



# Generalized Linear Models

## John Nelder's second legacy to Statistics

Roger Payne

VSN International, 5 The Waterhouse,  
Waterhouse Street, Hemel Hempstead, UK

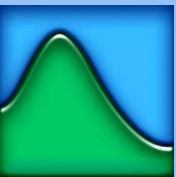
email: [Roger.Payne@vsni.co.uk](mailto:Roger.Payne@vsni.co.uk)

Palm Cove, 12<sup>th</sup> July 2011

# Early work

- Brooks, S.C. (1918). *Jour. Gen. Physiol.*, 1, 61
  - inherent variability among individuals of a population is considered to be responsible for the S-shaped character of the curve
- Bliss, C.I. (1934). The method of probits. *Science*, 79, 38-39.
  - "The effectiveness ... is of primary interest ... in the range of dosages approaching 100% kill. But in this region, the usual type of curve flattens to an asymptote, so that comparisons are commonly based upon dosages which kill only from 25 to 75 per cent. of the organisms."
  - "... curve is ordinarily fitted free hand ... introduces an indeterminate distortion due to the experimental errors and to unconscious bias.."
  - "... difficulties can be minimized if percentage kill and dosage are transformed to units which may be plotted as straight lines on ordinary cross-section paper and hence permit fitting by the customary technique of least squares of the straight-line regression equation."
- Bliss, C.I. (1934). The method of probits – a correction. *Science*, 79, 409-410
  - redefines the *probit* as  $5 + \text{NPI}(p)$

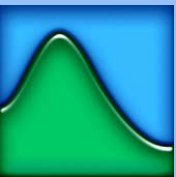
..



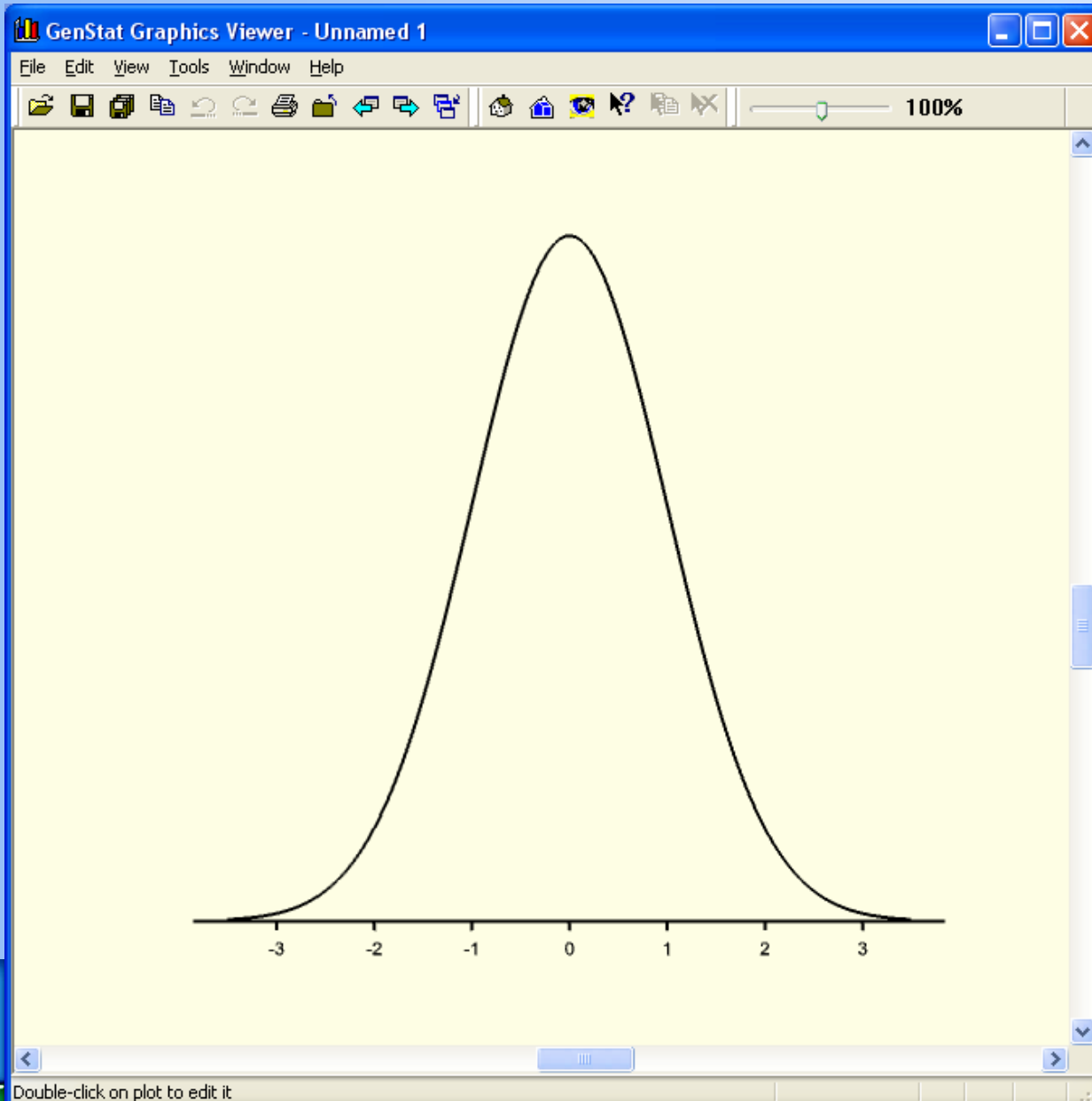
# Probit analysis

- models the relationship between a stimulus, like a drug, and a quantal response (success/failure)
- we assume there is a certain level of dose of the stimulus for each subject below which it will be unaffected, but above which it will respond
- this level of dose, known as its tolerance, will vary from subject to subject within the population
- for Probit transformation, it is assumed that the tolerance to the dose (or sometimes the the logarithm of the dose) has a Normal distribution
- so, if we plot the proportion of the population with each tolerance against (log) dose, we will obtain the familiar bell-shaped curve

..

