the requested structures were then copied into core. However, if subfiles of the same name existed in the userfile and workfile, the job failed. In 4.04, the requested structures will be copied directly into core from the subfile in the userfile so that it does not matter if subfiles in the workfile and the userfile have the same name.

At the same time as implementing the important changes described above, it was felt necessary to simplify the syntax. In the past, the number of the channel to which a userfile had been assigned was specified by the 'FILE' directive. The 'FILE' directive has been removed and replaced by a FILE option with a scalar setting. The setting 0 corresponds to the workfile and setting 1 to the userfile assigned to the first available channel: The default value is the lowest possible for each directive, i.e. 0 for 'DISPLAY', 'PUT' and 'GET', 1 for 'SAVE' and 'FETCH'.

In the revised syntax for the backing store directives given below, only changes from 4.03 Genstat are given in detail.

'PUT or GET/option list' subfile (\$ sequence of renaming segments of core structures)

where option list is

FILE=scalar,PRINT=name,LIST=name,COMP=name

The values for the FILE option are

0 = workfile (default)
1 = userfile assigned to first channel
2 = userfile assigned to second channel
etc

(the number of channels varies according to installation).

The settings for the PRINT option are now

Z = no printing of changed catalogues (default) (formerly N)
C = printing of changed catalogues (formerly Y)

'SAVE or FETCH/option list' (sequence of renaming segments of subfiles of the workfile)

'DISPLAY/option list' (list of subfiles)

where option list for 'SAVE', 'FETCH' and 'DISPLAY' is

FILE=scalar,PRINT=name,LIST=name.

The values of the option FILE are the same as in 'PUT' except that 0 is not permitted for SAVE or fetch where the default will be 1. The PRINT option is the same as in 'PUT', except for DISPLAY where the default is C.

'CLEAR/option list' (list of subfiles)

where option list is

PRINT=name,LIST=name.

The PRINT option is the same as in 'PUT'.

The examples below illustrate how programs need to be modified.

A) *STORING SOME STRUCTURES PERMANENTLY*

4.03
'PUT' SUB \$ DATA(1...10)
'FILE' PERM=48
'FILE ATTACHED TO CHANNEL 48:-
47 RESERVED FOR MACRO LIBRARY AT
ROTHAMSTED''
'SAVE' PERM \$ SUB
alternative
'FILE' PERM=48
'PUT' PERM,SUB \$ DATA(1...10)

4.04
'PUT' SUB \$ DATA(1...10)
'SAVE/FILE=2' SUB
'FILE ATTACHED TO SECOND AVAILABLE
BACKING STORE CHANNEL:- 48 AT
ROTHAMSTED''
alternative
'PUT/FILE=2' SUB \$ DATA(1...10)

Similar changes apply to 'GET' and 'FETCH' for retrieval.

B) *STORING STRUCTURES PERMANENTLY IN GROUPS IN ONE FILE*

4.03
'PUT' SUB(1) \$ DATA(1...5)
: SUB(2) \$ DATA(6...10)
'FILE' PERM=48
'SAVE' PERM \$ SUB(1,2)

4.04
'PUT' SUB(1) \$ DATA(1...5)
: SUB(2) \$ DATA(6...10)
'SAVE/FILE=2' SUB(1,2)

C) *STORING STRUCTURES PERMANENTLY IN MORE THAN ONE FILE*

4.03
'PUT' SUB(1) \$ DATA(1...5)
: SUB(2) \$ DATA(6...10)
'FILE' PERM(1)=48
: PERM(2)=49
'2ND FILE ATTACHED TO CHANNEL 49'

'SAVE' PERM(1) \$ SUB(1)
: PERM(2) \$ SUB(2)

4.04
'PUT' SUB(1) \$ DATA(1...5)
: SUB(2) \$ DATA(6...10)
'SAVE/FILE=2' SUB(1)
'SAVE/FILE=3' SUB(2)
'2ND FILE ATTACHED TO THIRD AVAILABLE
BACKING STORE CHANNEL:- 49 AT
ROTHAMSTED''

D) *REPLACING ONE SUBFILE IN A FILE, LEAVING OTHERS UNCHANGED*

4.03
'FILE' PERM = 48
'FETCH/LIST=ALL' PERM
'GET/LIST=ALL' SUB(1)
'CALC' DATA(1) = NEW*DATA(1)
'CLEAR' SUB(1)
'PUT' SUB(1) \$ DATA(1...5)
'SAVE/LIST=ALL' PERM

4.04
'FETCH/LIST=ALL,FILE=2'
'GET/LIST=ALL' SUB(1)
'CALC' DATA(1)=NEW*DATA(1)
'CLEAR' SUB(1)
'PUT' SUB(1) \$ DATA(1...5)
'SAVE/LIST=ALL,FILE=2'

*P.W. Lane
A.D. Todd
Rothamsted Experimental Station
Harpenden
Hertfordshire AL5 2JQ
U.K.*

THE HELP DIRECTIVE

A new directive has been introduced in Release 4.04 to make information available to people using Genstat interactively. It may also be used in batch, to print the information for reference; however, we also hope to make all the information available on a reference card.

The directive may be used as follows:

- (1) Index: Type 'HELP' followed by 'RUN' (or by the new synonym for 'RUN' - 'R'). This lists 16 section headings, e.g. data input, linear regression analysis, and gives a section number for each.
- (2) Index of section: Type 'HELP' *section number* 'R'. This gives general information about the section and a list of directive names.
- (3) Syntax of a directive: Type 'HELP' *directive name* 'R'. This gives the form of the directive, a one-line description of its purpose and a list of all options, option settings and keywords, with brief descriptions.
- (4) Further information: The information for some directives takes more than 20 lines, the maximum to be presentable on a video-terminal. To get more, the directive name can be given again with a suffix: possible suffices are listed at the end of the first section of output. For example, 'HELP' CALC(2) 'R' lists functions which summarise structure values.

With some versions of Genstat, the information will not be available because of limitation of storage. The text is presented in lines of 64 characters to ensure it fits on most terminals; there are about 1250 lines available, which take 80K bytes of storage. By contrast, it is hoped to print all the information, legibly, on a reference card measuring 21 x 60 cm, concertina-folded to give 12 pages.

*P.W. Lane
Rothamsted Experimental Station
Harpenden
Hertfordshire AL5 2JQ
U.K.*

DIALLEL CROSSES, PARTIALLY BALANCED INCOMPLETE BLOCK DESIGNS WITH TRIANGULAR ASSOCIATION SCHEMES AND RECTANGULAR LATTICES

INTRODUCTION

What have these designs in common? Apart from esoteric names, all three have a high degree of symmetry and yet cannot be completely analysed by the Genstat 'ANOVA' directive. The treatment structure for the first design cannot be specified, the 'ANOVA' algorithm finds the second design unbalanced and treatment means for the third are only 'approximate'. Of course, other methods of analysis are possible and macros are available to analyse (i) diallel crosses using regression procedures (N.G. Alvey and O. Mayo, personal communications) and (ii) rectangular lattices using 'ANOVA' on the dual of the design (E.R. Williams and D. Ratcliff, CSIRO Macro).

I suggest instead a technique which aids in the analysis and, I think, the understanding of all three designs (and no doubt others). It is based on the idea of repeating the data several times and generating treatment and block factors with each copy of the data. The treatment and block factors are chosen so that the repeated data has a 'balanced' structure and that the required analysis is a meaningful subset (usually the only one) of the analysis of the extended data. Thompson and Baker (1981) found a similar idea useful in embedding several types of data into a generalised linear model framework.

I will demonstrate the technique for each design in turn and then discuss the disadvantages of the method and point out how the Genstat 'ANOVA' algorithm might be improved.

DIALLEL CROSSES

In plant breeding work on p different parental lines, often, all reciprocal crosses can be made and the p^2 progeny used for the comparison of parental lines. Various models and analyses have been suggested (Yates (1947), Hayman (1954), Griffing (1956), Jinks and Mather (1971)). I first consider a simple model that exhibits the difficulty in fitting these classes of models. I take the case where the progeny mean, when a male of line i is crossed with a female of line j , is y_{ij} . One might think of fitting a factorial model

$$y_{ij} = \text{mean} + \text{male line } (i) \text{ effect} + \text{female line } (j) \text{ effect}$$

and this can be easily done in 'ANOVA'. However it might be reasonable to assume the male and female effects of line i to be similar and so want to fit a model

$$y_{ij} = \text{mean} + \text{line } (i) \text{ effect} + \text{line } (j) \text{ effect.} \quad (1)$$

However, each observation has 2 contributions from the line factor and so it is difficult to specify the treatment structure. Suppose, however, that we have two observations $y_{t_{ij1}}$ and $y_{t_{ij2}}$ with expectations

$$y_{t_{ij1}} = \text{mean} + \text{line } (i) \text{ effect}, \quad y_{t_{ij2}} = \text{mean} + \text{line } (j) \text{ effect.}$$

Each such pair $y_{t_{ij1}}$ and $y_{t_{ij2}}$ can be put in a different level of a factor DUMMY (for example put $y_{t_{ij1}}$ and $y_{t_{ij2}}$ into level $(i+(j-1)xp)$). For $p=3$ the data, treatment and block factors will be

$$\begin{array}{l}
 Y = (y_{11} \quad y_{12} \quad y_{13} \quad y_{21} \quad y_{22} \quad y_{23} \quad y_{31} \quad y_{32} \quad y_{33}) \\
 YT = (y_{t_{111}} \quad y_{t_{121}} \quad y_{t_{131}} \quad y_{t_{211}} \quad y_{t_{221}} \quad y_{t_{231}} \quad y_{t_{311}} \quad y_{t_{321}} \quad y_{t_{331}} \\
 \quad y_{t_{112}} \quad y_{t_{122}} \quad y_{t_{132}} \quad y_{t_{212}} \quad y_{t_{222}} \quad y_{t_{232}} \quad y_{t_{312}} \quad y_{t_{322}} \quad y_{t_{332}}) \\
 = (y_{11} \quad y_{12} \quad y_{13} \quad y_{21} \quad y_{22} \quad y_{23} \quad y_{31} \quad y_{32} \quad y_{33} \\
 \quad y_{11} \quad y_{12} \quad y_{13} \quad y_{21} \quad y_{22} \quad y_{23} \quad y_{31} \quad y_{32} \quad y_{33}) \\
 LINE = (1 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2 \quad 3 \quad 3 \quad 3 \\
 \quad 1 \quad 2 \quad 3 \quad 1 \quad 2 \quad 3 \quad 1 \quad 2 \quad 3) \\
 DUMMY = (1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \\
 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9) .
 \end{array}$$

Then an analysis of the $p^2 \times 2$ values y_{ijk}^t with LINE as treatment factor and DUMMY as a blocking factor will use 1 in the DUMMY stratum and, hence, give the required analysis in the DUMMY stratum. As noted by Wilkinson (1971) the resulting line effects have efficiency $(p-2)/(p-1)$.

Other more complicated models might be fitted. In Appendix 1 the code is given for Hayman's (1954) analysis using data in Mather and Jinks (1971, p. 256). This analysis includes comparisons of mean parental performance with mean progeny performance (i.e. y_{ii} versus y_{ij} ($j=i$) or 'pure' versus 'cross') and of line parental performance with line cross performance and fits a model similar to (1) to the reciprocal difference $(y_{ij} - y_{ji})$.

PARTIALLY BALANCED INCOMPLETE BLOCKS WITH TRIANGULAR ASSOCIATION SCHEMES (P.B.I.B.T.)

These designs (for example Clatworthy, 1973) have $t = n(n-1)/2$ treatments. The association scheme can be obtained from a symmetric $n \times n$ array A with the diagonal blank and each treatment occurring twice - once above and once below the diagonal. Treatments in the same row or column are first associates. For example, the design with $t=10$, $b=6$, $k=5$, $r=3$, with treatments

1, 3, 7, 8, 9	in the first block
1, 2, 6, 9, 10	in the second block
1, 4, 5, 8, 10	in the third block
2, 4, 6, 7, 8	in the fourth block
2, 3, 5, 7, 10	in the fifth block and
3, 4, 5, 6, 9	in the sixth block

is a P.B.I.B.T..

