Instructions for Using

First Bayes

Version 1.0

May 1994

CONTENTS

CHAPTER ONE. INTRODUCTION TO FIRST BAYES

1.1 What is First Bayes?

First Bayes is a programe intended to help with teaching and learning elementary Bayesian Statistics. It deals with quite simple and standard statistical models, with an emphasis on obtaining some understanding of how the Bayesian approach works. It is *not* a package for *doing* statistical analysis of practical data.

1.2 Main features of First Bayes

- Three standard one parameter models are offered—binomial data, Poisson data and normal data with known variance. A major feature of First Bayes is that such data may be analysed using an arbitrary mixture of distributions from the conjugate family. By this means, essentially arbitrary prior distributions can be defined, allowing the user to obtain an excellent understanding of how the likelihood and prior distribution are combined by Bayes' Theorem. Prior, likelihood and posterior can be plotted on a single "triplot".

- Distributions from any of thirteen standard families may be plotted and many useful summaries computed—mean, mode(s), median, quartiles (or any other quantiles), standard deviation, highest density intervals and probabilities of arbitrary intervals. Furthermore, arbitrary mixtures of distributions from any family can be defined and examined in the same way. Another key feature of First Bayes is that all these plots and summaries are available whenever such distributions arise in First Bayes. This includes prior, posterior and predictive distributions in analysis of one-parameter models, and corresponding one-dimensional marginal distributions in analysis of more complex models.

- Analysis of two simple kinds of linear model are also offered. One is the case of one or more normal samples with common but unknown variance (one-way analysis of variance), and the other is simple linear regression. Marginal distributions may be computed (and examined) for arbitrary linear combinations of the location parameters. In the case of regression, scatter and residual plots can be produced.

- Predictive distributions are available in a variety of forms for all analyses.

- For all analyses, the current posterior distribution can be turned into the prior distribution in order to add further data sequentially.

- Whenever the user makes changes to data or prior distributions, all analyses, distribution summaries and plots are automatically updated.

1.3 Genesis of First Bayes

First Bayes began life as a non-Windows program called MSB. I wrote MSB at Warwick University in 1987 in connection with an undergraduate course in elementary Bayesian statistics. The Windows user interface and a number of other features were added in 1994 after I moved to Nottingham University.

MSB was originally written in the APL language, and one impetus for writing First Bayes was the fact that powerful Windows-based development tools became available for APL. If you don't enjoy programming, you have probably never used APL. If you do enjoy programming but haven't used APL you are missing a real treat. If you are an APL user, you know what I mean! I discovered APL in 1976, and have been hooked ever since. I have used many other languages, but nothing compares with APL for power and elegance—APL is a joy to program in. But more than that, it provides such a wonderful development environment that complex applications can be built and debugged incredibly quickly.

All this discussion of APL is not just gratuitous advertising, but is relevant to the story of First Bayes. MSB was distributed by Warwick University as part of a grant I obtained from the Computer Board under the UK Government's "Computers in Teaching Initiative". The grant was to explore the use of APL in teaching statistics, and the development of MSB was as much a response to that initiative as an aid to the course I was teaching in Bayesian Statistics.

The graphics in MSB rested on a system, the Warwick APL Graphics System (WAGS), which was also distributed as part of that initiative. WAGS was originally written by Ewart Shaw at Warwick University. It was further developed and rewritten by Tuks Gillespie and myself in 1990 as part of another initiative, the APL Statistics Library (ASL). ASL was supportd by the British APL Association. In creating First Bayes in 1994, I further developed WAGS, porting it to the Windows environment.

APL is an interpreted language, and is not compiled. As a result any application written in APL needs an APL interpreter to run it. APL is a complex language and interpreters are not cheap. Another impetus for writing First Bayes was that the new Windows version of APL allowed applications to be distributed with a free "run-time only" interpreter. This has the advantage of allowing First Bayes to be distributed as a completely self contained package at no charge, which was never possible with MSB.

First Bayes uses the "Dyalog APL/W" APL interpreter, which is produced by Dyadic Systems Ltd. The file DYALOGRT.EXE on the First Bayes disk is the run-time version of Dyalog APL/W. The DYALOGRT.EXE supplied with First Bayes version 1.0 runs version 6.3 of Dyalog APL/W. If, however, you have the full Dyalog APL interpreter, version 6.3 or later, you can use that to run First Bayes instead of DYALOGRT.EXE. Just load the First Bayes workspace (called 1STBAYES) in the usual way. In that case, as an APL programmer, you will be able to gain access to the underlying APL code. I will be happy to provide advice on using First Bayes as a native APL application, but you must *not* distribute First Bayes after you have broken into it and modified or edited it in any way. To do so would violate my copyright on the code and the name "First Bayes".