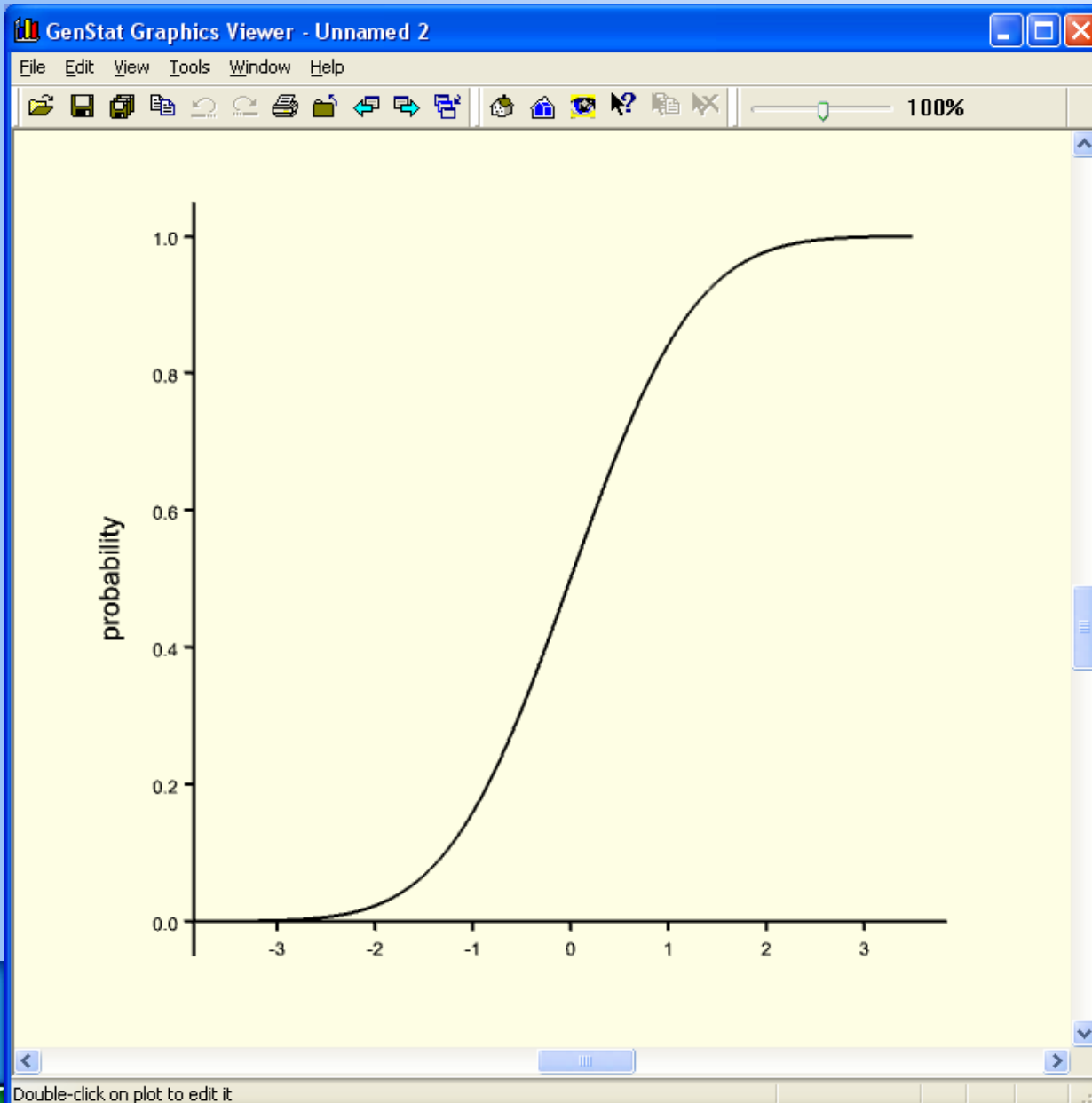


# Normal distribution



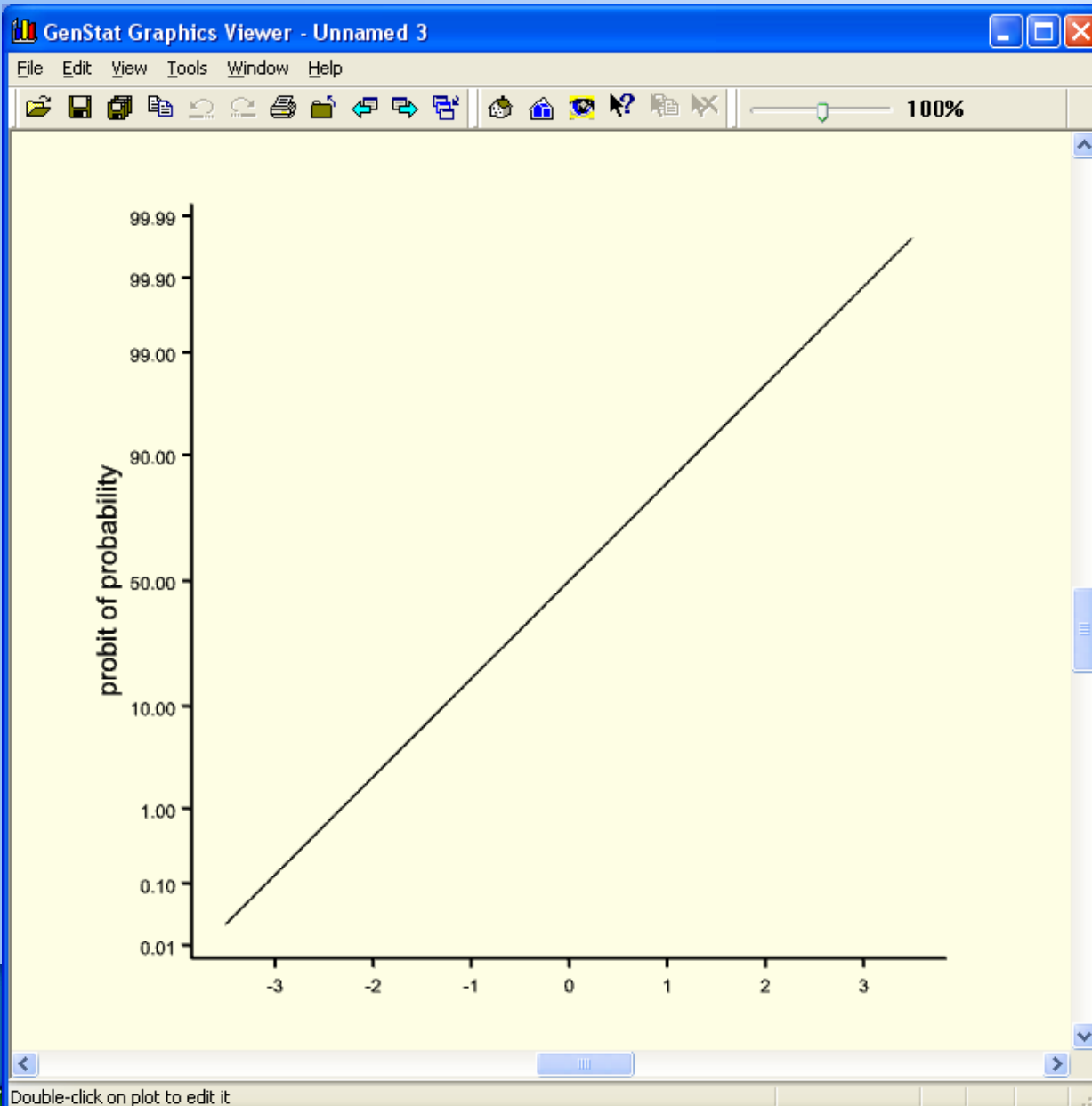
proportion of  
the population  
with each  
tolerance

# Cumulative Normal curve



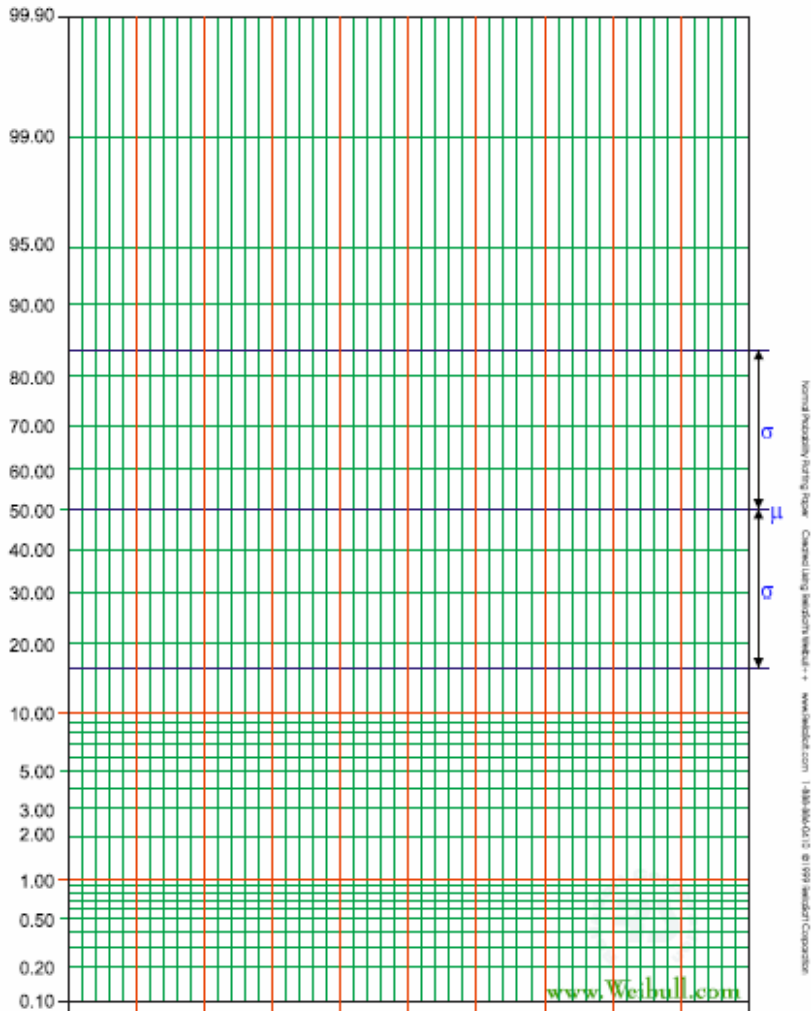
probability of response to dose  $x$  is the proportion of the population with a tolerance of less than  $x$  – we can read this off the cumulative Normal curve

# Probit transformation



take a probit transformation of the y axis (probability) to "stretch" the axis at the top and bottom to make the response into a straight line

# Probability paper



Galton (1899) *Nature*, 61, 102-104.

Hazen (1914) *Transactions of the American Society of Civil Engineers*, 77, 1539...

Whipple (1916) *Journal of Franklin Institute*, 182, 37-39 & 205-227.