The design has array A

*	1	2	3	4
1	*	5	6	7
2	5	*	8	9
3	6	8	*	10
4	7	9	10	*

and this indicates that 2, 3, 4, 5, 6 and 7 are the first associates of 1 and 3, 6, 8, 4, 7 and 9 are first associates of 10. As A is symmetric, the information on the association scheme can be found from the triangular array above (or below) the diagonal and hence the designs are said to have a triangular association scheme. The designs cannot be analysed directly by 'ANOVA' but they can be made balanced by imposing a pseudo-treatment structure. Because of the association scheme it is natural to use the rows

and columns of A to generate some pseudo-treatment structure. Suppose that treatment i (in the lower triangle of A) is in row r_i and column c_i . Then T.P. Speed has shown that P.B.I.B.T. designs have general balance with respect to the pseudo-treatment scheme

$$\begin{aligned} i\text{-th treatment effect} &= r_i\text{-th pseudo-treatment effect} \\ &+ c_i\text{-th pseudo-treatment effect} \\ &+ r_i, c_i \text{ pseudo-treatment interaction effect} \end{aligned}$$

with $(n-1)$ degrees of freedom for the pseudo-treatment effects and $(n-2)(n-1)/2$ degrees of freedom for the interaction effects.

This model is similar to the diallel model and can be fitted in a similar way. We make 2 copies of the data and use the rows of A to give the levels of the pseudofactor for the first copy and the columns of A to give the levels of the pseudofactor for the second. For example, the treatments

1 2 3 4 5 6 7 8 9 10

can be replaced by

2 3 4 5 3 4 5 4 5 5

with the first copy of the data and by

1 1 1 1 2 2 2 3 3 4

with the second copy of the data. Appendix 2 gives the Genstat code for the analysis of variance for this design, with some artificial data.

One problem is that Genstat does not know that each treatment is associated with 2 levels of the pseudofactor and it is not clear how the levels are calculated when deriving tables of means. For instance with the example above, Genstat associates levels 2, 2, 3, 3, 3, 3, 4, 4, 4, 5 with the treatments 1-10. The level for treatment i is the integer part of $[(r_i + c_i + 1)/2]$. In most other circumstances, the association of two levels of a pseudofactor with one treatment combination will only occur by mistake and it is intended to trap this in future releases of Genstat! However, correct tables of means and standard errors can easily be constructed using 'EXTRACT'.

A macro is available from the author to generate the levels of the pseudo-factors from treatment and block-factors for P.B.I.B.T. designs when A is not known, for example, if the treatments are randomised and no note is kept of the original order.

RECTANGULAR LATTICES

These designs, introduced by Harshbarger (1950) for $p(p-q)$ treatments in blocks of $(p-q)$, are a development of square lattices (Yates, 1936). Macros exist to analyse rectangular lattices but it is interesting to show how the method of the previous sections can also be used with these designs.

Consider a design for $12 = 4 \times 3$ treatments in 3 replications and block size of 3. A rectangular lattice design with these dimensions can be constructed from the following Graeco-Latin square

TP 2

	1	2	3	4
TP 1	δ 13 A	α 1 D	β 2 B	γ 3 C
	α 4 C	δ 14 B	γ 5 D	β 6 A
	β 7 D	γ 8 A	δ 15 C	α 9 B
	γ 10 B	β 11 C	α 12 A	δ 16 D

where the levels of TP3 are the Latin letters A, B, C, D and of TP4 are the Greek Letters α , β , γ , δ . We associate with each cell a treatment, the off-diagonal cells are numbered from 1 to 12, the diagonal from 13 to 16. Then, using rows, columns and treatments to generate blocks, a 3-replicate rectangular lattice design can be derived by ignoring the diagonal elements of the square. (It will become apparent later why we have bothered to label the diagonal treatments as 13 to 16.)

This is

Rep I	1	2	3	Rep II	4	7	10	Rep III	6	8	12
	4	5	6		1	8	11		2	9	10
	7	8	9		2	5	12		3	4	11
	10	11	12		3	6	9		1	5	7

This suggests using TP1 + TP2 + TP3 + TP4 as a pseudofactor structure for T and, although this allows Genstat to produce the correct analysis of variance, the treatment effects are only 'approximate'. This is because TP1, TP2 and TP3 are not mutually orthogonal.

T.P. Speed has suggested an alternative subdivision of the pseudofactor space which allows an orthogonal subdivision. In our case, rather than think of TP1, TP2, TP3 as representing 3 sets of 4 effects, we think of them as a two-way table indexed by COPY with 3 levels and TP with 4 levels, i.e.

		<i>TP1</i>	<i>TP2</i>	<i>TP3</i>
<i>COPY</i>		1	2	3
	1	1	1	A
<i>TP</i>	2	2	2	B
	3	3	3	C
	4	4	4	D

Speed has shown that *TP* and *TP.COPY* generate orthogonal subspaces (with 3 degrees of freedom and 6 degrees of freedom in our case). Note that the combinations of the pseudofactors *TP1*, *TP2*, *TP3* at level *i* of *TP* relate to the *i*-th diagonal of the Graeco-Latin square, so that the average effects of *TP* would give estimates to compare treatments 13 to 16 if the pseudo-factorial model was appropriate.

Hence, each treatment is now associated with 3 levels of *TP* and *TP.COPY*. As before, we can specify this model using 3 copies of the data and using *TPi* to give the level of *TP* with the *i*th copy of the data. Appendix 3 gives the Genstat code for a rectangular lattice taken from Kempthorne (1952, p.522). Again, there is a problem linking up treatment estimates with the 3 levels of *TP* and *COPY*.

DIFFICULTIES

There are three disadvantages of this method (i) multiple copies of, for instance, the yield variate are needed, (ii) effects and sums of squares in the dummy stratum are calculated, (iii) the linkage between treatments and pseudofactors is not complete.

G.N. Wilkinson has suggested that a simple modification of the algorithm should avoid at least the first two of the difficulties. He would allow the treatment factor length to be a multiple, say *c*, of the length of the analysed variate. When treatment totals are calculated, the program would go through the data *c* times and when sweeps are made, *c* terms would be taken from each data value. A similar technique could be used for deriving treatment estimates from pseudofactors.

REFERENCES

- Clatworthy, W.H. (1973) Tables of two-associate-class partially balanced designs (US Dept of Commerce).
- Griffing, B. (1956) A generalised treatment of the use of diallel crosses in quantitative inheritance. *Heredity*, 10, 31-50.
- Harshbarger, B. (1950) Triple rectangular lattices. *Biometrics*, 5, 1-13.

- Hayman, B.I. (1954) The analysis of variance of diallel tables. *Biometrics*, 10, 235-244.
- Jinks, J.L. and Mather, K. (1971) *Biometric genetics*. (Chapman and Hall).
- Kempthorne, O. (1952) *The design and analysis of experiments*. J. Wiley and Sons, New York.
- Thompson, R. and Baker, R.J. (1981) Composite link functions in generalised linear models. *J.R.S.S.*, 30, 123-131.
- Wilkinson, G.N. (1970) A general recursive algorithm for analysis of variance. *Biometrika*, 57, 19-46.
- Yates, F. (1936) A new method of arranging variety trials involving a large number of varieties. *Journal of Agricultural Science*, 26, 424-455.
- Yates, F. (1947) The analysis of dates from all possible reciprocal crosses between a set of parental lines. *Heredity*, 1, 287-301.

APPENDIX 1

```
'REFERENCE' DIALLEL
'CAPTION' ''
  P is the number of parental lines
  Sets up size parameters
  I is the index for male parent,
  J is the index for female parent,
  D is the index for copies of data,
  variates, factors, and calculations
  for PARENTS, Pure Versus Crosses,
  RECIPricals, DUMMY.
  Crosses (i,j) and (j,i) are at level
   $i \times (i-1)/2 + j$  of RECIP.
..
'SET' P=9 'SCALAR' N,N1,N2
'CALC' N=2*P*P : N1=P*P : N2=P*(P-1)/2
'RUN'
'UNITS' UNSN=1...N
'VARIATE' I=(P!(1...P))2 : J=((1...P)P)2 : D=N1!(1,2)
'VARIATE' PARV,PVCV,RECIPV,DUMV
'FACTOR' PARENTSSP : PVC$2 : RECIP$N2 : DUMMYSN1 : COPYS2
'CALC' PARV=I*(D.EQ.1) + J*(D.EQ.2)
'CALC' PVCV=(I.EQ.J) + 2*(I.NE.J)
'CALC' RECIPV=(I.GE.J)*(I*(I-1)/2 + J) + (I.LT.J)*(J*(J-1)/2 + I)
'CALC' DUMV=(I-1)*P + J
'GROUP' PARENTS,PVC,RECIP,DUMMY,COPY=INTPT(PARV,PVCV,RECIPV,DUMV,D)
'HEADING' H = ''
  THIS GIVES FOUR TIMES THE NUMBERS IN JINKS AND MATHER (page 256) ''
'DESCRIBE' YT $ ; H
'BLOCKS' RECIP/DUMMY
'TREATMENTS' PARENTS*PVC + PARENTS*COPY
'VARIATE' Y$N1 'VARIATE' YT
'READ/PRIN=DEM' Y 'EQUATE' YT=Y
'DESCIBE' YT$5
'ANOVA/PR=00010' YT
'RUN'
```

77.8 53.4 79.6 69.6 50.2 59.6 71.4 67.6 50.6
 47.8 54.1 50.0 46.2 43.0 52.4 46.8 41.2 40.4
 68.8 53.2 97.6 59.1 50.0 63.0 72.2 48.8 52.0
 72.2 47.0 62.4 68.2 46.8 58.7 54.4 44.6 50.0
 53.0 46.4 52.0 51.0 53.2 55.0 54.4 40.4 48.4
 56.8 48.2 60.6 63.8 48.3 54.0 55.4 44.8 49.6
 73.8 49.4 83.6 67.8 60.2 59.6 74.0 48.8 58.2
 53.6 38.6 55.6 44.2 38.4 37.6 45.4 30.6 43.6
 50.6 46.6 49.8 48.0 45.0 42.6 54.8 38.0 50.8

'EOD'

Y MNMINMAX 54.2616 30.6000 97.6000 81 VALUES 0 MISSING

1

***** ANALYSIS OF VARIANCE *****

VARIATE: YT

THIS GIVES FOUR TIMES THE NUMBERS IN JINKS AND MATHER (page 256)

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
RECIP STRATUM					
PARENTS	8	1.582E 4	72.85	1.978E 3	23.238
PVC	1	1.294E 3	5.96	1.294E 3	15.203
PARENTS.PVC	8	1.146E 3	5.27	1.432E 2	1.683
RESIDUAL	27	2.298E 3	10.58	8.511E 1	
TOTAL	44	2.056E 4	94.66	4.673E 2	
RECIP.DUMMY STRATUM					
PARENTS.COPY	8	5.420E 2	2.50	6.775E 1	3.068
RESIDUAL	28	6.183E 2	2.85	2.208E 1	
TOTAL	36	1.160E 3	5.34	3.223E 1	
RECIP.DUMMY.*UNITS* STRATUM					
PARENTS	8	0.000E 0	0.00	0.000E 0	
COPY	1	0.000E 0	0.00	0.000E 0	
PARENTS.COPY	8	0.000E 0	0.00	0.000E 0	
RESIDUAL	64	0.000E 0	0.00	0.000E 0	
TOTAL	81	0.000E 0	0.00	0.000E 0	
GRAND TOTAL	161	2.172E 4	100.00		
GRAND MEAN		54.26172			
TOTAL NUMBER OF OBSERVATIONS		162			

***** INFORMATION SUMMARY *****

MODEL TERM EF NON-ORTHOGONAL TERMS

RECIP STRATUM		
PARENTS	0.500	
PARENTS.PVC	0.875	PARENTS
RECIP.DUMMY STRATUM		
PARENTS.COPY	0.500	
RECIP.DUMMY.*UNITS* STRATUM		
PARENTS	0.500	RECIP
PARENTS.COPY	0.500	RECIP.DUMMY

'CLOSE'

APPENDIX 2

```
'REFERENCE'      ROW_COLUMN_ASSOCIATION
''
This file is RCA
''
'CAPTION'       ''
The row and column associations are now calculated by grouping the levels
of factor T ''
'UNITS'        $ 60
'INTEGER'      INT1 = -11,-1,2,-5,3,6,-8,4,7,9,-10
                INT2 = 1,2,3,-4,5,6,-7,8,-9,-10,-11
```

```
'FACTOR'      R $ 5
:             C $ 5
:             T $ 10 = 3(1),2(2),2(3),2,2(4),3,4,7,(6,5)2,
:             5,8,9,8,2(7),6,9,2(10),8,10,9
:             DUMMY $ 30 = (1...30)2
:             BLOCK $ 6 = (1...6)10
'RUN'
```

```
'GROUP'       R = GROUP(T ; INT1)
:             C = GROUP(T ; INT2)
'RUN'
```

```
'FACTOR'      TREAT $ 10 = T,T
:             PF1 $ 5 = R,C
'VARIATE'     Y = (3(-3),2(-2,-1),-2,2(0),-1,0,1,0,-1,0,2(-1),
:             1,2,3(1),0,2,2(3),1,3,2)2
:             BLCKT = (-5,-3,-1,1,3,5)10
:             INTERACT = (5(0),2(2),0,2(-2),(2,-2)2,0,-2,
:             5(0),2(2),-2,6(0))2
'PRINT/P'    DUMMY,BLOCK,TREAT,PF1,Y,BLCKT,INTERACT $ 10
'RUN'
```

