The computing you need

(as a postgraduate student in mathematical sciences)

Peter Green

Attitude

- Here to do statistics or mathematics
- Computer is a tool
- A modest collection of computing skills can
  - assist your efficiency, and
  - improve the accuracy and presentation of your work
  - stimulate your thinking
- Transferable skills

Plan

- Email/WWW – assumed
- Systems: Windows, Linux, Cygwin
- Editors
- Languages
- Scientific documents: LaTeX
- Web pages: HTML
- Presentations
- Computer support

Systems for your desktop PC

- Windows
- Linux on farm/Beowulf/SCONE
- Linux on desktop
- Windows + Cygwin (`www.cygwin.com`)

Text editors

- Windows
  - Notepad
  - Notetab
  - Emacs
- Linux
  - Vi
  - Emacs

→ pick one, become expert, and use it for everything!

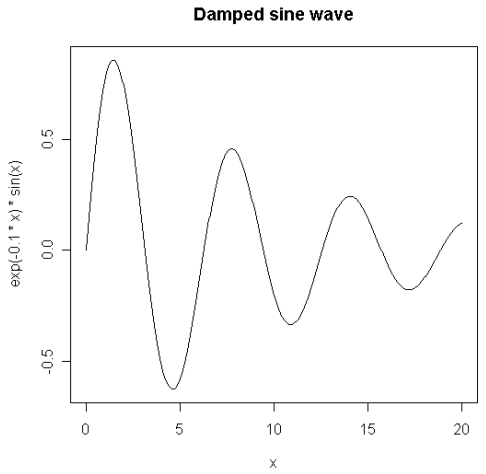
Languages

- High-level/'scripting'
  - R
  - Matlab
  - Python
  - Perl
- 'Low'-level
  - Fortran
  - C, C++
- Symbolic
  - Maple

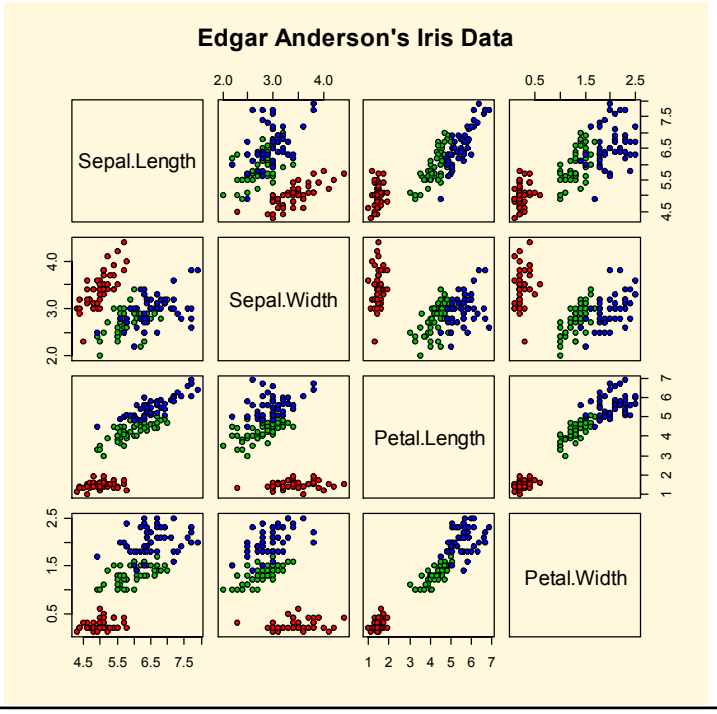
R (www.stats.bris.ac.uk/R)