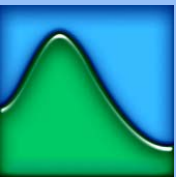
y-axis is on the probit scale, but labelled in percentages

(GenStat can plot like this by setting TRANSFORM option of YAXIS)

# Inspiration



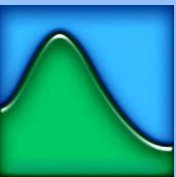
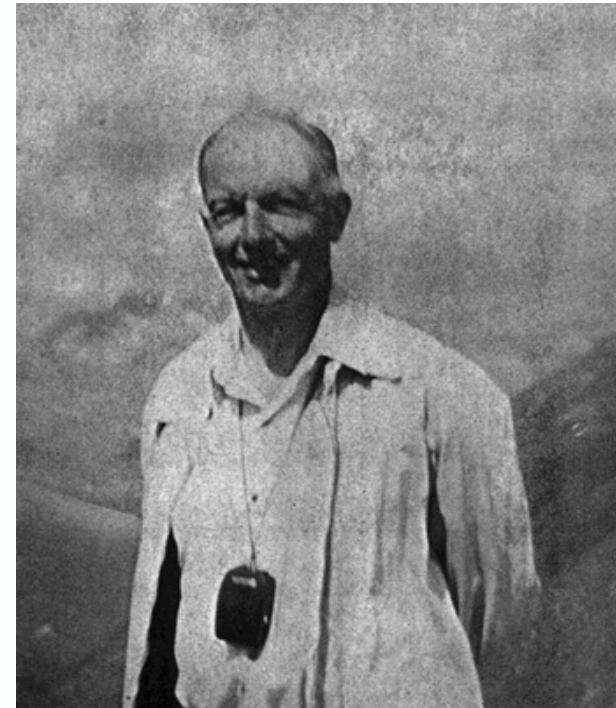
- Fisher, R.A. (1935). The case of zero survivors (Appendix to a paper by C.I. Bliss) *Ann. Appl. Biol.*, 22, 164-165.
  - Bliss "We were discussing probit analysis on a Sunday in Harpenden. ... empirical probits for zero and 100 % kill are minus and plus infinity ... I proposed assuming that half an insect died or lived so that they could be included ... He countered that these points were of little value and might be better omitted. I protested that they were as real observations as any others ... He was silent for a minute and then remarked 'When a biologist believes that there is information in an observation, it is up to the statistician to get it out.' After lunch that day he derived the maximum likelihood solution ... in which the empirical probits led to working probits based upon all the information in each quantal response, including those at zero and 100 percent."
  - from Joan Fisher Box's 1978 book *R.A. Fisher The Life of a Scientist*, p.272



# Refinements and further developments

- Tattersfield & Morris (1924). *Bull. Ent. Res.*, 14, 223-233.  
Abbott, W.S. (1925). *J. Econ. Ent.*, 18, 265-267.
  - adjustment for immunity and natural mortality
- Bliss, C.I. (1940). The relation between exposure time, concentration and toxicity in experiments on insecticides. *Ann, Ent. Soc. Amer.*, 33, 721-766.
  - probit plane (multiple regression) with  $\log(\text{concentration})$  and  $\log(\text{time})$
- Finney, D.A. *Probit Analysis* (1947)
  - full instructions for analysing probit lines, parallel lines and planes by hand
- Finney *Probit Analysis* 3<sup>rd</sup> ed (1971)
  - computer program to fit probits by maximum likelihood
- Ross, G.J.S. (1970) MLP
  - maximum likelihood program, with module for probit analysis (Jones & Payne)

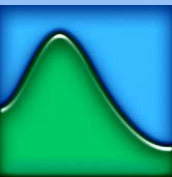
..



# Generalization

- Nelder. J.A. & Wedderburn, R.W.M. (1972). Generalized linear models. *J. Roy. Statist. Soc. A*
- ordinary regression - model  $y = \mu + \varepsilon$   
 $\mu$  is the mean predicted by a model  $\mu = X\beta$   
e.g.  $a + b \times x$   
 $\varepsilon$  is the residual with Normal distribution  $N(0, \sigma^2)$   
 $y$  (equivalently) has Normal distribution  $N(\mu, \sigma^2)$
- generalized linear model – still  $y = \mu + \varepsilon$   
but model now defines the *linear predictor*  $\eta = X\beta$   
 $\eta$  is related to  $\mu$  by the *link function*  $g()$   $\eta = g(\mu)$   
 $\varepsilon$  is the residual  
 $y$  has distribution from the *exponential family* (mean  $\mu$ )  
e.g. binomial, gamma, Normal, Poisson (and regression is weighted to take account of non-constant variances)  
iterative algorithm: estimate means, calculate weights, re-estimate means etc...

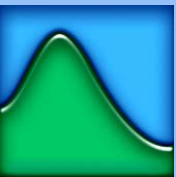
..



# Generalized linear models

- model (and link)
  - $E(y) = \mu$ , where  $g(\mu) = \eta = X\beta$
- distribution in exponential family (McC & N, p.28)
  - log likelihood  $l(\theta; y) = \sum \{y\theta - b(\theta)\} / a(\varphi) + c(y, \varphi)$
  - $\theta$  is the *canonical parameter* and  $\varphi$  is the *dispersion parameter*
    - *canonical link* is  $g()$  such that  $\eta = \theta$
    - e.g. identity for Normal, log for Poisson
  - $E(y) = \mu = b'(\theta)$  (i.e.  $\partial b / \partial \theta$ )
  - $\text{var}(y) = b''(\theta) a(\varphi)$
  - $b''(\theta)$  i.e.  $\partial^2 b / \partial \theta^2$  is the *variance function*  $V(\mu)$
  - usually dispersion function  $a(\varphi) = \varphi$  / prior-weight
- e.g. Normal distribution
  - $\log L(\theta; y) = \sum -(y-\mu)^2 / 2\sigma^2 - 1/2 \log(2\pi\sigma^2)$   
 $= \sum (y \times \mu - \mu^2 / 2) / \sigma^2 + \sum -y^2 / 2\sigma^2 - 1/2 \log(2\pi\sigma^2)$
  - so  $\theta = \mu$ ,  $b(\theta) = \theta^2 / 2$ ,  $\varphi = \sigma^2$ ,  $c(y, \varphi) = \sum -y^2 / 2\sigma^2 - 1/2 \log(2\pi\sigma^2)$
- e.g. Poisson
  - $\log L(\theta; y) = \sum y \log(\mu) - \mu - \log(y!)$
  - so  $\theta = \log(\mu)$ ,  $b(\theta) = \mu = \exp(\theta)$ ,  $\varphi = 1$ ,  $c(y, \varphi) = \sum \log(y!)$

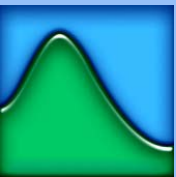
..



# Implementation – GLIM

- 1972
  - planned by the Working Party on Statistical Computing (set up with John as Chairman in 1966)
- 1974 – GLIM 1
  - same syntax for models as Genstat (Wilkinson & Rogers 1973)
  - simplified language – for interactive use
  - commands (directives) with fewer operands (no "options")
  - limited data structures – single data matrix (vectors same length)
  - allowed for offset
  - c. 50 sites)
- 1975 – GLIM 2
  - macro facility, scalar structures (%A... %Z), offsets
  - c. 130 sites
- 1978 – GLIM 3
  - vectors no longer all of same length, suffices, no need for maximal model, own links and error distributions, additional functions &c
  - c. 300 sites by 1980 – immensely influential

..



# GLIM

- example – probit lines from Finney (1971) p. 103

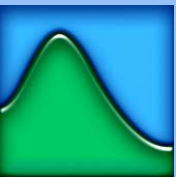
```
MS-DOS Command Prompt - glim-b
|GLIM release 4|
GLIM 4, update 8 for IBM etc. 80386 PC / DOS on 27-Mar-2004 at 20:02:25
(copyright) 1992 Royal Statistical Society, London

? $UNITS 14
? $FACTOR Drug 4
$FAC? $DATA Drug,Dose,Ntest,Nrespond
$DAT? $READ
$REA? 1 1.50 103 19    1 3.00 120 53    1 6.00 123 83
$REA? 2 1.50  60 14    2 3.00 110 54    2 6.00 100 81
$REA? 3 0.75  90 31    3 1.50  80 54    3 3.00  90 80
$REA? 4 5.00  60 13    4 7.50  85 27    4 10.00 60 32
$REA?                4 15.00 90 55    4 20.00 60 44 $

|Press F1 for help on windows , TAB to switch windows|

? $UNITS 14
? $FACTOR Drug 4
$FAC? $DATA Drug,Dose,Ntest,Nrespond
$DAT? $READ
$REA? 1 1.50 103 19    1 3.00 120 53    1 6.00 123 83
$REA? 2 1.50  60 14    2 3.00 110 54    2 6.00 100 81
$REA? 3 0.75  90 31    3 1.50  80 54    3 3.00  90 80
$REA? 4 5.00  60 13    4 7.50  85 27    4 10.00 60 32
$REA?                4 15.00 90 55    4 20.00 60 44 $
?
```

- GLIM 4 (with Genstat 5 Release 1.3 interface)