For more information about Dyalog APL/W, contact Dyadic Systems Ltd, Riverside View, Basing Road, Old Basing, Basingstoke, Hampshire RG24 0AL.

1.4 Numerical accuracy

The numerical computations in First Bayes are not intended to be highly accurate. The facility to work with arbitrary mixtures of distributions has been implemented by quite simplistic numerical integration. Probabilities for these distributions are typically computed to no more than about 3 decimal places. Highest density intervals and quantiles are only as accurate as the probabilities from which they are computed.

1.5 Disclaimer

I have tried to make First Bayes robust and reliable, but do not accept responsibility for loss or damage to data, software or hardware or any other consequence, direct or indirect of using First Bayes. Anyone using First Bayes implicitly acknowledges this disclaimer and accepts responsibility for all consequences.

CHAPTER TWO. INSTALLATION

2.1 Hardware and software requirements

First Bayes requires an IBM-PC or compatible computer running Microsoft Windows version 3.1 or later. It requires Windows to be running in 386 enhanced mode. For good performance, an 80486-DX33 or better processor is recommended. First Bayes will not run with an 80386 processor without an 80387 floating-point coprocessor. First Bayes is designed to run in 600 x 800 pixel "Super VGA" colour graphics. In monochrome or other graphics resolutions (the VGA standard 640 x 400 or even higher super VGA resolutions), First Bayes will not look at its best, and difficulties may be experienced in reading the forms and inputting data. (It is hoped that other resolutions will be catered for properly in a future version of First Bayes.) In particular, in the standard VGA resolution of 480 x 600 pixels the First Bayes distribution and analysis windows will fill the whole screen. Then it is important to be familiar with the procedure for switching between windows described in section 4.1.

2.2 Installation procedure

[The following description is for users who do not already have the Dyalog APL/W system. If you have Dyalog APL\W version 6.3 or later, you can instead just copy the file 1STBAYES.DCF into a directory on your WSPATH.]

Copy all the files from the floppy disk into an appropriate directory on your hard disk. (You might create a directory called 1b for the purpose.)

Edit the WIN.INI file in your main Windows directory (usually C:\WINDOWS) using any suitable editor (e.g. the EDIT editor supplied with recent releases of MS-DOS), to add the following lines:

```
[DyalogAPL]
dyalog=c:\1b
inifile=c:\1b\apl.ini
```

(Replace `c:\1b` with the name of the directory in which you have placed the First Bayes files if appropriate.)

In Windows, you next need to install the necessary fonts. Select the Control Panel icon, and then the Fonts icon when the control panel opens. Select the Add button. Then change the drive to a: (or whatever floppy drive you have the First Bayes disk in - an alternative is to select instead the directory c:\1b, or wherever you have copied the original files to). You should see two fonts that are on the disk - "APL font" and "Dyalog APL Normal (True Type)". Choose Select All and then OK. The two fonts should now be installed.

Finally, you need to make First Bayes accessible to Windows. In Windows, select File - New from the main Program Manager menus. Select Program Item and OK. Then type in the following text to the various boxes. Description - `First Bayes`. Command Line - `c:\1b\dyalogrt c:\1b\1stbayes`. Working Directory - `c:\1b`. (Again, replace `c:\1b` by the directory with First Bayes files in if necessary.) Now select Change icon. A new window will appear. Type `c:\1b\1b.1co` for the File name, and select OK twice to confirm this icon choice and exit this window. Select OK again to finish.

You will now have an icon labelled "First Bayes" in the current Program Manager group. (You can move it to another group or to another place in this group simply by dragging and dropping it.) Double clicking this icon will load and run First Bayes.

If you have more than 4MB of RAM available, it may be beneficial to instruct First Bayes that it can use more RAM. In the supplied file APL.INI, there is a line `maxws=2048`, telling First

Bayes that it has 2MB (2048KB) available. This can be increased generally up to the total amount of RAM less 2MB (expressed in KB).

2.3 Starting First Bayes

Once First Bayes has been installed it can be started by double-clicking on its icon, just like other Windows programs. First Bayes generally conforms to the Windows standard for keyboard and mouse usage. This manual assumes that you are familiar with Windows.