DUMMY	BLOCK	TREAT	PF1	Y	BLCKT	INTERACT
1	1	1	2	-3	-5	0
2	2	1	2	-3	-3	0
3	3	1	2	-3	-1	0
4	4	2	3	-2	1	0
5	5	2	3	-2	3	0
6	6	3	4	-1	5	2
7	1	3	4	-1	-5	2
8	2	2	3	-2	-3	0
9	3	4	5	0	-1	-2
10	4	4	5	0	1	-2
11	5	3	4	-1	3	2
12	6	4	5	0	5	-2
13	1	7	5	1	-5	2
14	2	6	4	0	-3	-2
15	3	5	3	-1	-1	0
16	4	6	4	0	1	-2
17	5	5	3	-1	3	0
18	6	5	3	-1	5	0
19	1	8	4	1	-5	0
20	2	9	5	2	-3	0
21	3	8	4	1	-1	0
22	4	7	5	1	1	2
23	5	7	5	1	3	2
24	6	6	4	0	5	-2
25	1	9	5	2	-5	0
26	2	10	5	3	-3	0
27	3	10	5	3	-1	0
28	4	8	4	1	1	0
29	5	10	5	3	3	0
30	6	9	5	2	5	0
1	1	1	1	-3	-5	0
2	2	1	1	-3	-3	0
3	3	1	1	-3	-1	0
4	4	2	1	-2	1	0
5	5	2	1	-2	3	0
6	6	3	1	-1	5	2
7	1	3	1	-1	-5	2
8	2	2	1	-2	-3	0
9	3	4	1	0	-1	-2
10	4	4	1	0	1	-2

11	5	3	1	-1	3	2
12	6	4	1	0	5	-2
13	1	7	2	1	-5	2
14	2	6	2	0	-3	-2
15	3	5	2	-1	-1	0
16	4	6	2	0	1	-2
17	5	5	2	-1	3	0
18	6	5	2	-1	5	0
19	1	8	3	1	-5	0
20	2	9	3	2	-3	0
21	3	8	3	1	-1	0
22	4	7	2	1	1	2
23	5	7	2	1	3	2
24	6	6	2	0	5	-2
25	1	9	3	2	-5	0
26	2	10	4	3	-3	0
27	3	10	4	3	-1	0
28	4	8	3	1	1	0
29	5	10	4	3	3	0
30	6	9	3	2	5	0

```
'BLOCKS'          BLOCK/DUMMY
'TREATMENT'       TREAT//PF1
'CALCULATE'       Y = Y + BLCKT
                  Y = Y + INTERACT
:
'ANOVA/PR=0.3013' Y
'RUN'
```

1

***** ANALYSIS OF VARIANCE *****

VARIATE: Y

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
BLOCK STRATUM					
TREAT	5	6.712E 2	72.33	1.342E 2	2
TOTAL	5	6.712E 2	72.33	1.342E 2	2
BLOCK.DUMMY STRATUM					
TREAT	9	2.568E 2	27.67	2.853E 1	1
RESIDUAL	15	0.000E 0	0.00	0.000E 0	0
TOTAL	24	2.568E 2	27.67	1.070E 1	1
BLOCK.DUMMY.*UNITS* STRATUM					
TREAT	4	0.000E 0	0.00	0.000E 0	0
RESIDUAL	26	0.000E 0	0.00	0.000E 0	0
TOTAL	30	0.000E 0	0.00	0.000E 0	0
GRAND TOTAL	59	9.280E 2	100.00		
GRAND MEAN		0.00			
TOTAL NUMBER OF OBSERVATIONS	60				

***** INFORMATION SUMMARY *****

MODEL TERM	EF	NON-ORTHOGONAL TERMS
BLOCK STRATUM		
TREAT	0.200	
BLOCK.DUMMY STRATUM		
PF1	0.375	
TREAT	0.800	BLOCK
BLOCK.DUMMY.*UNITS* STRATUM		
PF1	0.625	BLOCK.DUMMY

***** TABLES OF EFFECTS *****

VARIATE: Y

*** BLOCK STRATUM ***

TREAT	EFFECTS:					REP	6	ESE	*		
TREAT	1	2	3	4	5	6	7	8	9	10	
	-15.00	1.67	7.00	6.33	11.67	3.00	0.33	-8.33	-5.00	-1.67	

*** BLOCK.DUMMY STRATUM ***

PF1	EFFECTS:					REP	12	ESE	0.000		
PF1	1	2	3	4	5						
	-4.00	-2.00	0.00	2.00	4.00						

TREAT	EFFECTS:					REP	6	ESE	0.000		
TREAT	1	2	3	4	5	6	7	8	9	10	
	0.00	0.00	2.00	-2.00	-0.00	-2.00	2.00	-0.00	-0.00	-0.00	

*** BLOCK.DUMMY.*UNITS* STRATUM ***

PF1	EFFECTS:					REP	12	ESE	0.000		
PF1	1	2	3	4	5						
	0.00	0.00	0.00	0.00	0.00						

***** TABLES OF MEANS *****

VARIATE: Y

GRAND MEAN	0.00									
TREAT	1	2	3	4	5	6	7	8	9	10
PF1	2	2	3	3	3	3	4	4	4	5
	0.00	0.00	2.00	-2.00	-0.00	-2.00	2.00	-0.00	-0.00	-0.00

***** STANDARD ERRORS OF DIFFERENCES OF MEANS *****

TABLE	TREAT
REP	6
SED	0.000

***** STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION *****

STRATUM	DF	SE	CV%
BLOCK	0	*	*
BLOCK.DUMMY	15	0.000	*
BLOCK.DUMMY.*UNITS*	26	0.000	*

'CLOSE'

APPENDIX 3

```
'REFERENCE/NID=100,NUNN=100'      TRIPLE_RL
'CAPTION'      ''
This program uses an example of a triple rectangular lattice taken from
'The Design and Analysis of Experiments' by Kempthorne (page 522).
''
'LINES'      10
'UNITS'      $ 36
'FACTOR'      R $ 3 = 12(1...3)
              :      T4 $ 3
              :      B,T1,T2,T3 $ 4
              :      T $ 12
'READ/PRIN=DE,FORM=P'      Y,T,T1,T2,T3,B,T4
'RUN'
8.9  2 1 3 2 1 2
10.0 3 1 4 3 1 3
11.6 1 1 2 4 1 1
9.4  7 3 1 4 3 2
9.3  9 3 4 2 3 1
10.2 8 3 2 1 3 3
9.6  4 2 1 3 2 1
11.4 5 2 3 4 2 3
10.0 6 2 4 1 2 2
11.8 10 4 1 2 4 3
11.7 12 4 3 1 4 1
13.1 11 4 2 3 4 2
11.6 2 1 3 2 3 2
12.4 12 4 3 1 3 1
10.0 5 2 3 4 3 3
9.6  7 3 1 4 1 2
10.0 10 4 1 2 1 3
8.4  4 2 1 3 1 1
10.5 1 1 2 4 2 1
11.0 8 3 2 1 2 3
10.0 11 4 2 3 2 2
11.5 3 1 4 3 4 3
12.6 9 3 4 2 4 1
11.8 6 2 4 1 4 2
11.2 2 1 3 2 2 2
12.1 9 3 4 2 2 1
9.7  10 4 1 2 2 3
7.8  3 1 4 3 3 3
9.6  11 4 2 3 3 2
10.6 4 2 1 3 3 1
8.7  12 4 3 1 1 1
9.1  8 3 2 1 1 3
8.0  6 2 4 1 1 2
8.7  1 1 2 4 4 1
7.5  7 3 1 4 4 2
8.3  5 2 3 4 4 3
'EOD'
'PRINT/P'      Y,R,B,T,T1,T2,T3,T4 $ 10.2,7(10)
'LINES'      10
'BLOCKS'      R/R
'TREATMENTS'   T/((T1+T2+T3+T4)
'CAPTION'      ''
This analysis uses the data and gives an approximate table of means ''
'ANOVA'      Y
'RUN'
```

Y	R	B	T	T1	T2	T3	T4
8.90	1	1	2	1	3	2	2
10.00	1	1	3	1	4	3	3
11.60	1	1	1	1	2	4	1
9.40	1	3	7	3	1	4	2
9.30	1	3	9	3	4	2	1
10.20	1	3	8	3	2	1	3
9.60	1	2	4	2	1	3	1
11.40	1	2	5	2	3	4	3
10.00	1	2	6	2	4	1	2
11.80	1	4	10	4	1	2	3
11.70	1	4	12	4	3	1	1
13.10	1	4	11	4	2	3	2
11.60	2	3	2	1	3	2	2
12.40	2	3	12	4	3	1	1
10.00	2	3	5	2	3	4	3
9.60	2	1	7	3	1	4	2
10.00	2	1	10	4	1	2	3
8.40	2	1	4	2	1	3	1
10.50	2	2	1	1	2	4	1
11.00	2	2	8	3	2	1	3
10.00	2	2	11	4	2	3	2
11.50	2	4	3	1	4	3	3
12.60	2	4	9	3	4	2	1
11.80	2	4	6	2	4	1	2
11.20	3	2	2	1	3	2	2
12.10	3	2	9	3	4	2	1
9.70	3	2	10	4	1	2	3
7.80	3	3	3	1	4	3	3
9.60	3	3	11	4	2	3	2
10.60	3	3	4	2	1	3	1
8.70	3	1	12	4	3	1	1
9.10	3	1	8	3	2	1	3
8.00	3	1	6	2	4	1	2
8.70	3	4	1	1	2	4	1
7.50	3	4	7	3	1	4	2
8.30	3	4	5	2	3	4	3

This analysis uses the data and gives an approximate table of means

**** ANALYSIS OF VARIANCE ****

VARIATE: Y

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
STRATUM	2	16.1072	21.75	8.0536	8.329
.B STRATUM					
T	9	36.9025	49.84	4.1003	
TOTAL	9	36.9025	49.84	4.1003	4.240
.R.*UNITS* STRATUM					
T	11	8.4630	11.43	0.7694	0.796
RESIDUAL	13	12.5703	16.98	0.9669	
TOTAL	24	21.0333	28.41	0.8764	
RAND TOTAL	35	74.0430	100.00		
RAND MEAN		10.21			
TOTAL NUMBER OF OBSERVATIONS		36			

**** INFORMATION SUMMARY ****

DEL TERM	EF	NON-ORTHOGONAL TERMS
.B STRATUM		
T1	0.407	
T2	0.323	T1
T3	0.099	T1 T2

R.B.*UNITS* STRATUM

T1	0.593	R.B	
T2	0.556	R.B	T1
T3	0.494	R.B	T1 T2

ALIASED MODEL TERMS

T

***** TABLES OF MEANS *****

VARIATE: Y

GRAND MEAN 10.21

*** FOLLOWING TABLE OF MEANS IS ONLY APPROXIMATE ***

T	1	2	3	4	5	6	7	8	9	10	11
T1	1	1	1	2	2	2	3	3	3	4	4
T2	2	3	4	1	3	4	1	2	4	1	2
T3	4	2	3	3	4	1	4	1	2	2	3
T4	1	2	3	1	3	2	2	3	1	3	2
	11.24	10.03	9.25	9.47	9.98	9.42	10.18	11.29	10.69	9.75	10.60
T	12										
T1	4										
T2	3										
T3	1										
T4	1										
	10.66										

***** STANDARD ERRORS OF DIFFERENCES OF MEANS *****

TABLE	T
-----	-----
REP	3
SED	1.160
EXCEPT WHEN COMPARING MEANS WITH SAME LEVEL(S) OF:	
T1	0.991
T2	0.979
T3	0.954
T4	1.088

***** STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION *****

STRATUM	DF	SE	CV%
R	2	0.819	8.0
R.B	0	*	*
R.B.*UNITS*	13	0.983	9.6

```

'UNITS'          UNIT $ 108 = 1...108
'VARIATE'        NY = (Y)3
'FACTOR'         NTP $ 4 = T1,T2,T3
                  : COPY $ 3 = 36(1...3)
                  : NT4 $ 3 = (T4)3
                  : DUMMY $ 36 = (1...36)3
                  : NT $ 12 = (T)3
                  : NR $ 3 = (R)3
                  : NB $ 4 = (B)3
'DEVALUE'        Y,T1,T2,T3,T,R,B
'DESCRIBE'       NY $ 3
'LINES'          10
'TREATMENT'      NT//((NTP*COPY+NT4)
'BLOCKS'         NR/NB/DUMMY
'CAPTION'        ''

```

This analysis uses copies of the data and in the analysis of variance table gives three times the numbers in the above analysis ''