# GLIM

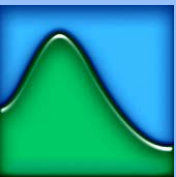
- common line

```
MS-DOS Command Prompt - glim-b
|GLIM release 4|
$REA? 4 5.00 60 13 4 7.50 85 27 4 10.00 60 32
$REA? 4 15.00 90 55 4 20.00 60 44 $
scaled deviance = 210.55 at cycle 3
residual df = 12

Scaled deviance is 210.6 on 12 d.f. from 14 observations
n.b. current and previous models are not necessarily nested

      estimate      s.e.      parameter
1      -0.2976      0.06629      1
2       0.2594      0.04164      LOGDOSE
scale parameter 1.000

|Press F1 for help on windows , TAB to switch windows|
$REA? 3 0.75 90 31 3 1.50 80 54 3 3.00 90 80
$REA? 4 5.00 60 13 4 7.50 85 27 4 10.00 60 32
$REA? 4 15.00 90 55 4 20.00 60 44 $
? $CALC Logdose = %LOG(Dose)
$CAL? $YUAR Nrespond
$YUA? $ERROR B Ntest
$ERR? $LINK P
? $FIT Logdose
$FIT? $DISPLAY DE $
?
```



# GLIM

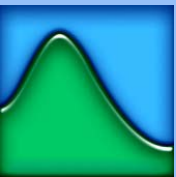
- parallel lines

```
MS-DOS Command Prompt - glim-b |GLIM release 4|
? $FIT + Drug
$FIT? $DISPLAY DE $
scaled deviance = 3.8680 <change = -206.7> at cycle 2
  residual df = 9      <change = -3  >

Scaled deviance is 3.9 on 9 d.f. from 14 observations
change is -206.7 for -3 d.f.

      estimate      s.e.      parameter
1      -1.379      0.1142      1
2       1.072      0.07495     LOGDOSE
3       0.2379     0.1083     DRUG<2>
4       1.360      0.1296     DRUG<3>
5      -1.180      0.1330     DRUG<4>
scale parameter 1.000

|Press F1 for help on windows , TAB to switch windows|
$ERR? $LINK P
? $FIT Logdose
$FIT? $DISPLAY DE $
?
? $FIT + Drug
$FIT? $DISPLAY DE $
? -
```



# GLIM

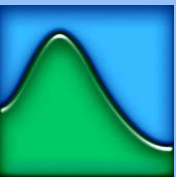
- different lines for each drug

```
MS-DOS Command Prompt - glim-b
|GLIM release 4|
? $FIT + Logdose.Drug
$FIT? $DISPLAY DE $
scaled deviance = 2.3344 <change = -1.534> at cycle 2
  residual df = 6      <change = -3      >

Scaled deviance is 2.334 on 6 d.f. from 14 observations
  change is -1.534 for -3 d.f.

      estimate      s.e.      parameter
1      -1.255      0.1710      1
2       0.9667      0.1321      LOGDOSE
3      -0.005823    0.2717      DRUG<2>
4       1.205      0.1971      DRUG<3>
5      -1.194      0.4019      DRUG<4>
6       0.2063      0.2108      DRUG<2>.LOGDOSE
7       0.2084      0.2065      DRUG<3>.LOGDOSE
8       0.05833     0.2014      DRUG<4>.LOGDOSE
scale parameter 1.000

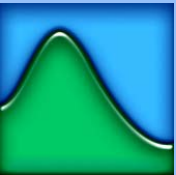
|Press F1 for help on windows , TAB to switch windows|
$FIT? $DISPLAY DE $
? $FIT + Logdose.Drug
$FIT? $DISPLAY DE $
? -
```



# Implementation – Genstat

- 1976
  - generalized linear models implemented in Release 3.09 by Ron Baxter and Peter Lane
- c. 1990
  - K system – John's procedures to recreate the GLIM environment in GenStat

..

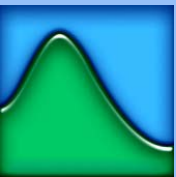


# Genstat 4.03e

- example – probit lines from Finney (1971) p. 103

```
C:\ Command Prompt - g
GENSTAT V MARK 4.03
<C>1980 LAWS AGRICULTURAL TRUST <ROTHAMSTED EXPERIMENTAL STATION>
MS-DOS VERSION BY C.E.M.S. < J.C. AND Y.M. >

>'REFERENCE' prohibit
  1 'REFERENCE' prohibit
>'UNIT' $ 14
  2 'UNIT' $ 14
>' ' define the data ' '
  3 ' ' define the data ' '
>'NAME' Dname = Morphine,Amidone,Phenadoxone,Pethidine
  4 'NAME' Dname = Morphine,Amidone,Phenadoxone,Pethidine
>'FACTOR' Drug $ Dname = 3<1...4>,4,4
  5 'FACTOR' Drug $ Dname = 3<1...4>,4,4
>'VARIATE' Dose = <1.5,3,6>2, 0.75,1.5,3, 5,7.5,10,15,20
  6 'VARIATE' Dose = <1.5,3,6>2, 0.75,1.5,3, 5,7.5,10,15,20
>: Ntest = 103,120,123,60,110,100,90,80,90,60,85,60,90,60
  7 : Ntest = 103,120,123,60,110,100,90,80,90,60,85,60,90,60
>: Nrespond = 19,53,83,14,54,81,31,54,80,13,27,32,55,44
  8 : Nrespond = 19,53,83,14,54,81,31,54,80,13,27,32,55,44
>'RUN' _
```

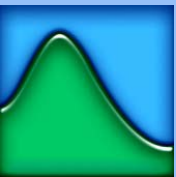


# Genstat 4.03e

- print the data

```
c:\ Command Prompt - g
10 'PRINT/P' Drug,Dose,Ntest,Nrespond
>'RUN'
11 'RUN'
```

Drug	Dose	Ntest	Nrespond
Morphine	1.5000E 0	1.0300E 2	1.9000E 1
Morphine	3.0000E 0	1.2000E 2	5.3000E 1
Morphine	6.0000E 0	1.2300E 2	8.3000E 1
Amidone	1.5000E 0	6.0000E 1	1.4000E 1
Amidone	3.0000E 0	1.1000E 2	5.4000E 1
Amidone	6.0000E 0	1.0000E 2	8.1000E 1
Phenadox	7.5000E -1	9.0000E 1	3.1000E 1
Phenadox	1.5000E 0	8.0000E 1	5.4000E 1
Phenadox	3.0000E 0	9.0000E 1	8.0000E 1
Pethidin	5.0000E 0	6.0000E 1	1.3000E 1
Pethidin	7.5000E 0	8.5000E 1	2.7000E 1
Pethidin	1.0000E 1	6.0000E 1	3.2000E 1
Pethidin	1.5000E 1	9.0000E 1	5.5000E 1
Pethidin	2.0000E 1	6.0000E 1	4.4000E 1

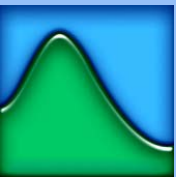


# Genstat 4.03e

- transform dose, define glm and fit constant

```
C:\ Command Prompt - g
Phenadox 3.0000E 0 9.0000E 1 8.0000E 1
Pethidin 5.0000E 0 6.0000E 1 1.3000E 1
Pethidin 7.5000E 0 8.5000E 1 2.7000E 1
Pethidin 1.0000E 1 6.0000E 1 3.2000E 1
Pethidin 1.5000E 1 9.0000E 1 5.5000E 1
Pethidin 2.0000E 1 6.0000E 1 4.4000E 1

>' transform dose to log base 10 ''
12 '' transform dose to log base 10 ''
>'CALCULATE' Logdose = LOG10(Dose)
13 'CALCULATE' Logdose = LOG10(Dose)
>' define the maximal model and binomial totals ''
14 '' define the maximal model and binomial totals ''
>'TERMS/TOTAL=Ntest' Nrespond + Drug*Logdose
15 'TERMS/TOTAL=Ntest' Nrespond + Drug*Logdose
>' define the glm and y-variate ''
16 '' define the glm and y-variate ''
>'Y/ERROR=BINOMIAL,LINK=PROBIT' Nrespond
17 'Y/ERROR=BINOMIAL,LINK=PROBIT' Nrespond
>' fit the constant ''
18 '' fit the constant ''
>'FIT/PRINT=Z,ANDEV=I'
19 'FIT/PRINT=Z,ANDEV=I'
>'RUN'
```





# Genstat 4.03e

- common line

```
C:\ Command Prompt - g

46 .....

***** REGRESSION ANALYSIS *****
ERROR DISTRIBUTION: BINOMIAL LINK FUNCTION: PROBIT

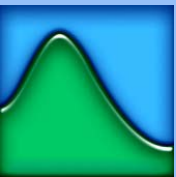
*** REGRESSION COEFFICIENTS ***
Y-VARIATE: Nrespond