CHAPTER THREE. THE MAIN MENUS

3.1 The title, menubar and backdrop

When you start First Bayes, you will see the First Bayes title and menubar at the top of the screen, and the rest of the screen is covered by the First Bayes backdrop. It is against the backdrop that windows you create in First Bayes are placed. It is possible to have several windows open simultaneously in First Bayes, and the backdrop helps to minimise the clutter by hiding any other Windows programs that are running (like Program Manager). (See Section 4.1 for how to regain access to other programs which are hidden by the First Bayes backdrop.)

If you prefer, the backdrop may be removed by selecting the "No backdrop" option from the "Options" menu on the First Bayes menubar. The backdrop will then disappear, and the title and menubar will be reduced to half their normal size. (The First Bayes title and menubar may be moved on the screen but cannot be resized or reduced to an icon.)

The various features of First Bayes are accessed from the menubar

3.2 File Menu

Selecting the "File" menu gives two menu items. The "Data" item opens the Data Form, for defining, saving and retrieving data. See Chapter Five.

The "Exit" item exits First Bayes (closing all its windows and removing the title, menubar and backdrop). An alternative way to exit is to double-click on the top left corner of the title bar, as usual in Windows.

3.3 Distributions menu

Selecting any item from the "Distributions" menu opens a Distribution Form to set up and examine a distribution from the corresponding family, or a mixture of distributions from that family. See Chapter Six.

3.4 Analysis menu

Selecting any item from the "Analyses" menu opens an Analysis Form to carry out an analysis of the corresponding type of data. See Chapters Seven and Eight.

3.5 Options menu

Selecting the "Options" menu gives two menu items. The action of the "No backdrop" item is described in Section 3.1. Selecting the "APL Data entry" item toggles on the facility to enter and edit data in the Data Form using the native APL language. (The presence of a tick mark against this item shows that the facility is switched on.) See Sections 5.5 to 5.8.

3.6 Help menu

Version 1.0 of First Bayes offers only one item, "About First Bayes", which displays a window of information about the First Bayes package. Future versions are intended to offer genuine on-line help from this menu.

CHAPTER FOUR.   USING FIRST BAYES

4.1   Selecting windows

When several windows have been opened from the menubar, the window you may want to see or work with at any time may have become wholly or partly obscured by other windows. If it is only partly obscured, clicking on some visible part will usually bring it to the front and make it the active window.

Whether or not the window you want is partly visible, you can switch forwards or backwards through the various First Bayes windows by repeatedly pressing Ctrl-Tab or Ctrl-Shift-Tab. One exception to this rule is explained in Section 7.2.

To leave First Bayes temporarily in order to use another Windows application, pressing Alt-Tab will cycle through the various Windows applications. This is also the way to return to First Bayes when you have finished with other tasks.

4.2   Working with the active window

You work with the currently active window in the usual Windows way. Buttons are pressed by clicking on them. A field for entering or editing data is accessed by clicking in that field and then typing from the keyboard. It is important to remember that First Bayes will need to know when you have finished working in an edit field, so nothing will happen until you leave the field, or in some cases, pressing the Enter key will also cause the edited data to be acted on. Alt key shortcuts and Tab or Shift-Tab are other ways to move around the windows as usual.

4.3   Closing and minimising or maximising windows

These operations also follow the usual Windows conventions. Double clicking the top left of the window frame will close it. There is usually a Quit button to perform the same function. Clicking on the maximise or minimise button in the top right of the window frame will either reduce it to an icon or magnify it to fill the whole screen. Double clicking an icon restores the window to full size.

Graphics windows are created by pressing appropriate buttons in other windows. The graph is linked to its parent window so that if data in the parent are changed then the graph will also be updated. If the relevant data in the parent window are somehow deleted, or if the parent window is closed, it is natural to close the graphics window. However, when this happens, you are asked to confirm that you want the graphics window closed. If you answer No, the graphics window stays open but is no longer linked to its parent window. The graph is then fixed.

4.4   Printing graphics

Each graphics window has its own menubar with two menu items on it. One is a Quit "button", which will close the window. The other is a Print "button". Clicking on that will send the graph to your default printer.

CHAPTER FIVE.   DATA

5.1 Data stored on file

When you open the Data window from the menubar, the window appears as a form with three main regions. On the left is a region concerned with data stored on file. In order for stored data to become available for use in the First Bayes session they must be loaded into memory. This is done by selecting the required item of data in the list box and then pressing the Load button (beneath the box).