'ANOVA/PR=03013' NY

'RUN'

This analysis uses copies of the data and in the analysis of variance table gives three times the numbers in the above analysis

***** ANALYSIS OF VARIANCE *****

VARIATE: NY

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
NR STRATUM	2	4.832E	1	21.75	2.416E 1
NR.NB STRATUM					
NT	9	1.107E	2	49.84	1.230E 1
TOTAL	9	1.107E	2	49.84	1.230E 1
NR.NB.DUMMY STRATUM					
NT	11	2.539E	1	11.43	2.308E 0
RESIDUAL	13	3.771E	1	16.98	2.901E 0
TOTAL	24	6.310E	1	28.41	2.629E 0
NR.NB.DUMMY.*UNITS* STRATUM					
NT	11	0.000E	0	0.00	0.000E 0
RESIDUAL	61	0.000E	0	0.00	0.000E 0
TOTAL	72	0.000E	0	0.00	0.000E 0
GRAND TOTAL	107	2.221E	2	100.00	
GRAND MEAN		10.214			
TOTAL NUMBER OF OBSERVATIONS		108			

***** INFORMATION SUMMARY *****

MODEL TERM EF NON-ORTHOGONAL TERMS

NR.NB STRATUM

NTP	0.012
NTP.COPY	0.198

NR.NB.DUMMY STRATUM

NTP	0.099	NR.NB
NTP.COPY	0.247	NR.NB

NR.NB.DUMMY.*UNITS* STRATUM

NTP	0.889	NR.NB	NR.NB.DUMMY
NTP.COPY	0.556	NR.NB	NR.NB.DUMMY

ALIASED MODEL TERMS

NT

***** TABLES OF EFFECTS *****

VARIATE: NY

*** NR.NB STRATUM ***

NTP	EFFECTS:				REP	27	ESE
	1	2	3	4			
NTP	-7.625	3.575	-1.025	5.075			

NTP.COPY	EFFECTS:				REP	9	ESE
	1	2	3				
COPY							
NTP							
1	0.969	-1.356	0.388				
2	-1.456	-1.531	2.988				
3	-1.881	1.494	0.388				
4	2.369	1.394	-3.762				

** NR.NB.DUMMY STRATUM ***

NTP EFFECTS: REP 27 ESE 1.0430

NTP	1	2	3	4
	-0.725	0.800	0.062	-0.137

NTP.COPY EFFECTS: REP 9 ESE 1.1425

COPY	1	2	3
NTP			
1	-0.160	-0.640	0.800
2	-1.055	1.570	-0.515
3	1.025	-0.145	-0.880
4	0.190	-0.785	0.595

* NR.NB.DUMMY.*UNITS* STRATUM ***

NTP EFFECTS: REP 27 ESE 0.0000

NTP	1	2	3	4
	0.000	0.000	0.000	0.000

COPY EFFECTS: REP 36 ESE 0.0000

COPY	1	2	3
	0.000	0.000	0.000

TP.COPY EFFECTS: REP 9 ESE 0.0000

COPY	1	2	3
NTP			
1	0.000	0.000	0.000
2	0.000	0.000	0.000
3	0.000	0.000	0.000
4	0.000	0.000	0.000

***** TABLES OF MEANS *****

VARIATE: NY

GRAND MEAN 10.214

NT	1	2	3	4	5	6	7	8	9	10	11
NTP	2	2	3	2	3	2	3	2	3	2	3
COPY	2	2	2	2	2	2	2	2	2	2	2
NT4	1	2	3	1	3	2	2	3	1	3	2
	10.517	10.058	10.067	10.517	10.067	10.058	10.058	10.067	10.517	10.067	10.058

NT 12

NTP 3

COPY 2

NT4 1

10.517

***** STANDARD ERRORS OF DIFFERENCES OF MEANS *****

TABLE	NT
REP	9
SED	0.4014

***** STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION *****

STRATUM	DF	SE	CV%
NR	2	0.8192	8.0
NR.NB	0	*	*
NR.NB.DUMMY	13	0.9833	9.6
NR.NB.DUMMY.*UNITS*	61	0.0000	0.0

'CLOSE'

*R Thompson
Agricultural Research Council
Unit of Statistics
University of Edinburgh
James Clerk Maxwell Building
The King's Buildings
Mayfield Road
Edinburgh EH9 3JZ
U.K.*

A REGRESSION MODEL FOR GENOTYPICAL STABILITY

INTRODUCTION

Yates and Cochran (1938) have described a model for relating the performance of a set of genotypes across a set of environments. The model has been used in studies of crop variety performance (Finlay and Wilkinson, 1963, Eberhart and Russell, 1966) and also in the evaluation of variety-isolate interactions in studies of disease resistance and pathogen virulence (Leonard and Moll, 1979, Jenks and Leonard, 1981).

The model takes the form

$$Y_{ij} = g_i + b_i I_j + d_{ij}$$

where Y_{ij} is response of the i th genotype grown in the j th environment,
 g_i is the mean of the i th genotype,
 b_i is the coefficient of regression for the response of the i th genotype on I_j , an environmental index, and
 d_{ij} is deviation from regression.

The environmental index is variously taken as the mean response of all genotypes within an environment or as the mean response of a specific genotype within an environment. Generally, the coefficients, b_i , are used to characterise stability, or general adaptation, of the set of genotypes over the series of environments that are covered by the test. For further details on interpretation of the model, the reader is referred to the literature.

MACROS

The model is readily fitted by Genstat. Three Macros are used.

ERINDEX computes the environmental indices. There are four input parameters. *Y* is the response variate, *GEN* and *ENV* are factors for genotype and environment respectively. *INDEX* is a scalar taking a value of zero if environmental means are to be used for indexing. If a specific genotype is to be used as the index, then *INDEX* must be set to the formal level of that genotype. *ERINDEX* must be used before either of the following macros.

ERFIT partitions variation into genotypical effects, a general linear environmental effect, specific linear effects due to the interaction of individual genotypes with the linear effect of environments and deviations from regression.

ERPLOT may be used to produce graphical displays and will also perform a test for linearity on each genotype.

EXAMPLE

The following example was taken from Leonard and Moll (1979) who used the data of Clifford and Clothier (1974). Four varieties of barley were inoculated with each of eighteen cultures of leaf rust. The rust cultures were treated as separate environments. Output comprises a table of the environmental indices, a regression analysis for the entire set of data and graphs of response versus index for each genotype. Note that the residual sum of squares for the maximal model was zero for this particular analysis as a consequence of the analysis being conducted on the published two-way table of means rather than on replicated data.

THE MACROS

```
'MACRO' ERINDEX $
'' ARGUMENTS IN: GEN,ENV,Y,INDEX OUT: EINDEX,IMEANS''
'LOCA' TINDEX,L1
'SCAL' L1
'JUMP' L1*(INDEX.EQ.0)
'REST' Y $ GEN=INDEX
'LABE' L1 'TABL' TINDEX $ ENV 'TABU' VARIATE=Y ; MEANS=TINDEX
'PRIN' TINDEX
'REST' Y
'VARI' IMEANS $ ENV 'EQUATE' IMEANS=TINDEX
'CALC' EINDEX=ELEM(IMEANS;FLOAT(ENV))
'DEVA' TINDEX
'ENDMACRO'
```

```
'MACRO' ERFIT $
'' ARGUMENTS IN: GEN,ENV,Y,EINDEX''
'CAPT' ''*** STABILITY ANALYSIS ***''
'TERM/DVSET=F' Y+GEN*EINDEX+GEN/ENV 'Y' Y
'FIT/Z,ANDEV=I' GEN
'ADD/Z' GEN.ENV
'DROP/Z' GEN.ENV
'ADD/Z' EINDEX*GEN
'ADD/Z,ANDEV=T' GEN.ENV
'DROP/C,INT=N,' EINDEX+GEN.ENV
'ENDMACRO'
```

```

'MACRO' ERPLOT $
' ARGUMENTS IN: Y,EINDEX,IMEANS,GEN,GLEV,ENV''
'LOCA' HEAD1,HEAD2,HEAD3,HEAD4,T1,YMEANS,VR,WDF,WSQ,DDF,DSQ
'SCAL' VR 'TABLE' T1 $ ENV 'VARI' YMEANS $ ENV
'HEAD' HEAD1='GENOTYPE''
: HEAD2='F-STATISTIC FOR NON-LINEARITY='''
: HEAD3=''' DF=''' : HEAD4='MEAN='''
'FOR' I=1...GLEV
' REST' Y $ GEN=I
' TABU' VARIATE=Y ; MEANS=T1 'EQUATE' YMEANS=T1
' GRAPH/NCF=80,NRF=30' YMEANS ; IMEANS
' TERM' Y+EINDEX+ENV 'Y' Y
' FIT/Z' ENV ; DF=WDF ; DEV=WSQ ; EINDEX ; DF=DDF ; DEV=DSQ
' CALC' DDF=DDF-WDF : VR=(DSQ-WSQ)/DDF*(WDF/WSQ)
' PRIN/C,LABR=1' HEAD1,I $ 9,3,/
: HEAD2,VR,HEAD3,DDF,WDF $ 30,7.2,5,5,5,/
' CALC' VR=MEAN(Y) 'PRIN/C,LABR=1' HEAD4,VR $ 5,8.3
'REPEAT'
'ENDMACRO'

```

EXAMPLE RUN

```

'UNIT' $ 72
'VARI' Y=36.0,30.0,29.4,32.5,33.6,43.8,39.1,41.2,41.4,32.4,47.7,45.3,
57.8,51.0,65.3,63.9,37.0,46.5,54.0,55.8,29.0,25.0,44.1,27.9,
42.2,46.1,28.1,23.7,19.2,20.9,21.2,23.6,59.4,41.4,49.9,40.7,
66.1,47.5,67.3,42.6,47.6,44.1,53.9,45.9,50.8,50.7,51.9,41.7,
35.8,29.6,41.4,13.5,38.2,44.7,48.4,25.2,36.3,46.5,46.7,28.2,
19.6,19.9,21.3,15.2,60.1,40.6,64.2,42.9,54.7,44.7,68.9,32.2
'SCAL' GLEV=4 : ELEV=18 : INDEX=0
'FACT' GEN $ GLEV : ENV $ ELEV
'GENE' ENV,GEN
'USE/R' ERINDEX $
'USE/R' ERFIT $
'USE/R' ERPLOT $
'RUN'

```

	TINDEX									
ENV	1	2	3	4	5	6	7	8	9	10
	3.1975E 1	3.9425E 1	4.1700E 1	5.9500E 1	4.8325E 1					
ENV	3.1500E 1	3.5025E 1	2.1225E 1	4.7850E 1	5.5875E 1					
ENV	4.7875E 1	4.8775E 1	3.0075E 1	3.9125E 1	3.9425E 1					
ENV	1.9000E 1	5.1950E 1	5.0125E 1							