```

	ESTIMATE	S.E.	T
CONSTANT	-0.29757	0.06629	-4.49
Logdose	0.59721	0.09589	6.23

```
*** STANDARD ERRORS BASED ON SCALE PARAMETER WITH VALUE 1.000

>' fit parallel lines ''
48 '' fit parallel lines ''
>'ADD/PRINT=C' Drug
49 'ADD/PRINT=C' Drug
>'RUN'
```



# Genstat 4.03e

- parallel lines

```
C:\ Command Prompt - g

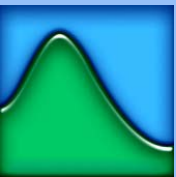
***** REGRESSION ANALYSIS *****
ERROR DISTRIBUTION: BINOMIAL LINK FUNCTION: PROBIT

*** REGRESSION COEFFICIENTS ***

Y-VARIATE: Nrespond

ESTIMATE          S.E.          T
CONSTANT          -1.37926      0.11421      -12.08
Logdose           2.46849      0.17260      14.30
Drug Amidone      0.23787      0.10835       2.20
Drug Phenadox    1.35955      0.12964      10.49
Drug Pethidin    -1.17991     0.13303      -8.87
*** STANDARD ERRORS BASED ON SCALE PARAMETER WITH VALUE 1.000

>' try different slopes ''
 27 ' try different slopes ''
>'ADD/PRINT=C' Logdose.Drug
 28 'ADD/PRINT=C' Logdose.Drug
>' RUN'
```



# Genstat 4.03e

- different lines for each drug

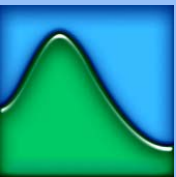
```
C:\ Command Prompt - g

***** REGRESSION ANALYSIS *****
ERROR DISTRIBUTION: BINOMIAL LINK FUNCTION: PROBIT

*** REGRESSION COEFFICIENTS ***
Y-VARIATE: Nrespond

ESTIMATE          S.E.          T
CONSTANT          -1.2548701    0.1709866    -7.34
Logdose           2.2259541    0.3041332     7.32
Drug Amidone      -0.0058154    0.2716760    -0.02
Drug Phenadox     1.2046237    0.1970555     6.11
Drug Pethidin    -1.1938599    0.4019285    -2.97
Logdose.Drug Amidone  0.4749701    0.4853492     0.98
Logdose.Drug Phenadox  0.4798320    0.4753822     1.01
Logdose.Drug Pethidin  0.1343163    0.4636877     0.29
*** STANDARD ERRORS BASED ON SCALE PARAMETER WITH VALUE 1.000

>' ADD/PRINT=Z,ANDEV=T'
  30 ' ADD/PRINT=Z,ANDEV=T'
>' RUN'
```



# Genstat 4.03e

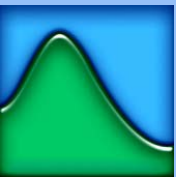
- andev table

```
C:\ Command Prompt - g

*** SUMMARY ANALYSIS OF DEVIANCE ***
Y-VARIATE: Nrespond

TERMS
DEU .
TIO
INITIAL MODEL
CONSTANT
RESIDUAL DF DEVIANCE CHANGE DF DEVIANCE MEAN CHANGE MN RA
13 * * *
MODIFICATIONS TO MODEL
13 249.9579 0 *
+Logdose 12 210.5501 1 39.4078 39.4078 101.29
+Drug 9 3.8680 3 206.6821 68.8940 177.08
+Logdose .Drug 6 2.3344 3 1.5336 0.5112 1.31
6 2.3344 0 0.0000 *
**DENOMINATOR OF RATIO IS RES.DEU./RES.DF FROM LINE ABOVE, = 0.3891

>
>'STOP'
```



# Implementation – K system

GenStat

File Edit Search Run Data Spread Graphics Stats Options Window Help

Output

```

5 kinit
The K system is now initialized. For on-line help type KHELP.

6 kun 14          "$UNITS"
7 kyv Nresponded "$YUAR"
8 kerr b; Ntested "$ERROR B"
9 klin p          "$LINK P"
10 kfit Logdose   "$FIT"
model re-initialized
dev. = 216.3 d.f. = 12 changes = * & *
11 kpe            "$DISPLAY DE $"
                estimate      s.e.          t
1  Constant      5.474          0.4720     11.597
2  Logdose       -2.678          0.4712     -5.684
Note: s.e.s assume dispersion = 1.000
12 kadd Drug      "$FIT +"
dev. = 39.94 d.f. = 9 changes = -176.4 & -3
13 kpe            "$DISPLAY DE $"
                estimate      s.e.          t
1  Constant      11.081         1.0782     10.278
2  Logdose       -11.328        1.3791     -8.214
3  Drug Amidone  -0.746         0.3172     -2.353
4  Drug Phenadoxone -4.437        0.4860     -9.129
5  Drug Pethidine  5.430         0.7915      6.861
Note: s.e.s assume dispersion = 1.000
14 kadd Logdose.Drug "$FIT +"
dev. = 29.37 d.f. = 6 changes = -10.58 & -3
15 kpe            "$DISPLAY DE $"
                estimate      s.e.          t
1  Constant      16.377         3.601        4.548
2  Logdose       -18.269        4.666        -3.916
3  Drug Amidone  -5.223         4.337        -1.204
4  Drug Phenadoxone -12.120        3.664        -3.308
5  Drug Pethidine  1.894         5.047         0.375
6  Logdose.Drug Amidone  5.880         5.615         1.047
7  Logdose.Drug Phenadoxone 11.971        4.882         2.452
8  Logdose.Drug Pethidine  5.559         5.430         1.024
Note: s.e.s assume dispersion = 1.000

```

D:\rp\papers\Welder80\kprobit.gen

```

spload 'probit.gsh'
calc Logdose = LOG10(Dose)
open 'ksys.glb'; channel=1; file=procedure
kinit
kun 14          "$UNITS"
kyv Nresponded "$YUAR"
kerr b; Ntested "$ERROR B"
klin p          "$LINK P"
kfit Logdose   "$FIT"
kpe            "$DISPLAY DE $"
kadd Drug      "$FIT +"
kadd Logdose.Drug "$FIT +"
kpe            "$DISPLAY DE $"

```

kprobit.gen Server Ready Ln 15, Col 1 D:\rp\papers\Welder80 INS

# Implementation – 14<sup>th</sup> Edition

The screenshot displays the GenStat software interface. The main window shows the results of a regression analysis. The 'Generalized Linear Models' dialog box is open, showing the model configuration. A 'Spreadsheet [Drug.gsh]' window is also open, displaying the data used for the analysis.

**Regression analysis**

Response variate: R  
 Binomial totals: N  
 Distribution: Binomial  
 Link function: Probit  
 Fitted terms: Constant + Logdose + Drug + Logdose.Drug

**Summary of analysis**

Source	d.f.	deviance	mean deviance	deviance ratio	approx chi pr
Regression	7	247.624	35.3748	35.37	<.001
Residual	6	2.334	0.3891		
Total	13	249.958	19.2275		

Dispersion parameter is fixed at 1.00.

*Message: deviance ratios are based on dispersion parameter with value 1.*

**Estimates of parameters**

Parameter	estimate	s.e.	t(*)	t pr.
Constant	-1.255	0.171	-7.34	<.001
Logdose	2.226	0.304	7.32	<.001
Drug Amidone	-0.006	0.272	-0.02	0.983
Drug Phenadoxone	1.205	0.197	6.11	<.001
Drug Pethidine	-1.194	0.402	-2.97	0.003
Logdose.Drug Amidone	0.475	0.485	0.98	0.328
Logdose.Drug Phenadoxone	0.480	0.475	1.01	0.313
Logdose.Drug Pethidine	0.134	0.464	0.29	0.772

*Message: s.e.s are based on dispersion parameter with value 1.*

**Generalized Linear Models**

Available Data: Dose, Drug, Logdose, N, R  
 Analysis: General Model  
 Response Variate: R  
 Maximal Model: Logdose \* Drug  
 Model to be Fitted: Logdose \* Drug  
 Distribution: Binomial  
 Link Function: Probit  
 Binomial Totals: N

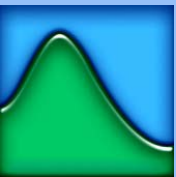
**Spreadsheet [Drug.gsh]**

Row	Drug	Dose	N	R
1	Morphine	1.5	103	19
2	Morphine	3	120	53
3	Morphine	6	123	83
4	Amidone	1.5	60	14
5	Amidone	3	110	54
6	Amidone	6	100	81
7	Phenadoxone	0.75	90	31
8	Phenadoxone	1.5	80	54
9	Phenadoxone	3	90	80
10	Pethidine	5	60	13
11	Pethidine	7.5	85	27

# Algorithm

- algorithm (see McCullagh & Nelder, 1983, §2.5)
  - initial estimate  $\eta_0 = g(y)$
  - adjusted dependent variable  $z = \eta_0 + (y - \mu_0) (d\eta/d\mu)_0$
  - (iterative) weight  $W$  given by  $W^{-1} = (d\eta/d\mu)^2_0 V(\mu)_0$
  - estimate  $\beta = (X^t\phi WX)^{-1} X^t\phi Wz$
  - recalculate  $\mu_0$ , and repeat until convergence
  - Fisher scoring algorithm – variant of Newton-Raphson optimization
- results
  - $\text{var}(\beta) = \phi(X^tWX)^{-1}$
  - deviance =  $-2\phi \{ \log L(\mu; y) - \log L(y; y) \}$
- so a GLM is characterized by its
  - link function  $g(\mu)$
  - derivative of link function  $d\eta/d\mu$
  - variance function  $V(\mu)$
  - deviance function  $d(y; \mu)$

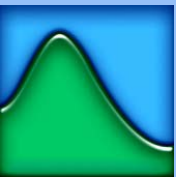
..