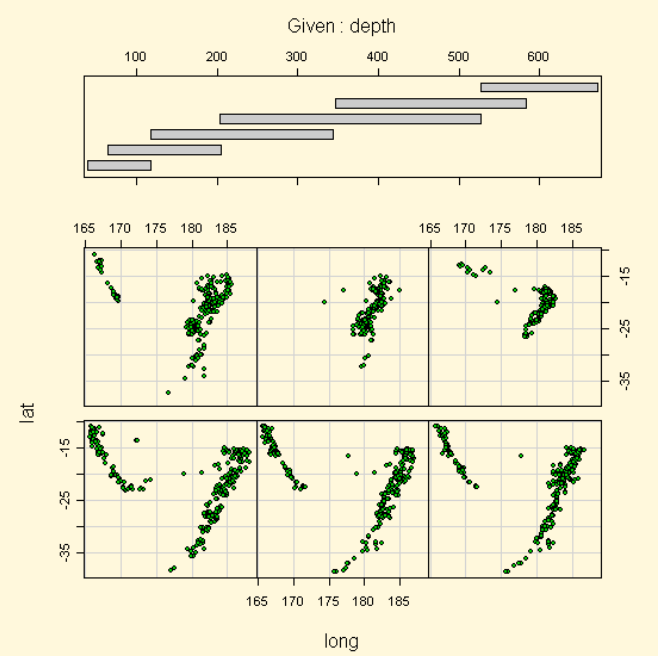
- (Programmable) 'calculator'
- Graphics 'calculator'
- Produces publication quality graphics
- Programming language
- Statistics system
- Tool for post-processing results
- Extendable


```
> 2+3
[1] 5
> sum((1:1000)^(-2))
[1] 1.643935
> pi^2/6
[1] 1.644934
```




```
> x<-seq(0,20,len=401)
> plot(x,exp(-0.1*x)*sin(x),type='l')
> title('Damped sine wave')
```


Edgar Anderson's Iris Data
Sepal.Length
Sepal.Width
Petal.Length
Petal.Width


Given : depth
lat
long

The ‘3k+1’ problem

14,7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1

```
nmax<-100

n<-rep(-1,nmax)
n[1]<-0
for(i in 2:nmax)
{
       ni<-0
       j<-i
       repeat
       {
              if(j%%2==0) j<-j/2 else j<-3*j+1
              ni<-ni+1
              if(j<=nmax && n[j]>=0) {n[i]<-n[j]+ni; break}
       }
}
```


The ‘3k+1’ problem

```
> n
  [1]   0   1   7   2   5   8  16   3  19   6  14
 [12]   9   9  17  17   4  12  20  20   7   7  15
 [23]  15  10  23  10 111  18  18  18 106   5  26
 [34]  13  13  21  21  21  34   8 109   8  29  16
 [45]  16  16 104  11  24  24  24  11  11 112 112
 [56]  19  32  19  32  19  19 107 107   6  27  27
 [67]  27  14  14  14 102  22 115  22  14  22  22
 [78]  35  35   9  22 110 110   9   9  30  30  17
 [89]  30  17  92  17  17 105 105  12 118  25  25
[100]  25
```

Fortran, C, C++

- Compiled general-purpose languages like these are vastly faster than interpreted languages
- Essential for serious computer-intensive applications
  - large systems of differential equations
  - simulation, etc
- Good strategy – combine with scripting languages (prototyping, as front-end, for post-processing…)

Symbolic computing: Maple

```
> diff( x^2+x^4-3*x+2, x );
```

$$2x + 4x^3 - 3$$

```
> int( sin(x), x );
```

$$-\cos(x)$$

```
> limit( sin(x) / x, x = 0 );
```

$$1$$

Scientific typesetting: LaTeX

- Strongly discourage Word, etc
- LaTeX
  - Mark-up language
  - Professional quality
  - Tables, figures, cross-referencing, indexing, custom styles…
  - Free
  - Custom editors if you need them

Plain text

Part of my online demo document:

```
This is a simple example \LaTeX\ document, demonstrating
plain text, mathematical notation (both inline and
displayed), definition and use of a `newcommand',
a figure, a table and some displayed program text.

Integrating out $w$, we find
```

This is a simple example LaTeX document, demonstrating plain text, mathematical notation (both inline and displayed), definition and use of a 'newcommand', a figure, a table and some displayed program text.

Integrating out w, we find

A formula

```
\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx =
\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}
```

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

A table

```
\begin{table}[h]
\caption{This is a simple table.\label{dtba}}
\vspace*{5mm}\centering\leavevmode
\begin{tabular}{|c|r|}
\hline
$x$ & $\exp(-0.3x)\cos(x)$ \\
\hline
0 & 1.0000 \\
2 & $-0.2284$ \\
4 & $-0.1969$ \\
6 & 0.1587 \\
\hline
\end{tabular}
\end{table}
```

Table 1: This is a simple table.

x	$\exp(-0.3x)\cos(x)$
0	1.0000
2	-0.2284
4	-0.1969
6	0.1587

Web pages & HTML

- HyperText Markup Language
- Custom editors if you need them
- Please make a web page – NOW!
- **www.stats.bris.ac.uk/~YOU** or
- **www.maths.bris.ac.uk/~YOU** is an external address for your directory **public_html** on Hyperion
- Files in this directory and subdirectories are interpreted and displayed by your customer's browser
- You can learn a lot from View | Source