*** STABILITY ANALYSIS ***

*** SUMMARY ANALYSIS OF VARIANCE ***

Y-VARIATE: Y

	RESIDUAL		CHANGE		MEAN VARIANCE	
	DF	SS	DF	SS	CHANGE	RATIO
INITIAL MODEL						
CONSTANT	71	12882.79	*	*		
MODIFICATIONS TO MODEL						
+GEN	68	11661.80	3	1220.99	407.00	9.76
+GEN.ENV	0	-0.00	68	11661.80	171.50	4.11
-GEN.ENV	68	11661.80	-68	-11661.80	171.50	4.11
+EINDEX	67	2918.92	1	8742.88	8742.88	209.61
+EINDEX.GEN	64	2669.44	3	249.48	83.16	1.99
**DENOMINATOR OF RATIO IS RES. SS /RES.DF FROM LINE ABOVE,=					41.71	
+GEN.ENV	0	0.00	64	2669.44	41.71	1.00

```

*** WHILE DROPPING TERM EINDEX
REFITTED 1 PARAMETER(S) OF TERM EINDEX.GEN
*** WHILE DROPPING TERM CONSTANT
REFITTED 1 PARAMETER(S) OF TERM GEN

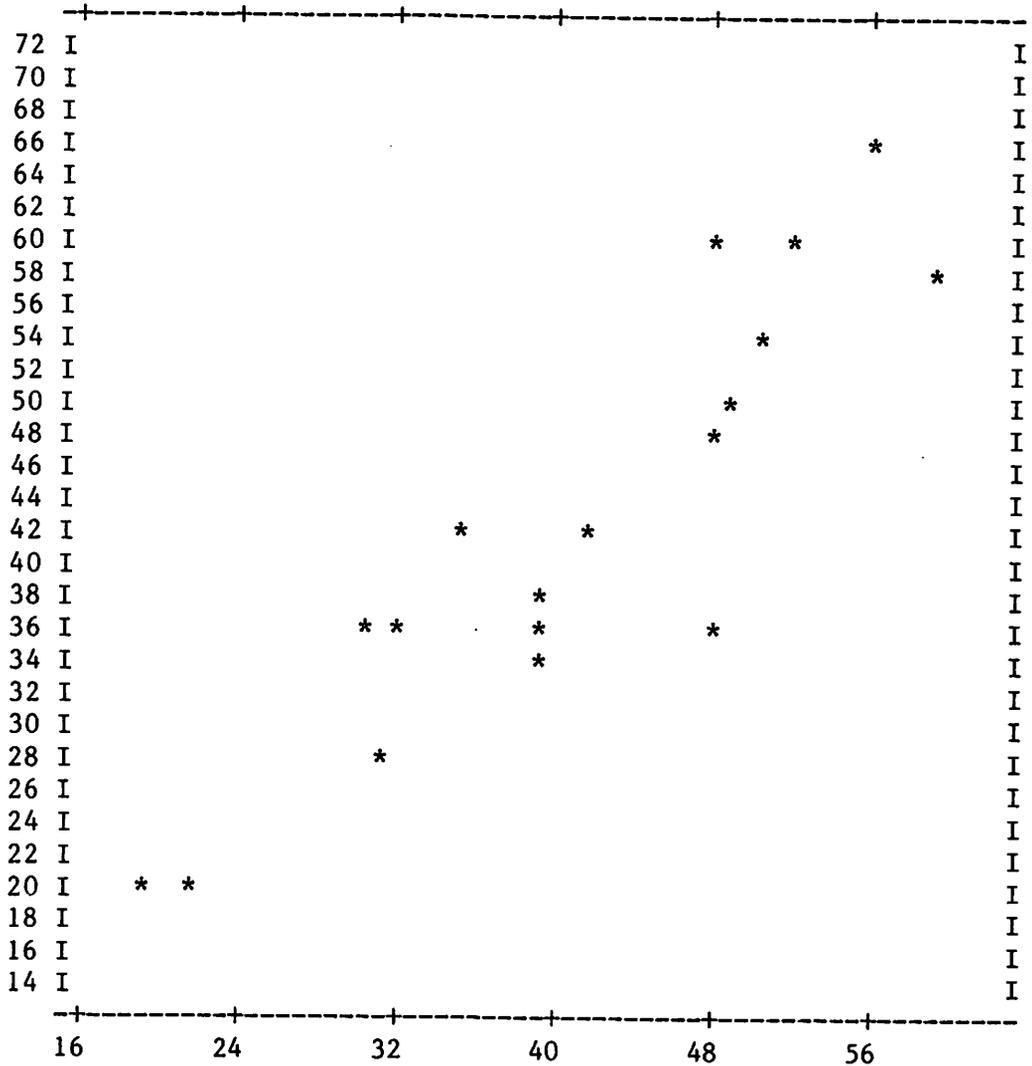
```

***** REGRESSION ANALYSIS *****

*** REGRESSION COEFFICIENTS ***

Y-VARLATE: Y

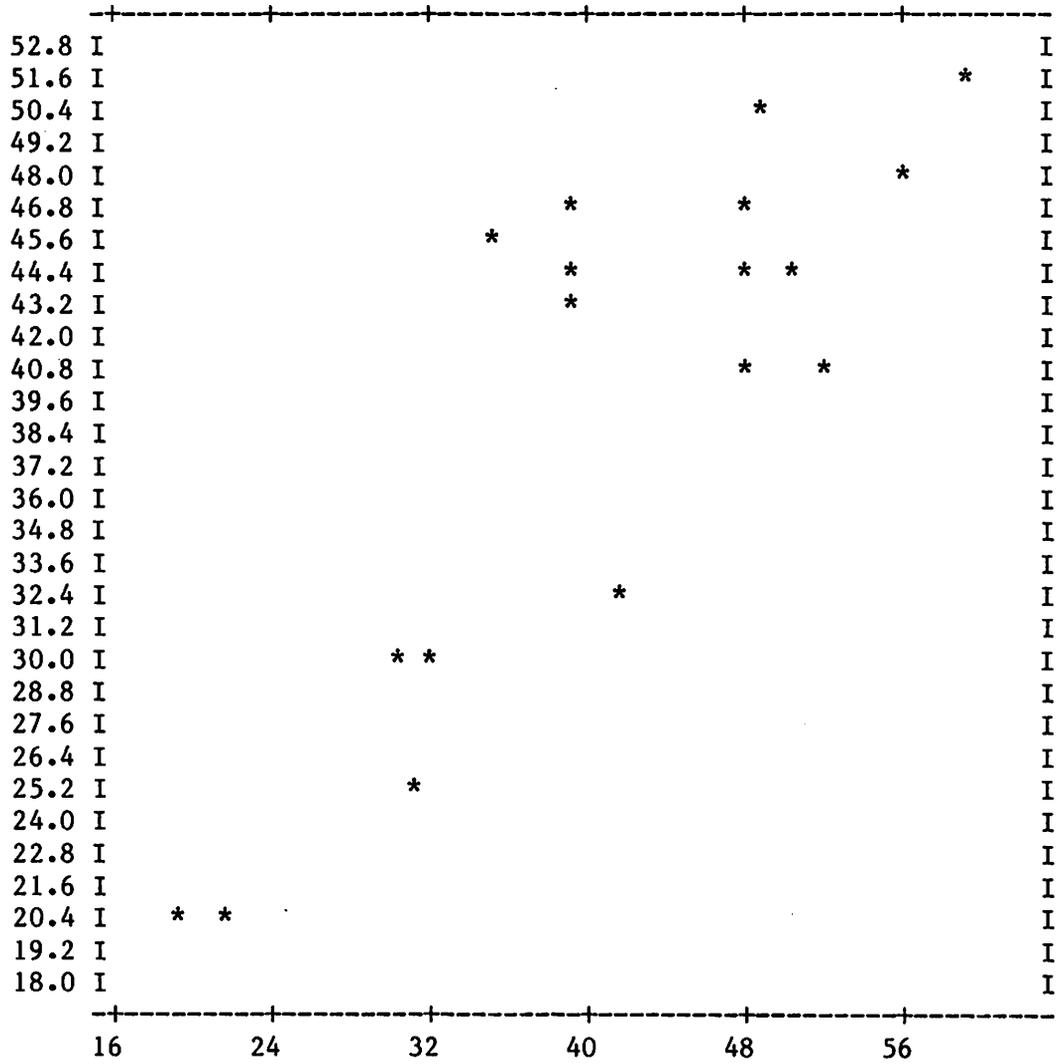
	ESTIMATE	S.E.	T
GEN 1	-2.02730	5.87034	-0.35
GEN 2	8.52237	5.87034	1.45
GEN 3	-2.60626	5.87034	-0.44
GEN 4	-3.88881	5.87034	-0.66
EINDEX.GEN 1	1.08466	0.13814	7.85
EINDEX.GEN 2	0.74720	0.13814	5.41
EINDEX.GEN 3	1.20435	0.13814	8.72
EINDEX.GEN 4	0.96379	0.13814	6.98