# Quasi-likelihood

- Wedderburn, R.W.M. (1974) Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, 61, 439-447.
- legitimizes *heterogeneity factor* of probit analysis
- characterized by same 4 properties as a GLM
  - assume independent responses  $y_i$  with  $E(y) = \mu = g^{-1}(\eta)$
  - variance function  $V$  and deviance function  $d$
- definition (Mc & N Ch9; LNP Ch3)
  - quasi-(log)likelihood is a function  $q(\mu; y)$  whose derivative  $U$  has form
  - $U = \partial q(\mu; y) / \partial \mu = (y - \mu) / \phi V(\mu)$
- following properties shared with log-likelihood derivative
  - $E(U) = 0$ ,  $\text{var}(U) = 1 / \{\phi V(\mu)\}$ ,  $-E(\partial U / \partial \mu) = 1 / \{\phi V(\mu)\}$
- if  $Q$  integral of  $U$  exists, it behaves like a log-likelihood fn
  - *quasi-deviance*  $d(\mu; y) = -2\phi Q(\mu; y) = -2 \int_{\mu}^y (y-u)^2 / V(u) du$
- estimation again by iteratively reweighted least squares

..

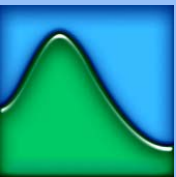


# Quasi-likelihood

- comments

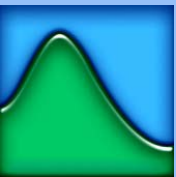
- makes assumptions only on 1<sup>st</sup> and 2<sup>nd</sup> moments
- is a true likelihood iff response  $y$  is from the exponential family
  - with log-density  $q(\mu; y) = \theta y - b(\theta) + c(y)$
  - where  $\mu = b'(\theta)$ ,  $V(\mu) = b''(\theta)$
- for 1<sup>st</sup> order inferences (e.g. asymptotic Normality of estimates) QL behaves "largely" like a true likelihood
  - further info, see Lee & Nelder (1999) *Canadian J. Stats*, 27
- is essentially an estimating equations approach
  - and, if you replace the variance with a general covariance matrix of responses, you get the method of Generalized Estimating Equations
- Wedderburn's estimation of mean  $\mu$  assumes  $\varphi$  known
  - $\varphi$  can be estimated separately by method of moments as  $1/(n-p) \sum \{(y-\mu)^2/V(\mu)\}$
  - essentially this uses a pseudo-likelihood approach
- a proper framework for estimation of  $\varphi$  requires *extended quasi-likelihood*

..



# Extended quasi-likelihood

- Nelder, J.A. & Pregibon, D. (1987). *Biometrika*, 74, 221-232
- need function  $Q(\mu, \sigma^2; y) = q(\mu; y) + h(\sigma^2; y)$  that
  - operates as a quasi-likelihood for known  $\sigma^2$
  - has properties of a log-likelihood with respect to  $\sigma^2$
- derivation (N & P, 1987; McC & N §9.6; LNP §3.5)
  - $Q(\mu, \sigma^2; y) = -1/2 d(\mu; y) / \sigma^2 + h(\sigma^2; y)$
  - suppose  $h(\sigma^2; y) = -1/2 h_1(\sigma^2) - h_2(y)$
  - need  $E(\partial Q / \partial \sigma^2) = 0$ , so  $1/2 E\{d(\mu; y)\} / (\sigma^2)^2 - 1/2 h_1'(\sigma^2) = 0$
  - i.e.  $(\sigma^2)^2 h_1'(\sigma^2) = E\{d(\mu; y)\}$
  - 1<sup>st</sup> order approximation  $E\{d(\mu; y)\} = \sigma^2$ , so  $h_1(\sigma^2) = \log(\sigma^2) + \text{const.}$
- thus  $Q(\mu, \sigma^2; y)$ 
  - $= -1/2 d(\mu; y) / \sigma^2 - 1/2 \log(\sigma^2) + \text{const}$
  - $= -1/2 d(\mu; y) / \sigma^2 - 1/2 \log(\sigma^2 V(y))$
- in effect EQL treats deviance contributions as  $\sigma_i^2 \times \chi^2_1$  variables (= gamma with mean  $\sigma_i^2$  & variance  $2\sigma_i^2$ )
- cautionary note
  - bias may occur as  $\int \exp(Q) dy$  may require a normalizing const  $c \neq 1$
  - so GenStat HG procs use exact likelihood, where feasible, instead



# Modelling mean & variance

- 2 interlinked *joint GLMs* (fitted by MD system, Nelder 199?)
- fit alternately until convergence (LNP §3.6)

Table 3.2 *GLM attributes for joint GLMs*

Components	$\beta$ (fixed)	$\gamma$ (fixed)
Response	$y$	$d^*$
Mean	$\mu$	$\phi$
Variance	$\phi V(\mu)$	$2\phi^2$
Link	$\eta = g(\mu)$	$\xi = h(\phi)$
Linear Pred.	$X\beta$	$G\gamma$
Dev. Comp.	$d$	$\text{gamma}(d^*, \phi)$
Prior Weight	$1/\phi$	$(1 - q)/2$

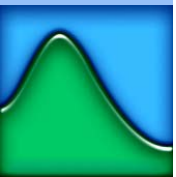
$$d_i = 2 \int_{\hat{\mu}_i}^y (y - s) / V(s) ds,$$

$$d^* = d / (1 - q),$$

$$\text{gamma}(d^*, \phi) = 2 \{ -\log(d^*/\phi) + (d^* - \phi)/\phi \},$$

This gives the EQL procedure if  $q = 0$ , and

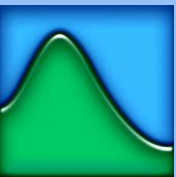
the REML procedure if  $q$  is the GLM leverage (Lee and Nelder, 1998).



# Prediction

- 1982
  - Lane & Nelder (1982) Analysis of covariance and standardization as instances of prediction. *Biometrics*, 38, 613-621
  - get predictions that seem more "natural" if you form fitted values for all factor combinations and average them on back-transformed scale
- 1983 – PREDICT directive Genstat Release 4.04
- 1985
  - Morris, G.E.L. (1985) The presentation of treatment responses from block experiments after analysis of variance of transformed data. *Annals of Applied Biology*, 107, 571-580
  - in a randomized-block experiment, back-transformed means predict treatment values that would be expected in absence of block effects, means of back-transformed values predict values expected if each treatment had been applied over all the plots
- 1989 – BACKTRANSFORM option Genstat 5 Release 2
- 2004 – VPREDICT directive 7<sup>th</sup> Edition
- 2006 – HGPREDICT procedure 9<sup>th</sup> Edition

..

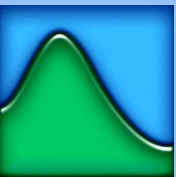


# Multiple sources of error variation

- e.g. split-plot
  - Yates (1935) *JRSS suppl*, 2, 185-247.
  - oats example from Yates (1937) *TC35*
- designed in 2 stages
  - divide each block into 3 *whole plots* & randomly allocate varieties (V1-3)
  - divide each *whole plot* into 4 *sub-plots* and randomly apply nitrogen (N0-3)
- different error terms needed
  - varieties compared with whole-plots
  - nitrogen (& variety × nitrogen interaction) compared with sub-plots
- generally balanced design
  - Nelder (1965) *Proc. Royal Soc, A*, 283, 147-178.
  - Nelder (1968) *Biometrika*, 54, 449-460.
  - Payne & Tobias (1992) *Scand. J. Stats.*, 19, 3-23.