The selected data item may also be deleted from the file by pressing the Delete button (also beneath the Data or File list box). You will be asked to confirm the action, because data deleted from file are lost. However, see Section 5.3 for information on saving data from memory to file.

5.2 The display region

The central region of the window displays the data currently selected in either the Data on File list box or the Data Loaded list box (whichever was last selected).

Notice that a data item in First Bayes has three constituent parts. It has a name, which may be as long as you wish, and made up of any combination of letters (upper case or lower case), the numbers 0 to 9 and the underscore characters. A name may not include spaces, and must begin with a letter. A data item also has a description, which may also be as long as you wish (and may also include spaces), or can be empty. Finally of course a data item has data! A data item in First Bayes refers to a *vector* of numbers. These appear in the Data field in the central portion of the window. Although they may typically occupy several lines (and perhaps more lines than the Data field can show at a time), they are thought of as just a single line of data, wrapping round from one line to the next as necessary.

This central region is also used for editing data in memory or for defining new data. See the next Section for details.

5.3 Data loaded

The right-hand region refers to data currently loaded in memory. It has a list showing the names of the data items in memory, and an array of buttons to the right and beneath.

- Save. The Save button stores the currently selected data item to the file on disk. If an item with that name is already on file the new item replaces it. The new item now becomes the *first* item on the file, and so appears at the top of the Data on File list.

- Erase. The Erase button removes the currently selected item from memory. You cannot, however, remove the "Nul" data item, which is required always to be loaded in First Bayes to denote an empty data vector.

- Edit. The Edit button makes the central display region, showing the currently selected data item, active. You can then edit the description or data associated with that name in memory. (Data with the same name on file are *not* changed unless you subsequently use the Save button to over-write them.) You can also edit the name, to create a new data item in memory (or over-write one already existing with the new name) as a copy or edited copy of the currently selected item. The OK and Cancel buttons also become active—see below.

- New. The New button clears the central display region, and then acts like the Edit button, allowing you to create a new item in memory from scratch.

- OK. The OK button is active and set to the default button (and so may be pressed simply by pressing the Enter key) when the central region is active (and if a valid Name is showing). Pressing it sets the data item from the display into memory, over-writing any existing item of that name. The central region, and the OK and Cancel buttons, then become inactive again.

- Cancel. The Cancel button is active when the central display region is active (and may be pressed also by pressing the Esc key). Pressing it aborts the editing operation. The central region is reset to show the data item currently selected, and becomes inactive (together with the OK and Cancel buttons).

- Quit. The Quit button closes the Data window.

5.4 Effect of changes

If data in memory, that are currently being used by an analysis window, are edited or deleted, then appropriate changes are made automatically to that window when it is next selected.

5.5 APL data entry

Normally, when you ask for edited data to be stored in memory, by pressing the OK button, First Bayes checks that the Data field contains only a sequence of numbers separated by spaces. Any special characters will cause an error message and you will be expected to correct the Data field. If, however, "APL data entry" is turned on from the "Options" menu, this checking is not done. The contents of the Data field are treated as an expression in the native APL language. This allows two special facilities. First, it allows data to be generated randomly using one of First Bayes' built in random number generators. Second, it allows someone familiar with the APL language to perform complex operations on data. These facilities are described in the next two sections. Notice, however, that an important consequence of turning off the data checking is that if the Data field does not contain a valid APL expression when the OK button is pressed, then First Bayes will crash and exit immediately, losing all your calculations in progress.

5.6 Random numbers

When "APL data entry" is on, random numbers can be created as data by typing one of the following expressions (and nothing else) in the Data field. The words must be in capital letters.

- <p> BINGEN <n>, where <p> is a probability and <n> is a positive integer, both separated from the word BINGEN by one or more spaces. This generates <n> binary numbers, i.e. zero or one, with each having probability <p> of being one or probability 1 - <p> of being zero.

- <m> POISGEN <n>, where <m> is a positive number and <n> is a positive integer, both separated from the word POISGEN by one or more spaces. This generates <n> random numbers from the Poisson distribution with mean <m>.

- <m> <v> NORMGEN <n>, where <m> is any number, <v> is a positive number and <n> is a positive integer, all separated from each other and from the word NORMGEN by one or more spaces. If <m> is to include a minus sign it should be enclosed in parentheses. This generates <n> random numbers from the normal distribution with mean <m> and variance <n>.