GENOTYPE 1

F-STATISTIC FOR NON-LINEARITY= * DF= 16 0

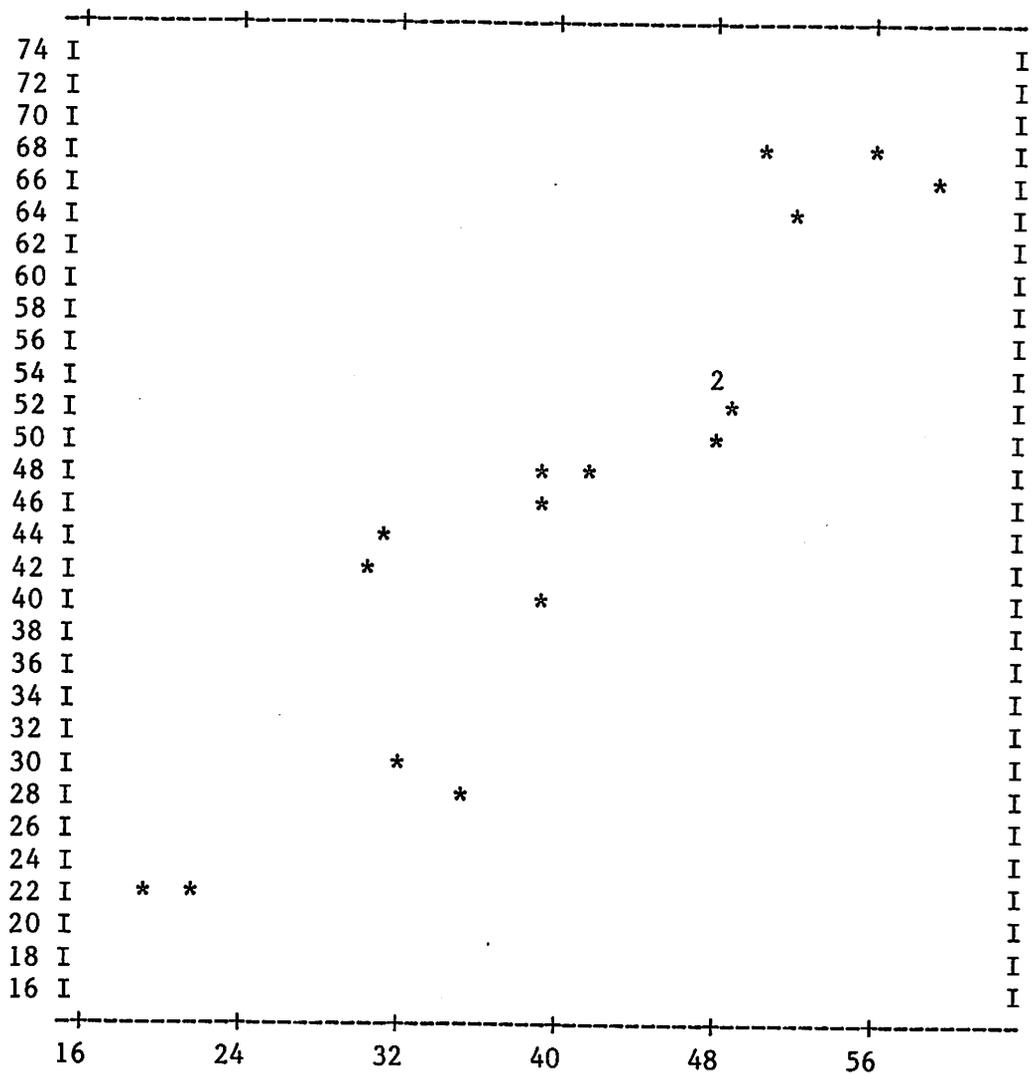
MEAN= 42.489



GENOTYPE 2

F-STATISTIC FOR NON-LINEARITY= * DF= 16 0

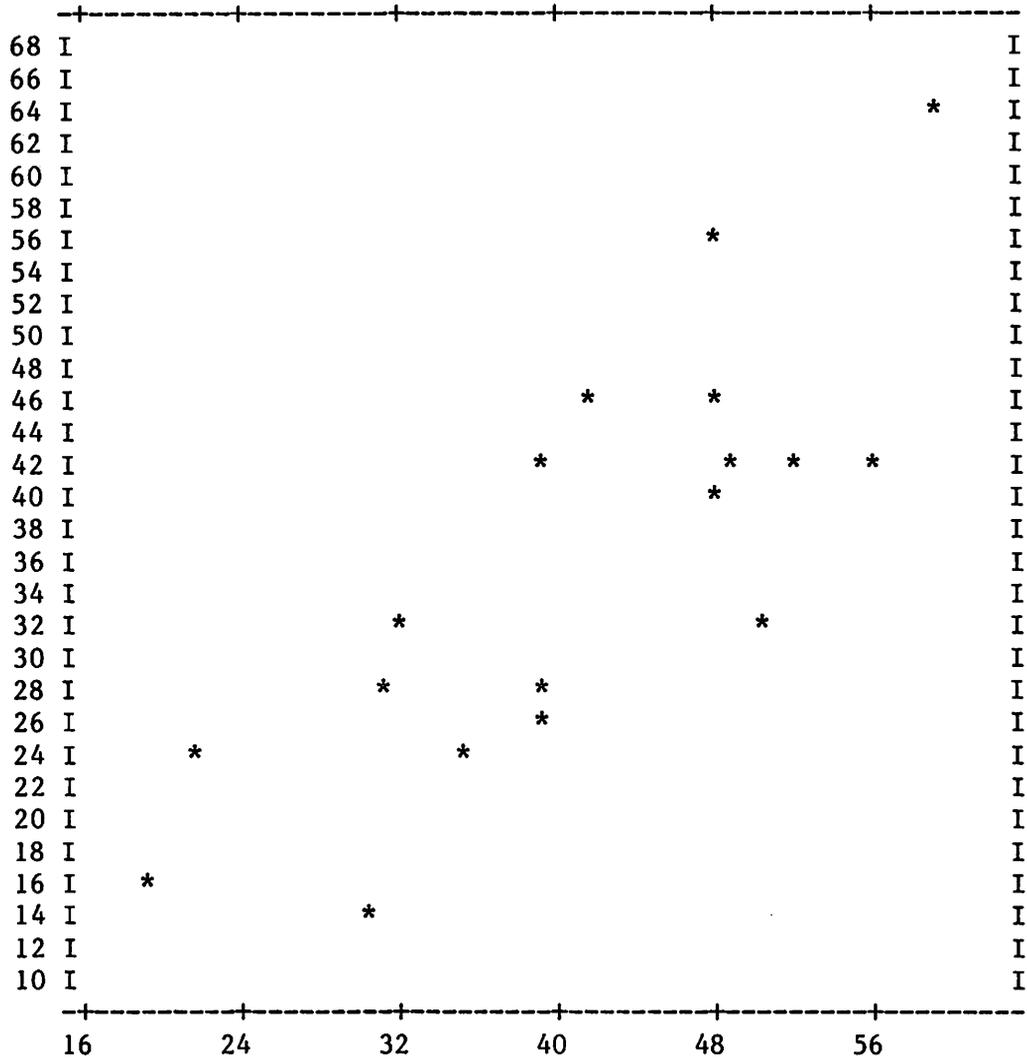
MEAN= 39.189



GENOTYPE 3

F-STATISTIC FOR NON-LINEARITY= * DF= 16 0

MEAN= 46.822



GENOTYPE 4

F-STATISTIC FOR NON-LINEARITY= * DF= 16 0

MEAN= 35.667

REFERENCES

- Clifford, B.C. (1974) Physiologic specialisation of *Puccinia hordei* in barley hosts with non-hypersensitive resistance. *Trans. Britt. Mycol. Soc.*, 63, 421-430.
- Eberhart, S.A. (1966) Stability parameters for comparing varieties. *Crop Sci.*, 6, 36-40.
- Finlay, K.W. (1963) Analysis of adaptation in a breeding programme. *Aust. J. Agr. Res.*, 14, 742-754.
- Jenns, A.E. (1981) Estimation of specific resistance by regression analysis of disease in host lines inoculated with several pathogen isolates. (Abs.) *Phytopathology*, 71, 883.
- Leonard, K.J. (1979) Durability of general resistance: Evaluation of cultivar x isolate interactions. *Proc. 9th Int. Cong. Plant Protect.*, 1, 190-193.
- Yates, F. (1938) The analysis of groups of experiments. *J. Agric. Sci.*, 28, 556-580.

J. Barnard
J. Norelli
New York State Agricultural Experiment Station
Cornell University
Geneva
New York 14456
USA

NON-HIERARCHICAL CLASSIFICATION - CLASSIFY (AND CLASSF)

'CLASSIFY' has a number of pitfalls for the unwary; those which I know about I describe below, if users find any others they might care to write to this Newsletter or inform me.

1. NUMERICAL ACCURACY

The various criteria available in 'CLASSIFY' are not calculated afresh after each step in the transfer and swapping process, instead their values are updated. At any stage it might be the case that a particular transfer or swap (move) improves the criterion by a fairly accurately known small amount; however, if the criterion value is relatively large some of the accuracy of the improvement is lost. Over many moves this loss of accuracy can become quite large and, as with all problems involving numerical accuracy, can vary markedly across different types of computer. If an accurate determination of the criterion value is important, it would be safest to save the resulting classification (in a FACTOR called FINAL, say) and use the extra statement

'CLASSIFY/PRINT = CI, ITERATE = N' VSET; START = FINAL

which will print the Criterion value for the Initial classification and will Not iterate from the classification given in FINAL.

2. LOCAL OPTIMA

There is absolutely no guarantee that the global, rather than a local, optimum will be found. Like being burgled - it always happens to someone else. Beware: it can, and will, happen to you, especially if you do not guard against non-global optima and if your data set is large (what constitutes large depends on the problem - as few as two dozen units might cause problems). The classical approach is to run the classification a few times from randomly chosen starting points and compare the results (NB (1) above). Alternatively, you might choose a 'good' starting classification and hope for success, or even use various 'good' starting points and compare results. These two approaches, 'random' and 'good' starting points, are well known; what I was not aware of until recently was that there also exist 'bad' starting points that must be avoided (see below). Since 'random' can be 'good', 'bad' or 'indifferent', I suggest that a variety of 'good' starting points be used and results compared. I stress 'variety'; do not use a few minor variations of one 'good' starting point ('CLASSIFY' could most probably find these anyway, given the 'good' one).

3. BAD STARTING CLASSIFICATIONS

These comments stem from one example I have seen recently. A set of binary variates for over 100 units was used and no starting classification was provided. About a third of the units were identical and, unfortunately, were allocated amongst several groups in the pseudo-random starting classification used by 'CLASSIFY'. Of course, the identical units should all be in the same group; however, a local optimum was found for which the identical units were still spread around several groups. Clearly, any pair of these units in different groups could be swapped without altering the

criterion value, generating many classifications with equal criterion values and thus a mass of output (which was why I became aware of this problem). The moral of this story is to make sure that identical units start in the same groups. More practically, very similar units ought to start in the same group. (I know that that is precisely why you use 'CLASSIFY', but it is worth a little trouble to weed out the 'bad' starting points.)

4. GOOD STARTING CLASSIFICATIONS

There are a number of ways of getting a reasonable starting point. The macro CLASSF chooses a set of mutually distant points as nuclei and assigns units to their nearest nucleus to form a set of groups. This ought to be quite a good starting point under normal conditions but it does take a while to run and can fail (see below). Doing a preliminary hierarchical classification, using 'HIERARCHY', and assigning units according to the resulting clusters is another method that could be used. This requires two separate runs: the first determines the grouping, or equivalently the clustering threshold level at which the grouping is to be taken; the second does the non-hierarchical classification with the starting classification either fed in 'by hand' or obtained by rerunning the hierarchical classification with the GL option set to the required threshold level. A different scheme, which has the advantage of simplicity, is to take the mean over all the variates for each unit and use this to determine a grouping; eg for variates with values in the range zero to ten

```
'VARIATE LIMS = 1...9
'CALC'   ROWMEAN = VMEAN(VSET)
'GROUP'  MYSTART = LIMITS(ROWMEAN;LIMS)
```

will give a classification into (up to) ten groups. A less crude method, based on the same idea, would be to produce a graph of the first two principal component scores and group the units according to that. Finally, what might be the best method: use your own preconceived ideas. If you set the PRINT letter *I*, you can see how far wrong you were!

5. MACRO CLASSF

This operates as follows. For n units and p variables the units are considered as n points in a p -dimensional space. The first two nuclei are the two most distant points; the third nucleus is the point most distant from the line joining the first pair of nuclei; the fourth nucleus is the point most distant from the plane defined by the first three nuclei and so on. When the required number of nuclei have been found, the units are assigned to their nearest nucleus. This works well, but there is a restriction that no more than $p+1$ groups can be formed, eg when $p=2$ all the points lie on a plane which is precisely the plane defined by the first three nuclei, so a fourth nucleus cannot be found. As a consequence, when all the points lie in a k -dimensional subspace, no more than $k+1$ groups will be formed, even if more were asked for by the user. In the current version of the macro, no comment is made when this occurs; however, future versions (ie with Genstat 4.04 and later) produce a warning and print out the classification (into $k+1$ groups) when this occurs.

I hope that these remarks are useful to future classifiers; I am sure that the editors would welcome letters from users with other comments.

*P.G.N. Digby
Rothamsted Experimental Station
Harpenden
Hertfordshire
AL5 2JQ
U.K.*

GENSTAT AND PRIME CPL

The New Zealand Ministry of Agriculture and Fisheries has a PRIME computer network on which Genstat can be used in either batch or interactive mode.

For applications in which a general Genstat program can be used with certain parameters, options, calculations or printing specified by the user, we have found it useful to combine Genstat with PRIME CPL (Prime's Command Procedure Language).

Using CPL data allows people, who are not familiar with Genstat programming, to use Genstat even for quite complex applications.

One application where we have used Genstat and CPL together was in a program to investigate properties of three class attribute sampling plans.

The program calculates the operating characteristic for the scheme specified, assuming that samples are normally distributed. It then graphs the O.C. curve and prints the probabilities if this option has been set.

The following listings show the program, a sample run and the output from the run.

We have written a more general version of this program in which there is a choice between several different sampling distributions, and also an option to compare more than one scheme at a time.

Many of the CPL commands are self-explanatory. They can occur outside the Genstat program or even amongst the Genstat commands: in either case CPL evaluates them before running the Genstat program.

Some of the CPL commands used in the example are:-

R filename	Used to run the CPL program.
&	Most CPL directives begin with an ampersand.
&S	Used to set a value to a CPL variable.
%x%	References to variables are enclosed in % signs.
[]	Function calls are enclosed in brackets.
QUERY	Asks a question with supplied prompt.

RESPONSE Inputs data in response to prompt.

TYPE Prints a line of text on the screen.

PROGRAM LISTING

```
TYPE                            /* 3 CLASS SAMPLING PLANS
TYPE CHARACTERISTICS OF SAMPLING SCHEMES
&S N := [RESPONSE 'Number of samples' ]
&S C := [RESPONSE 'Acceptance number' ]
TYPE
TYPE LIMITS OF ALLOWABLE QUALITY
&S M1 := [RESPONSE 'Best quality defined as marginal' ]
&S M2 := [RESPONSE 'Best quality defined as bad' ]
TYPE
TYPE RANGE OF QUALITY OF INTEREST
&S X0 := [RESPONSE 'Lower limit for mean lot values' ]
&S X1 := [RESPONSE 'Upper limit for mean lot values' ]
&S S := [RESPONSE 'Standard deviation' ]
TYPE
&S NN := [RESPONSE 'Number of values calculated (Default=21, Max=99)' 21]
&IF %NN% > 99 &THEN %NN% := 99
&DATA GENSTAT -L2 CLASS3.OUT
'REFE/PR=N' CLASS3_SCHEMES
'SCAL' NN = %NN% : N = %N% : C = %C% : M1 = %M1% : M2 = %M2%
:    X0 = %X0% : X1 = %X1% : S = %S% : STEP,J(1...C)
'VARI' MN_VALUE,P_ACCEPT,P_GOOD,P_MARG,P_G_OR_M,PQ,TT $NN
'VARI' Y=1...NN    'VARI' BVAL=0,1,*,*
'HEAD' H1='OPERATING CHARACTERISTIC CURVE FOR 3-CLASS SAMPLING PLAN'
'HEAD' H2='PROBABILITIES OF ACCEPTANCE FOR THIS 3-CLASS PLAN'
'HEAD' H3=' (INFINITE LOT SIZES)'
'HEAD' HY='PROBABILITY OF ACCEPTANCE'    'HEAD' HX='MEAN LOT COUNT'
'HEAD' HEADY = ''
NUMBER OF SAMPLES                            %N%
ACCEPTANCE NUMBER                            %C%
LIMIT FOR MARGINALS                            %M1%
LIMIT FOR BADS                                %M2%
LOWER LIMIT FOR MEAN LOT VALUES            %X0%
UPPER LIMIT FOR MEAN LOT VALUES            %X1%

NORMAL DISTRIBUTION

STANDARD DEVIATION    %S%            STEP SIZE VALUE ''
'CALC' J(1...C)=1...C
'CALC' STEP=(X1-X0)/(NN-1)
'CALC' MN_VALUE=X0+(Y-1)*STEP
'CALC' P_GOOD,P_G_OR_M=NPI((M1,M2-MN_VALUE)/S)
'CALC' P_MARG=P_G_OR_M - P_GOOD
'CALC' PQ=P_MARG/P_GOOD
'CALC' TT,P_ACCEPT=P_GOOD*N
'FOR' J=J(1...C)
'CALC' TT=TT*(N-J+1)*PQ/J
'CALC' P_ACCEPT=P_ACCEPT+TT
'REPE'
'OUTPUT' 2 $80
'GRAPH/BV=BVAL,ATY=HY,ATX=HX' P_ACCEPT;MN_VALUE
'PRINT/C,LABR=1' H1,H3,HEADY,STEP $/,2(0),/,0,8.3
&IF [QUERY 'Do you want probabilities printed. (Default=Yes)' TRUE] &THEN &DO
  'PAGE'
  'PRINT/C,LABR=1' H1,H3,HEADY,STEP,H2,H3 $/,2(0),/,0,8.3,4/,2(0)
  'PRINT/P' MN_VALUE,P_ACCEPT
  &IF [QUERY 'Do you want extra probabilities printed' ] &THEN ~
    [UNQUOTE ',P_GOOD,P_MARG,P_G_OR_M']
    $15.2,(12.4)4
&END
'OUTPUT' 1
'RUN'
'CLOSE'
'STOP'
&END
&RETURN
```

EXAMPLE RUN

OK, R CLASS3

CHARACTERISTICS OF SAMPLING SCHEMES

Number of samples: 25
Acceptance number: 5

LIMITS OF ALLOWABLE QUALITY

Best quality defined as marginal: 7
Best quality defined as bad: 9

RANGE OF QUALITY OF INTEREST

Lower limit for mean lot values: 3
Upper limit for mean lot values: 8
Standard deviation: 1.5

Number of values calculated (Default=21, Max=99):

Do you want probabilities printed. (Default=Yes)?