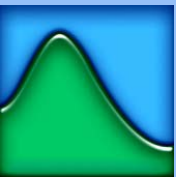
V3 N3	V3 N2	V3 N2	V3 N3
V3 N1	V3 N0	V3 N0	V3 N1
V1 N0	V1 N1	V2 N0	V2 N2
V1 N3	V1 N2	V2 N3	V2 N1
V2 N0	V2 N1	V1 N1	V1 N2
V2 N2	V2 N3	V1 N3	V1 N0
V3 N2	V3 N0	V2 N3	V2 N0
V3 N1	V3 N3	V2 N2	V2 N1
V1 N3	V1 N0	V1 N2	V1 N3
V1 N1	V1 N2	V1 N0	V1 N1
V2 N1	V2 N0	V3 N2	V3 N3
V2 N2	V2 N3	V3 N1	V3 N0
V2 N1	V2 N2	V1 N2	V1 N0
V2 N3	V2 N0	V1 N3	V1 N1
V3 N3	V3 N1	V2 N3	V2 N2
V3 N2	V3 N0	V2 N0	V2 N1
V1 N0	V1 N3	V3 N0	V3 N1
V1 N1	V1 N2	V3 N2	V3 N3

..



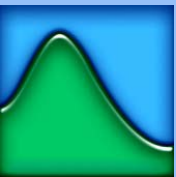
# Multiple sources of error variation

- non-Normal data that are subject to sources of error variation present particular challenges
  - counts or binomial proportions from split-plot experiments
  - continuous non-Normal data from split-plot experiments
    - e.g. Cochran & Cox (1957, *Experimental Designs*, 2<sup>nd</sup> ed) cake data (see Lee, Nelder & Pawitan, 2006, Section 6.4.1)
  - time-course experiments with treatments applied to subjects (and possibly at times within subjects)
  - survey-type data with random terms (e.g. factors with levels selected from an infinite population)
- methods include ANOVA or REML (with transformations), GEEs, GLMMs, HGLMs, HGNNLMs, DHGLMs ...
- 



# Generalized linear mixed models

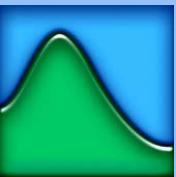
- Schall, R. (1991). Estimation in GLMs with random effects. *Biometrika*, 78, 719-727.
  - Gilmour, Anderson & Rae (1985) *Biometrika*, 72, 593-599.
  - Breslow, N.E. & Clayton, D.G. (1993). *JASA*, 88, 9-25.
- generalized linear model  $E(y) = \mu$   
model defines the *linear predictor*  $\eta = X \beta$   
 $\eta$  is related to  $\mu$  by the *link function*  $g()$   $\eta = g(\mu)$   
 $y$  has distribution from the *exponential family* (mean  $\mu$ )
- generalized linear mixed model  $E(y) = \mu$   
model defines the *linear predictor*  $\eta = X \beta + \sum_i Z_i v_i$   
containing additional vectors of random effects  $v_i$   
 $v_i$  have Normal distributions with mean 0 and variance  $\sigma_i^2$
- fitted by iteratively reweighted REML
- procedure GLMM into Genstat 5 R3.2 by Sue Welham 1992
- 



# Hierarchical generalized linear models

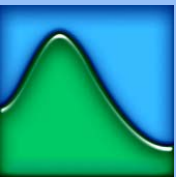
- 1994
  - Nelder, J.A. & Lee, Y., Double generalized linear models. Wagga Wagga conference (abstracts p.24) "class includes GLMMs and conjugate DGLMs", uses h-likelihood "avoids integration necessary for the use of marginal likelihood.."
- 1996
  - Lee & Nelder, Hierarchical generalised linear models, *JRSS B*, 58
  - Nelder, J.A., Construction of sub-systems in GENSTAT. Adelaide conference (abstracts p.21) describes K and HG systems
- 1999
  - Nelder & Lee, Generalized REML via GLMs. Lorne conference (abstracts p.11) algorithm for fitting HGLMs by 2 interconnected GLMs
- 2001
  - Lee & Nelder, HGLMs: synthesis of GLMs, random-effect models and structured dispersion, *Biometrika*, 88, 987-1006
  - Lee & Nelder, Modelling and analysing correlated non-normal data. *Statistical Modelling*, 1, 3-16
  - Nelder, Payne & Candy, Extensions of GLMs for Modelling Discrete Data. Workshop at Surfers Paradise conference

..



# Hierarchical generalized linear models

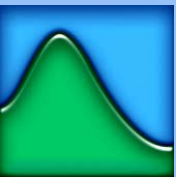
- expected value  $E(y) = \mu$   
link  $\eta = g(\mu)$   
and distribution from *exponential family* (Normal, Binomial, Poisson or gamma as before)
- but linear predictor  $\eta = X\beta + \sum_i Z_i V_i$   
now contains additional vectors of random effects  $V_i$  with either Normal, beta, gamma or inverse gamma distributions – and with their own link functions
  - Normal-identity gives a GLMM but HGLM algorithms use much improved Laplace approximations in their use of adjusted profile likelihood to address problems of small data sets
- references
  - Lee, Y. & Nelder, J.A. (1996, 2006).
  - Lee, Y., Nelder, J.A. & Pawitan, Y. (2006). *Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood*. Chapman & Hall.
  - Payne (2011) *Guide to Regression, Nonlinear and Generalized Linear Models in GenStat*.



# Implementation

- as in ordinary linear mixed models, we need to
  - estimate effects for the treatment (fixed) terms, and residuals for the random (blocking) terms
  - estimate variance components for the random terms
- GLMMs
  - use REML linear mixed models algorithm with weights being recalculated each time it re-estimates the means (i.e. estimate means and variance components by REML, (re)calculate weights, re-estimate means etc...)
- HGLMs
  - iterates between 2 interlinked generalized linear models, to
    1. estimate effects and residuals (with weights based on variance components)
    2. estimate the variance components (using residuals from GLM 1: note GLM 2 is a gamma-log GLM so variance components must be positive)
  - allows you to "model" the dispersion (variance components can vary over the experiment according to other factors e.g. location)
  - can use familiar GLM model-checking techniques to assess each aspect
  - produces likelihood goodness of fit statistics for each aspect

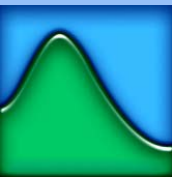
..



# HGLM algorithm

Components	$\beta$ (fixed)	$\gamma$ (fixed)
Response	$y$	$d^*$
Mean	$\mu$	$\phi$
Variance	$\phi V(\mu)$	$2\phi^2$
Link	$\eta = g(\mu)$	$\xi = h(\phi)$
Linear Pred.	$X\beta + Zv$	$G\gamma$
Dev. Comp.	$d$	$\text{gamma}(d^*, \phi)$
Prior Weight	$1/\phi$	$(1 - q)/2$

Components	$u$ (random)	$\lambda$ (fixed)
Response	$\psi_M$	$d_M^*$
Mean	$u$	$\lambda$
Variance	$\lambda V_M(u)$	$2\lambda^2$
Link	$\eta_M = g_M(u)$	$\xi_M = h_M(\lambda)$
Linear Pred.	$v$	$G_M \gamma_M$
Deviance	$d_M$	$\text{gamma}(d_M^*, \lambda)$
Prior Weight	$1/\lambda$	$(1 - q_M)/2$



# Example – Cochran & Cox cakes

Spreadsheet [Cake.gsh]

Row	Replicate	Batch	Recipe	Temperature	Angle
1	1	1 1	1	175	42
2	1	1 1	1	185	46
3	1	1 1	1	195	47
4	1	1 1	1	205	39

D:\G5\HGLM\Nelder\cakes.gen\*

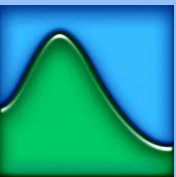
```
open 'master/ksys.gpi'; 2 : input 2
kinit
kload 'master/hgsys.gpi'
hginit
kload 'cakes.dat'
hgopen
hgsum g;r;r*t;;;bl+mp
hgsud
hgfit[zed=10]
hgpe
hgpe[d]
hgfa1
hgmc[d][m] 1
summ y
kerr b;!(270(500))
hgst
hgfm
hgfd
hgst
```

- breaking angle of chocolate cakes (LNP §5.5)
- split plot: Replicate/Batch/Cake
- treatment factors: Recipe (whole-plot factor, between Batches), Temperature (sub-plot factor, within Batches)
- gamma GLMM (§6.6.5)

# HGLMs in GenStat 14<sup>th</sup> Edition

- procedures (Payne, Lee, Nelder & Noh 2011)
  - HGFIXEDMODEL – defines the fixed model for an HGLM or DHGLM
  - HGRANDOMMODEL – defines the random model for an HGLM
  - HGDRANDOMMODEL – adds random terms into the dispersion models of an HGLM, so that the whole model becomes a DHGLM
  - HGNONLINEAR – defines nonlinear parameters for the fixed model
  - HGANALYSE – fits a hierarchical generalized linear model (HGLM) or a double hierarchical generalized linear model (DHGLM)
  - HGDISPLAY – displays results from an HGLM or DHGLM
  - HG PLOT – produces model-checking plots for an HGLM or DHGLM
  - HGPREDICT – forms predictions from an HGLM or DHGLM analysis
  - HGGRAPH – plots predictions from an HGLM or DHGLM analysis
  - HGKEEP – saves information from an HGLM or DHGLM analysis
  - HGWALD – gives Wald tests for fixed terms that can be dropped
  - HGFTEST – calculates likelihood tests for fixed terms
  - HGRTEST calculates likelihood tests for random terms
- menus
  - Stats | Regression Analysis | Mixed Models | Hierarchical Generalized Linear Models
  - cover the standard situations, but not dispersion modelling nor DHGLMs

..