- CURGEN <n>, where <n> is a positive integer separated from the word CURGEN by one or more spaces. This generates <n> random numbers from the *current* distribution, which is whatever distribution was last used by First Bayes. To ensure that the current distribution is the one you intend to draw random numbers from, you should set that distribution up in a Distribution window (opened from the "Distributions" menu on the menubar) immediately before using the Data window. (Analysis windows use distributions in such a way that the effect on the current distribution can be difficult to predict. So avoid interacting with Analysis windows before using the CURGEN expression.)

When you press OK to define these data, you will not see the data immediately. To view the random data, select another data item in either of the list boxes, and then select again the newly created or edited data item.

5.7 Other APL expressions

When "APL data entry" is on, APL expressions can be used to create data in complex ways. In general you need to understand the APL language, and preferably be familiar with the APL keyboard mapping in Dyalog APL/W, to make the most of this feature. (Remember, any error will cause First Bayes to crash.) However, a number of useful cases are given in the next Section for the general user.

First some important general points. Many of the APL operations require special characters not normally available from the keyboard. The "+" key to the right of the numeric keypad toggles the keyboard between its normal state and a special APL keyboard. If you use this to obtain special characters *do not forget* to toggle back to the normal keyboard, or you are likely to get strange characters appearing where you do not want them!

Negative numbers are denoted in APL by the special "high minus" sign, which can be obtained as Shift-2 in the APL keyboard. An alternative is to enclose every negative number in parentheses, when you can use the ordinary minus sign. An ordinary minus sign not isolated in this way can have unpredictable effects if you do not understand APL.

A data item in memory can be referred to by using its name preceded by a lower case d (with no space between).

The operations described in the next section (and many more) can be combined into complex APL expressions if you understand the language, but otherwise you are strongly advised to use them just one at a time. Save the result each time as a new named item, press the New button to create a new item and apply the next operation to the saved item by name. In this way you can build up complex operations step by step.

5.8 Simple and useful expressions

In the following, <data>, <data1> or <data2> means an item of data appearing either as a list of numbers separated by spaces or as a named item in memory (see above).

- <log symbol> <data> takes the natural logarithm of each number in <data>. <log symbol> is obtained as Ctrl-Shift-8 from the APL keyboard, and appears as a star in a circle.

- <base> <log symbol> <data> takes the logarithm to base <base> of each number in <data>. <base> must be a single positive number.

- <exp symbol> <data> takes the exponential of each number in <data>. <exp symbol> is obtained as Shift-P from the APL keyboard, and appears as a star.

- <num> <exp symbol> <data> takes <num> raised to the power of each number in <data>. <num> must be a single number.

- <data> <exp symbol> <power> raises each number in <data> to the power <power>, which must be a single number.

- <num> <take symbol> <data> takes a starting or ending portion of <data>. <take symbol> is obtained as Shift-Y from the APL keyboard and appears as an up-pointing arrow. <num> must be a single integer. If positive, <num> numbers are taken from the start of <data>, and if negative then -<num> numbers are taken from the end of <data>. If you ask for more numbers than there are in <data>, the extra will be made up of zeros.

- <num> <drop symbol> <data> achieves the same effect as the lase case but by specifying not how many numbers to take from <data> but how many not to take. <take symbol> is obtained as Shift-U from the APL keyboard, and appears as a down-pointing arrow. If <num> is positive, <num> numbers are dropped from the start of <data>, and if negative then -<num> numbers are dropped from the end.

- <num> <repeat symbol> <data> takes <num> numbers from <data>, repeating them if necessary. <repeat symbol> is obtained as Shift-R from the APL keyboard, and appears as a Greek letter rho. <num> must be a positive integer.. <num> numbers are taken from the start of <data>, but if there are not enough numbers in <data> more will be taken by going back to the start of <data> and reusing them as often as necessary. For example, 11 <repeat symbol> 0 1 will create the data 0 1 0 1 0 1 0 1 0 1 0.