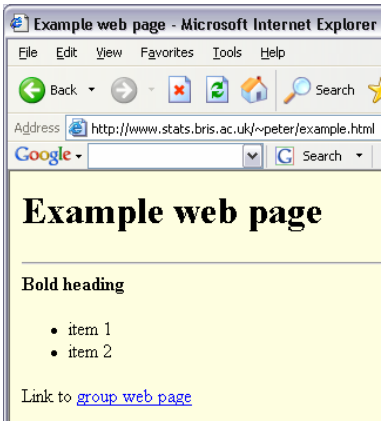

```
<html>

<head>
<title>Example web page</title>
</head>

<body bgcolor="lightyellow">
<h1>Example web page</h1>
<hr>
<b>Bold heading</b>
<ul>
<li>item 1
<li>item 2
</ul>

Link to <a href="http://www.stats.bris.ac.uk/">group web
page</a>
</body>

</html>
```



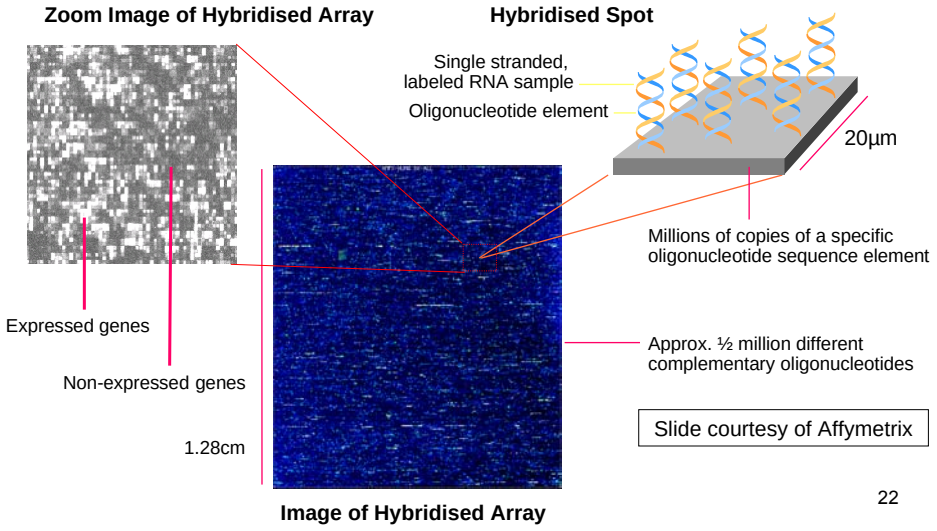
Presentations

- Overheads
- Data projector

- PowerPoint
- Acrobat
  - LaTeX → pdf

Gene expression using Affymetrix chips

Powerpoint



Slide courtesy of Affymetrix

LaTeX/pdf

The Dirichlet process - view 1

Given a probability distribution G_0 on an arbitrary measure space Ω, and a positive real α, we say the random distribution G on Ω follows a Dirichlet process,

$$G \sim DP(\alpha, G_0)$$

if for all partitions $\Omega = \bigcup_{j=1}^{m} B_j$ ($B_j \cap B_k = \emptyset$ if $j \neq k$), and for all m,

$$(G(B_1), \ldots, G(B_m)) \sim \text{Dirichlet}(\alpha G_0(B_1), \ldots, \alpha G_0(B_m))$$

Even if G_0 is continuous, G is a.s. discrete, so i.i.d. draws $\{\theta_g, g = 1, 2, \ldots, n\}$ from G exhibit ties.

Computer support

- support-maths@bristol.ac.uk
- Peter, Bob and Hugo
  - They have a huge workload!
  - Courtesy helps
  - If you have broken the rules, how tolerant do you expect them to be?
- Hardware: don't dismantle your PC!
- Systems: you have a balance of freedom & responsibility
- Programming: learn it yourself!

Sources

www.stats.bris.ac.uk/~peter/sources.html

- Demo LaTeX document
- Demo LaTeX slides
- Demo of making pdf slides from LaTeX, for display with data projector
- Demo web page
- Information about R
- Link to Cygwin
- Links to Library, Web of science, ...