Do you want extra probabilities printed? Y

>'REFE/PR=N' CLASS3_SCHEMES

>'SCAL' NN = 21 : N = 25 : C = 5 : M1 = 7 : M2 = 9

> : X0 = 3 : X1 = 8 : S = 1.5 : STEP,J(1...C)

>'VARI' MN_VALUE,P_ACCEPT,P_GOOD,P_MARG,P_G_OR_M,PQ,TT \$NN

>'VARI' Y=1...NN 'VARI' BVAL=0,1,*,*

>'HEAD' H1='OPERATING CHARACTERISTIC CURVE FOR 3-CLASS SAMPLING PLAN''

>'HEAD' H2='PROBABILITIES OF ACCEPTANCE FOR THIS 3-CLASS PLAN''

>'HEAD' H3='(INFINITE LOT SIZES)''

>'HEAD' HY='PROBABILITY OF ACCEPTANCE'' 'HEAD' HX='MEAN LOT COUNT''

>'HEAD' HEADY = ''

>NUMBER OF SAMPLES 25

>ACCEPTANCE NUMBER 5

>LIMIT FOR MARGINALS 7

>LIMIT FOR BADS 9

>LOWER LIMIT FOR MEAN LOT VALUES 3

>UPPER LIMIT FOR MEAN LOT VALUES 8

>

>NORMAL DISTRIBUTION

>

>STANDARD DEVIATION 1.5 STEP SIZE VALUE ''

>'CALC' J(1...C)=1...C

>'CALC' STEP=(X1-X0)/(NN-1)

>'CALC' MN_VALUE=X0+(Y-1)*STEP

>'CALC' P_GOOD,P_G_OR_M=NPI((M1,M2-MN_VALUE)/S)

>'CALC' P_MARG=P_G_OR_M - P_GOOD

>'CALC' PQ=P_MARG/P_GOOD

>'CALC' TT,P_ACCEPT=P_GOOD**N

>'FOR' J=J(1...C)

>'CALC' TT=TT*(N-J+1)*PQ/J

>'CALC' P_ACCEPT=P_ACCEPT+TT

>'REPE'

>'OUTPUT' 2 \$80

>'GRAPH/BV=BVAL,ATY=HY,ATX=HX' P_ACCEPT;MN_VALUE

>'PRINT/C,LABR=1' H1,H3,HEADY,STEP \$/,2(0),/,0,8.3

> PAGE

> 'PRINT/C,LABR=1' H1,H3,HEADY,STEP,H2,H3 \$/,2(0),/,0,8.3,4/,2(0)

> 'PRINT/P' MN_VALUE,P_ACCEPT

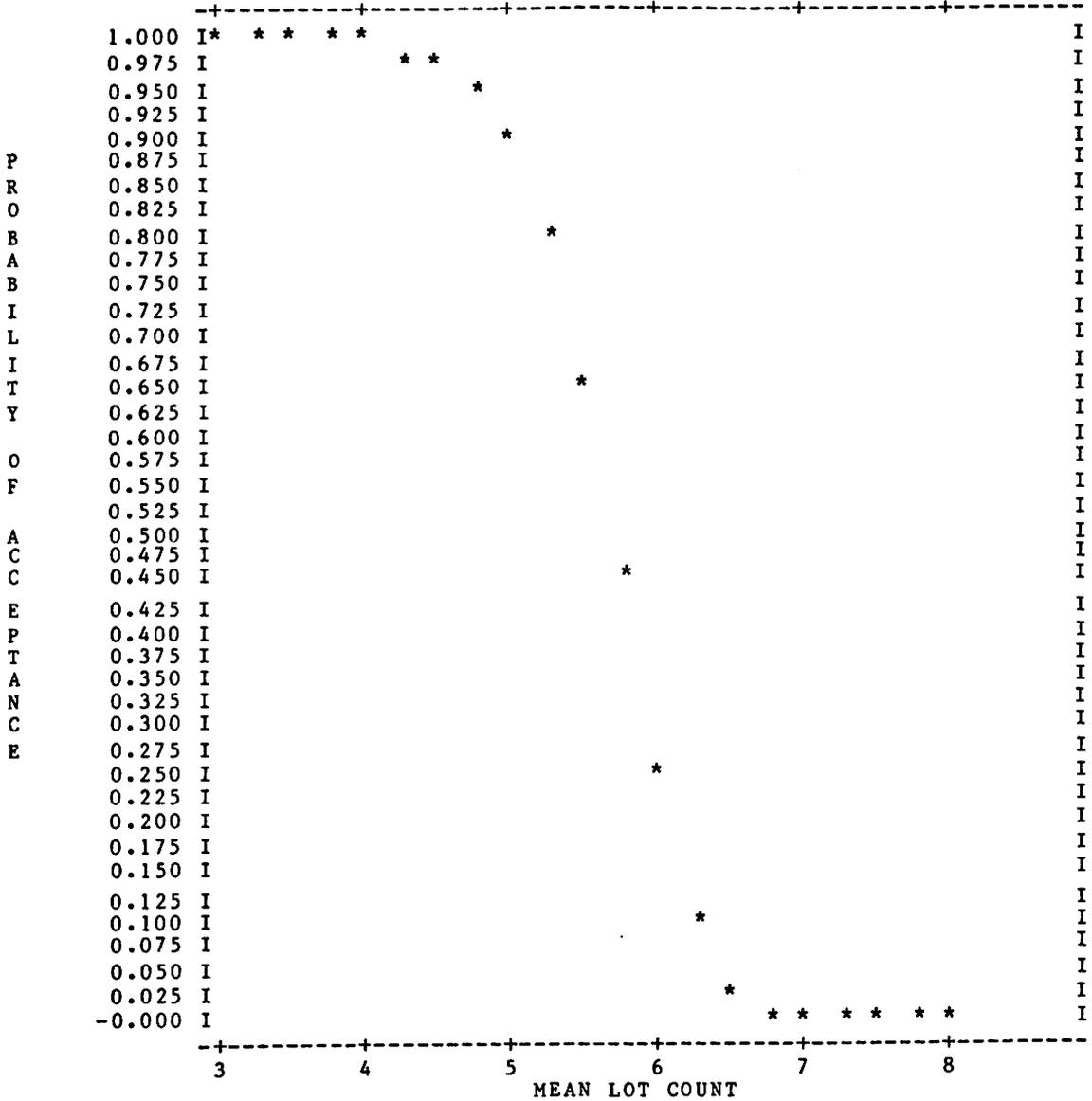
> ,P_GOOD,P_MARG,P_G_OR_M

> \$15.2,(12.4)4

>'OUTPUT' 1

>'RUN'

EXAMPLE OUTPUT



OPERATING CHARACTERISTIC CURVE FOR 3-CLASS SAMPLING PLAN (INFINITE LOT SIZES)

NUMBER OF SAMPLES 25
 ACCEPTANCE NUMBER 5
 LIMIT FOR MARGINALS 7
 LIMIT FOR BADS 9
 LOWER LIMIT FOR MEAN LOT VALUES 3
 UPPER LIMIT FOR MEAN LOT VALUES 8

NORMAL DISTRIBUTION

STANDARD DEVIATION 1.5 STEP SIZE VALUE 0.250

OPERATING CHARACTERISTIC CURVE FOR 3-CLASS SAMPLING PLAN (INFINITE LOT SIZES)

NUMBER OF SAMPLES	25
ACCEPTANCE NUMBER	5
LIMIT FOR MARGINALS	7
LIMIT FOR BADS	9
LOWER LIMIT FOR MEAN LOT VALUES	3
UPPER LIMIT FOR MEAN LOT VALUES	8

NORMAL DISTRIBUTION

STANDARD DEVIATION	1.5	STEP SIZE VALUE	0.250
--------------------	-----	-----------------	-------

PROBABILITIES OF ACCEPTANCE FOR THIS 3-CLASS PLAN (INFINITE LOT SIZES)

MN_VALUE	P_ACCEPT	P_GOOD	P_MARG	P_G OR M
3.00	0.9992	0.9962	0.0038	1.0000
3.25	0.9984	0.9938	0.0061	0.9999
3.50	0.9969	0.9902	0.0097	0.9999
3.75	0.9942	0.9849	0.0149	0.9998
4.00	0.9893	0.9772	0.0223	0.9996
4.25	0.9808	0.9666	0.0326	0.9992
4.50	0.9660	0.9522	0.0464	0.9986
4.75	0.9398	0.9332	0.0645	0.9977
5.00	0.8915	0.9088	0.0874	0.9962
5.25	0.8025	0.8783	0.1155	0.9938
5.50	0.6543	0.8413	0.1488	0.9902
5.75	0.4542	0.7977	0.1872	0.9849
6.00	0.2509	0.7475	0.2297	0.9772
6.25	0.1037	0.6915	0.2752	0.9666
6.50	0.0304	0.6306	0.3217	0.9522
6.75	0.0061	0.5662	0.3670	0.9332
7.00	0.0008	0.5000	0.4088	0.9088
7.25	0.0001	0.4338	0.4445	0.8783
7.50	0.0000	0.3694	0.4719	0.8413
7.75	0.0000	0.3085	0.4891	0.7977
8.00	0.0000	0.2525	0.4950	0.7475

June Atkinson
Biometrics Section
Ministry of Agriculture and Fisheries
Wellington
New Zealand


```
***** WARNING *****                                05/05/82
'TABLE'  IF MORE THAN 12 TABLES ARE DECLARED IN A TABLE
         STATEMENT A FORTRAN ERROR MAY OCCUR ON COMPUTERS
         WHICH CHECK ARRAY BOUNDS (E.G. UNIVAC, ICL 2980).
         THIS CAN BE AVOIDED BY SPLITTING THE STATEMENT
         INTO SEVERAL STATEMENTS.
```

(SEE ERROR EX3)

```
***** ERROR *****                                  27/09/82
'ESTIMATE' A FAULT OCCURS, WHEN ESTIMATING PARAMETERS
          IN TRANSFER FUNCTION MODELS, IF ARIMA MODELS ARE
          SPECIFIED FOR SOME INPUT SERIES BUT NOT FOR ALL.
```

(SEE ERROR TS2)

```
***** ERROR *****                                  27/09/82
'EXPAND'  A FAULT OCCURS WHEN EXPANDING MODELS WHICH HAVE
          FEW PARAMETERS COMPARED TO THE NUMBER IN THE
          EXPANDED FORM. AN EXAMPLE IS THE STANDARD
          AIRLINE - PASSENGER MODEL.
```