# GLMM for cakes

**Hierarchical Generalized Linear Models**

Available Data: Angle, Batch, Recipe, Replicate, Temperature

Response Variate: Angle

Fixed Model

Model: Recipe \* Temperature

Distribution: Gamma

Link Function: Logarithm

Binomial Totals:

Random Model

Model: Replicate / Batch

Distribution: Normal

Link Function: Identity

Run Options... Save... Further Output...  
Cancel Defaults Predict...

- like the ordinary *Linear Mixed Models* menu
  - extra box to select the *Distribution* of the basic GLM
  - extra box to select the *Link Function* of the basic GLM
  - extra box to specify the binomial totals
  - extra box to select the *Distribution* of the random terms
  - extra box to select the *Link Function* of the random terms

# Output

## Hierarchical generalized linear model

Response variate: Angle

### Mean model

Fixed terms: Recipe\*Temperature  
Distribution: gamma  
Link: logarithm  
Random terms: Replicate/Batch  
Distribution: normal  
Link: identity  
Dispersion: free

### Dispersion model

Distribution: gamma  
Link: logarithm

### Estimates from the mean model

	estimate	s.e.	t(252)
constant	3.3537	0.05975	56.127
Recipe 2	-0.0787	0.05618	-1.402
Recipe 3	-0.0540	0.05618	-0.962
Temperature 185	0.0933	0.05055	1.845
Temperature 195	0.0614	0.05055	1.215
Temperature 205	0.1491	0.05055	2.950
Temperature 215	0.2874	0.05055	5.686
Temperature 225	0.1928	0.05055	3.815
Recipe 2 .Temperature 185	-0.0094	0.07149	-0.131
Recipe 2 .Temperature 195	0.0978	0.07149	1.368
Recipe 2 .Temperature 205	0.0286	0.07149	0.399
Recipe 2 .Temperature 215	-0.0495	0.07149	-0.693
Recipe 2 .Temperature 225	0.0819	0.07149	1.145
Recipe 3 .Temperature 185	-0.0512	0.07149	-0.717
Recipe 3 .Temperature 195	0.0745	0.07149	1.042
Recipe 3 .Temperature 205	-0.0389	0.07149	-0.544

← regression (GLM) style output

← descriptions of models

← parameter estimates (here just for fixed effects)

..

# Output

Recipe 3 .Temperature 195	0.0745	0.07149	1.042
Recipe 3 .Temperature 205	-0.0389	0.07149	-0.544
Recipe 3 .Temperature 215	-0.0649	0.07149	-0.907
Recipe 3 .Temperature 225	0.0576	0.07149	0.806

## Estimates from the dispersion model

← dispersion model estimates variance components (on the log scale)

## Estimates of parameters

Parameter	estimate	s.e.	t(**)	antilog of estimate
phi	-3.9581	0.0947	-41.79	0.01910
lambda Replicate	-3.514	0.394	-8.92	0.02978
lambda Replicate.Batch	-5.406	0.343	-15.78	0.004492

*Message: s.e.s are based on dispersion parameter with value 1.*

## Likelihood statistics

-2 × h(y v)	1508.533
-2 × h	1352.777
-2 × P <sub>v</sub> (h)	1616.117
-2 × P <sub>β,v</sub> (h)	1697.849
-2 × EQD(y v)	1507.674
-2 × EQD	1351.917
-2 × P <sub>v</sub> (EQD)	1615.258
-2 × P <sub>β,v</sub> (EQD)	1696.989

← likelihood statistics allow you to assess various aspects of the model

← fixed model by  $-2 \times P_v(h)$

← dispersion models by  $-2 \times P_{\beta,v}(h)$

← (EQD statistics are alternatives, calculated a different way)

Fixed parameters in mean model	18
Random parameters in mean model	60
Fixed dispersion parameters	3
Random dispersion parameters	0

..

# Is there an interaction?

## Wald tests for dropping HGLM fixed terms

Term	Wald statistic	d.f.	approx. pr.
Recipe.Temperature	9.328	10	0.501

89 HGFTEST

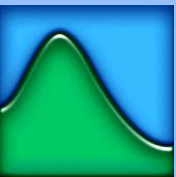
## Likelihood tests for dropping HGLM fixed terms

Term	Test statistic	d.f.	pr.
Recipe.Temperature	9.321	10	0.502

← **Wald test**

← **or more accurate test, using change in  $-2 \times P_v(h)$  (HGFTEST refits model without interaction but with dispersion estimates from original model)**

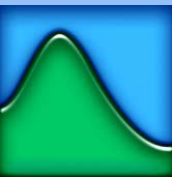
..



# Technical comments

- in GLMMs fixed effects are estimated by numerical integration of the likelihood over distributions of the random effects
  - standard REML-based algorithms use Laplace approximations
    - can be unreliable (biased) with small or binary (0/1) data sets
  - Bayesian algorithms use Markov chain Monte Carlo (MCMC) methods
    - require expertise to drive, can be computationally expensive
- with some HGLMs the likelihood factorizes
  - so no numerical integration or approximations are required
  - these are conjugate HGLMs
    - Normal-Normal, Poisson-gamma, binomial-beta, gamma-inverse gamma
    - random effect distribution is *conjugate* to the basic GLM distribution
  - intuitively and mathematically appealing
- improved Laplace approximations are available when HGLM algorithm is used to fit a GLMM, which reduce biases
  - see Salamander example

..



# DHGLMs (double HGLMs)

	Aug. GLM	GLM	Aug. GLM	GLM
Components	$\beta$ (fixed)		$\gamma$ (fixed)	
Response	$y$		$d^*$	
Mean	$\mu$		$\phi$	
Variance	$\phi V(\mu)$		$2\phi^2$	
Link	$\eta = g(\mu)$		$\xi = h(\phi)$	
Linear Pred.	$X\beta + Zv$		$G\gamma + Fb$	
Dev. Comp.	$d$		$\text{gamma}(d^*, \phi)$	
Prior Weight	$1/\phi$		$(1 - q)/2$	
Components	$u$ (random)	$\lambda$ (fixed)	$a$ (random)	$\alpha$ (fixed)
Response	$\psi_M$	$d_M^*$	$\psi_D$	$d_D^*$
Mean	$u$	$\lambda$	$a$	$\alpha$
Variance	$\lambda V_M(u)$	$2\lambda^2$	$\alpha V_D(a)$	$2\alpha^2$
Link	$\eta_M = g_M(u)$	$\xi_M = h_M(\lambda)$	$\eta_D = g_D(a)$	$\xi_D = h_D(\alpha)$
Linear Pred.	$v$	$G_M \gamma_M$	$b$	$G_D \gamma_D$
Deviance	$d_M$	$\text{gamma}(d_M^*, \lambda)$	$d_D$	$\text{gamma}(d_D^*, \alpha)$
Prior Weight	$1/\lambda$	$(1 - q_M)/2$	$1/\alpha$	$(1 - q_D)/2$

- Lee & Nelder (2006) extended HGLMs to become *double hierarchical generalized linear models* by allowing random terms in the linear predictor of the GLMs used to model the dispersion parameter  $\phi$  (so this, itself is now an HGLM)

$$d_i = 2 \int_{\mu_i}^y (y - s) / V(s) ds,$$

$$d_{Mi} = 2 \int_{u_i}^{\psi_M} (\psi_M - s) / V_M(s) ds,$$

$$d_{Di} = 2 \int_{a_i}^{\psi_D} (\psi_D - s) / V_D(s) ds,$$

$$d^* = d / (1 - q_0),$$

$$d_M^* = d_M / (1 - q_M),$$

$$d_D^* = d_D / (1 - q_D),$$

$$\text{gamma}(d^*, \phi) = 2\{-\log(d^*/\phi) + (d^* - \phi)/\phi\} \text{ and}$$

$(q, q_M, q_D)$  are leverages as described in Section 7.2.2.

..

# Conclusion

- Generalized Linear Models show Rothamsted stats at its best
  - identifying and solving real and challenging practical problems
  - providing a solution that includes an implementation (as well as published papers)
  - having applications that go way beyond the original problem
- they will perhaps be John's most enduring legacy



John Ashworth Nelder

8 October 1924 – 7 August 2010

<http://www.vsni.co.uk/2010/home-pages/john-nelder/>