(SEE ERROR TS3)

```
***** GENSTAT 4.03 ERROR NOTICE NO. 11 15.7.82
**
**
***** ERROR DG12
***** MODULE DGA SUBROUTINE GRAFF
**
***** DECLARE SYMNO AS INTEGER UNDER LOCALS
***** 2 LINES BELOW LABEL 21, OMIT 1 LINE:
**
21 CONTINUE
   IF(NFM.NE.1) GO TO 30
   DO 20 I=1,MVZ
     INTERP(I)=1
**
***** ON LINE LABELLED 34
**
C     HEADING
C
* 34 IF((KCNT.GT.1).OR.(IAC.GT.1)) GO TO 25
     IJK=VALOR(3)+1
     JB=IJK*NBV(4)
*     SYMNO=ISDATA(IJK)*NBV(4)+JB
     JA=JB
* 25 DO 35 J=1,NVZ
     NOLV(J)=1
37 JA=JA+1
   CALL GBYT(IDATA,JA,KA)
*   IF(JA.GE.SYMNO) JA=JB
   IF(KA.EQ.KCH(47)) GO TO 37
38 JORIG(J)=ISAD
**
***** END OF ERROR DG12
**
**
```

```
***** ERROR DG13
***** MODULE DGB SUBROUTINE LPLOT(F)
**
***** 5 LINES BELOW LABEL 90
**
      IF(ICPNT(1).NE.1) GO TO 120
      IF(ISYM.EQ.KOLON) GO TO 300
*     IF(KB.EQ.KCH(55)) GO TO 190
*     IF(KB.EQ.KCH(49)) GO TO 190
      KSYM=KCH(56)
      IF(.NOT.PRIME) GO TO 190
**
***** END OF ERROR DG13
**
**
***** ERROR MV3
***** MODULE MVC SUBROUTINE TQLA
**
***** BETWEEN STATEMENTS LABELLED 40 AND 50
***** THE FIRST AND LAST PAIRS OF STATEMENTS BELOW REMAIN UNALTERED
***** DECIDE WHETHER THE DOUBLE(*) OR SINGLE(CS) PRECISION VERSION
***** IS REQUIRED AND CHANGE THE REMAINING STATEMENTS APPROPRIATELY
**
      P=(DATA(IA)-DATA(IA-1))/(DTWO *DATA(IB))
C.SD 2
CS   IF(ABS(P).LT.DONE) GO TO 42
CS   R=ABS(P)*SQRT((1/P)*(1/P)+DONE)
*    IF(DABS(P).LT.DONE) GO TO 42
*    R=DABS(P)*DSQRT((1/P)*(1/P)+DONE)
      GO TO 44
C.SD 1
CS 42 R=SQRT(P*P+DONE)
* 42 R=DSQRT(P*P+DONE)
      44 PR=P+R
      IF(P.LT.DZERO) PR=P-R
**
**
***** END OF ERROR MV3
**
**
***** ERROR EX3
***** MODULE EX SUBPROGRAM EXECD
**
***** AT LABEL 39:
**
      39 IF(NUNK.LE.0) GO TO 391
      IUNK=IUNK+1
      SCALNO(1)=ISDATA(IUNK)
      GO TO L39, (10,110)
      391 IF(GETATT(3,0).NE.0) GO TO 1000
**
**
***** END OF ERROR EX3
**
```

```
**
***** ERROR TS2
***** MODEL TSA SUBPROGRAM ESTIMA
**
***** AT LABEL 230
**
C      MISSING VALUES IN LIST 3 ARE REPRESENTED BY ZEROES
C      IN ARRAY AT KMODL2
*      IF(N.EQ.NINPUT) GO TO 280
C      FOR MISSING VALUES IN LIST 2, SET UP DEGENERATE MODELS
      OLDMOD = .FALSE.
      IVEC = 0
      IVOR = -1
      IVAL = 2
      ISWIT = 1
230 IF(VECSET(IVEC,ITYPE,IMODE,IVAL,IVOR,ISWIT,2).NE.0) GO TO 1000
**
**
***** END OF ERROR TS2
**
**
***** ERROR TS3
***** MODULE TSB SUBPROGRAM TSINT
**
***** AT LABEL 90
**
90 IP = IP*(1+ISDATA(IORD-2))*(1+ISDATA(IORD-1))
   IQ = IQ*(1+ISDATA(IORD))
   IR = IR+ISDATA(IORD-2)+ISDATA(IORD)
   IF(IORD.LE.IORIG+4) IORD = IORD-1
   IORD = IORD+4
   IF(IORD.LE.IMAX) GO TO 90
*     MAXPAR = IP+IQ+4
   IF(KODE.EQ.2) MAXPAR = IR
C     GET TEMPORARY WORKSPACE FOR NEW MODEL
**
**
***** END OF ERROR TS3
**
**
***** END OF ERROR NOTICE NO. 11
```

METHOD OF ISSUING SOURCE CORRECTIONS

Some site representatives at the 1982 NAG Users' Association meeting expressed dissatisfaction with the present method of issuing GENSTAT source corrections, and this is now being reviewed.

One improvement to be introduced at release 4.04 is that every revision issued will be recorded by a source comment identifying the sub-release:
e.g.

C RELEASE 4.04B (date)

and this extended release identification will be printed at the head of each GENSTAT job.

It would be possible to combine this with the present system, issuing an additional such comment with each error notice. However, other methods of issuing source corrections are possible and users' comments, either in favour of the current system or suggesting improvements, would be welcomed. These may be sent to the author or to the GENSTAT co-ordinator at NAG.

*H.R. Simpson
Rothamsted Experimental Station
Harpenden
Hertfordshire AL5 2JQ
U.K.*

- NOTICES -

GENSTAT CONFERENCE 1983

The third International Conference of users of the Genstat statistical package will take place at Orsay (near Paris) France, during 10-12 October 1983. Sessions of selected contributed papers are planned on:

- new facilities and future developments
- macros
- comparisons of Genstat with other packages
- interesting applications of Genstat
- the use of Genstat in education.

Participants will be able to view the Genstat video tapes, see demonstrations of the latest Genstat release and consult with expert users of the program.

Intending contributors should submit a one-page abstract, in English or in French, to:

R.W. Payne,
Statistics Department,
Rothamsted Experimental Station,
Harpenden,
Hertfordshire AL5 2JQ,
U.K.

before 1 April 1983.

If you would like to receive further information about the Conference, please complete and return the form below.

To: R.W. Payne, Statistics Department, Rothamsted Experimental Station,
Harpenden, Hertfordshire AL5 2JQ, U.K.

I am interested in attending the third Genstat Conference. Please send further information.

Name: _____

Organisation: _____

Address: _____

(If you intend to submit a paper, please give the title below:)

"AN INTRODUCTION TO GENSTAT"

This new book, by Norman Alvey, Nick Galwey and Peter Lane, has been issued in paperback by Academic Press, who have provided the following description:

Introduction to Genstat provides detailed guidance for students and researchers. No prior knowledge of computing is assumed and the statistical concepts used are explained. However, some familiarity with statistics will help the reader to follow the practical examples which illustrate the text. Many users will find all they need to know about Genstat in this book but for others it will serve as a helpful lead-in to the manual which documents the language. End of chapter exercises encourage the reader to write and run programs in the Genstat language. Solutions are provided in an appendix.

Copies may be obtained from bookshops for £8.50 (U.K. only) or US \$16.00 (elsewhere). Copies may also be ordered from NAG from £8.50 plus postage and packing.

GERMAN GENSTAT MANUAL

A German edition of the GENSTAT 4.03 manual is now available from NAG.

Produced in a single binder, this retails at £25 plus postage and packing.

Orders should be sent to

*Numerical Algorithms Group Ltd
NAG Central Office
Mayfield House
256 Banbury Road
Oxford OX2 7DE
U.K.*

ITSM ANNOUNCEMENTS

**CALL FOR PAPERS - 9TH INTERNATIONAL TIME SERIES MEETING (ITSM):
NOTTINGHAM UNIVERSITY (ENGLAND) 11-15 APRIL, 1983**

The following provides a first list of those (some still tentative) whom we hope will be speaking at this Conference (44 people from 16 countries, so far):

R.J. Bhansali (U.K.)	E.J. Godolphin (U.K.)
W.R. Bell (U.S.A.)	J.R.M. Hosking (U.K.)
R.J. Bennett (U.K.)	R.R. Hyatt (U.S.A.)
M. Borgard (France)	P.B. Kenny (U.K.)
W. Bruggeman (Belgium)	A.J. Lawrence (U.K.)
D.W. Bunn (U.K.)	J. Lillestøl (Norway)
M. Daub (Canada)	A. Lopez C. (Chile)
Y. Hosoya (Japan)	R.M. Loynes (U.K.)
A.J. de Hoyos (Brazil)	E.U. Makov (U.K.)
M. Deistler (Austria)	A. Maravall (Spain)
R.F. Galbraith (U.K.)	A. Milhoj (Denmark)

W. Mohr (West Germany)	T.M.F. Smith (U.K.)
G.V.L. Narasimham (U.S.A)	P. Stoica (Romania)
R.Y. Pei (U.S.A.)	T. Subba Rao (U.K.)
J. Pemberton (U.K.)	G. Thury (Austria)
D.A. Pierce (U.S.A.)	A.R. Tremayne (U.K.)
D.S. Poskitt (U.K.)	A.M. Walker (U.K.)
M.A. Rahim (Canada)	G.C. Watkins (U.S.A.)
D. Ray (India)	M. West (U.K.)
M.J. Routman (U.S.A.)	H.O. Wold (Sweden)
S. Shahabuddin (U.S.A.)	Y. Yajima (Japan) and
J.Q. Smith (U.K.)	S.K. Zaremba (U.K.)

Further contributed papers can still be considered, and 100-word abstracts (or enquiries concerning attendance) should be sent as soon as possible to the Convenor, Oliver D. Anderson, from whom details and application forms can also be obtained.

SCHEDULE OF INTERNATIONAL TIME SERIES MEETINGS (ITSM), 1982-5

General Interest ITSM, Cincinnati (U.S.A.)	19-22 August, 1982
General Interest ITSM, Nottingham (U.K.)	11-15 April, 1983
Special Topics ITSM, Valencia (Spain)	20-24 June, 1983
(Time Series Applications in Medicine and the Bio Sciences)	
Special Topics ITSM, Toronto (Canada)	11-14 August, 1983
(Hydrological, Geophysical and Spatial Time Series)	
General Interest ITSM, Toronto (Canada)	20-24 August, 1983
General Interest ITSM, Philadelphia (U.S.A.)	16-19 August, 1984
General Interest ITSM, Fort Collins (U.S.A.)	9-12 August, 1985

Fine Programmes of invited speakers are planned for these events, and 100-word abstracts for contributed papers and offers to act as session chairman should be sent to Oliver Anderson as soon as possible. As is usual with ITSMs, all Proceedings will be published.

The last four Conferences satellite the Annual ASA Joint Statistical Meetings: at each of these ASA meetings we hope to organise a special Time Series Analysis and Forecasting (TSA&F) session, featuring distinguished experts from outside North America.

If you would like to contribute, help (or just participate) at any of the above ITSM (or special TSA&F sessions), please write to Oliver D. Anderson.

SCOPE OF GENERAL INTEREST ITSM

Suitable time series topics for presentation include: Statistical Methodology; Applications to Economics and in Econometrics; Rational Expectations; Government, Business and Industrial Examples; Finance and Accountancy; the Hydrosociences, such as Limnology, Hydrology, Water Quality Regulation and Control, and the Modelling of Marine Environments; Persistence and Fractional Differencing; the Geosciences, especially such areas as Oil Exploration and Seismology; Civil Engineering and allied disciplines; Point Processes; Spatial and Space-Time Processes - their theory and application - especially in Geography and related areas, such as city planning or energy demand forecasting; Biology and Ecology;

Environmental Studies - Air and River Pollution; Medical Applications and Biomedical Engineering; Psychology; Speech Communication and Recognition; Acoustics; Radar; Sonar; Astronomy; Nuclear Science; Data Communication and Telephony; Irregularly Spaced Data (including Outliers and Missing Observations); Robust and Nonparametric Methods; Seasonal Modelling and Adjustment; Calendar Effects; Causality; Bayesian Approaches; Distributed Lags; Box-Jenkins Univariate ARIMA, Transfer-Function, Intervention, Vector ARIMA and other Multivariate Modelling; State Space; Nonlinear Models; Identification Problems; Estimation; Diagnostic Checking; Signal Extraction; Comparative Studies; Spectral Analysis, especially for the Physical Sciences; Business Cycle and Expectations Data; Data Revisions; Computer Software and Numerical Analysis; Forecasting (including new topics such as Traffic Forecasting and Safety, and Forecasting in Agriculture); and, no doubt, many other areas of the subject.

Suggestions (and offers of help) are always most welcome.

*Oliver D Anderson
ITSM Convenor
9 Ingham Grove
Lenton Gardens
Nottingham NG7 2LQ
U.K.*

GENSTAT NEWSLETTER ORDER FORM

To order future issues of the Genstat Newsletter, please complete the form below and return it to:

The Genstat Co-ordinator
NAG Central Office
Mayfield House
256 Banbury Road
OXFORD
OX2 7DE
U.K.

(Each Genstat site representative *automatically* receives one copy of each issue, free of charge.)

Please note that each subscription to the Newsletter costs £5.00 per annum (2 issues). This price includes 2nd class/surface postage. Postage at other rates will be charged at cost.

Back issues of the Newsletter are available on microfiche (24X). The first contains issues 1 - 6 and each subsequent fiche contains two issues - 7/8, 9/10 etc. Each fiche costs the same as a year's subscription to the Newsletter (£5.00).

To: NAG Ltd., Mayfield House, 256 Banbury Road, OXFORD OX2 7DE, U.K.

Please supply me with ... copies of each future issue of the GENSTAT Newsletter * for ... years until further notice beginning with issue number

(Minimum subscription period: 2 years)

Please supply me with ... microfiche of each of the following issues

Enclosed is my remittance for

Please invoice me.

Signature

Name and address for posting

(please type or print

Special mailing instructions

Cheques to be made payable to the Numerical Algorithms Group Ltd.

* delete one alternative
tick one box