- <data1>/<data2> selects numbers from <data2> as specified by <data1>. <data1> and <data2> *must* be of the same length, and the numbers in <data1> *must* be non-negative integers. The result is obtained by looking at the first number in each, then the second number in each, and so on. The result begins by being empty, then at each step if <n1> and <n2> are the corresponding numbers in <data1> and data2> then <n1> copies of <n2> are added on the end of the result so far. For example, 0 1 0 1 0/1 2 3 4 5 produces the result 2 4. And 0 1 2 1 0/1 2 3 4 5 produces 2 3 3 4. This operation is most commonly used with <data1> a list of zeros and ones created by an operation of the type described next. Another use is to generate binary data as needed for the Binomial sample analysis (see Chapter Seven), by for example 27 8 / 0 1.

- <num> <test symbol> <data> carries out a test on each number in data by comparing it with <num>. <test symbol> is *one* of the usual mathematical symbols for "less than", "less than or equal to", "equal", "greater than or equal to", "greater than" or "not equal to", obtained as Shift-3 to Shift-8 respectively in the APL keyboard. <num> must be a single number. If the relationship defined by the chosen <test symbol> is true for a number in <data> the result is a one, otherwise it is a zero. For instance, 1 < 0.2 0.8 1 2 with produce the result 0 0 0 1. Combining with the previous operation, data can be selected according to a test criterion. For instance, if NUMS is a data item and the result of 0 < dNUMS is saved as item NUMTEST, then the result of dNUMTEST/dNUMS will be to select only the positive numbers in NUMS.

- <data1>, <data2> joins the numbers of <data1> and <data2> to produce a single data item.

- <data1> + <data2> adds the numbers in <data1> to those in <data2>. <data1> and <data2> must be the same length. Alternatively if either <data1> or <data2> is a single number, it is added to each number in the other. Subtraction, multiplication and division are obtained using the minus, "times" and "divide" symbols. The "times" symbol is obtained by pressing the key with the equals sign from the APL keyboard, and the "divide" symbol by pressing that key shifted.

CHAPTER SIX.   DISTRIBUTIONS

6.1 Examining distributions

When you open a Distribution form by selecting any of the thirteen distribution types from the "Distributions" menu, it has two main regions. The right-hand region is for defining a particular distribution by setting its parameter values. The left-hand region is the larger of the two, and is for examining various properties of the distribution specified on the right. We begin by describing the parts of the left-hand region, because it is an important component of the First Bayes approach. Identical regions appear on all the types of Analysis forms, for examining prior posterior, marginal or predictive distributions as the case may be.

The top part always shows the name of the specified distribution, including its parameter values, and some key summaries of it—its mean, median, mode(s), standard deviation, variance and quartiles. If the name is too long to appear in the box, left and right scroll buttons will appear beside it. Click on a button to scroll by one character, or double-click to scroll by ten characters at a time.

The lower part allows you to select various other summaries to examine.

- Probabilities. If you enter two numbers, say *a* and *b*, in the Probability edit field, the probability contained in the interval [*a*, *b*] will be shown alongside.

- Highest density intervals. If you enter a number, say *p*, in the HDI edit field, the *p*% highest density interval will be computed and shown alongside. For a unimodal distribution this will always be two numbers giving the ends of the interval. For a multimodal distribution, there might be two or more separate intervals comprising the "highest density interval", whose ends are shown as successive pairs of numbers.

- Quantiles. If you enter a number, say *q*, in the "% ile" field, the *q*-th percentile will be shown alongside. (The 25-th, 75-th and 50-th percentiles are of course already shown as the quartiles and median.)

- Calculate. The calculations above are not done, and the results displayed, until you indicate that you have finished editing. You do this either by leaving that edit field, for instance by pressing Tab to move to the next object in the form, or by pressing the Enter key. The Enter key actually activates the Calculate button when you are in this part of the form, and there is rarely a need for you to press the Calculate button explicitly by clicking on it.

- Plot. Pressing the plot button causes a window to be opened showing a plot of the probability mass function or density function.

Everything in this region will be updated if you make any changes to the distribution specified in the right-hand region. (The same is true of the identical regions which appear on Analysis forms.)

6.2 Mixtures

In the centre of the right-hand region are edit fields for you to enter values for the parameters of the distribution, and below these are various buttons. Above the parameters is an edit field called Parameter set, and two small buttons. The explanation of this, and of the Weight field lies in one of the key features of First Bayes.

First Bayes allows arbitrary mixtures of distributions from the same family to be defined and examined. For instance 0.4 *N* (0, 1) + 0.6 *N* (2, 2) is a mixture with weight 0.4 given to a standard normal distribution and weight 0.6 to the *N* (2, 2) distribution. A mixture distribution is therefore defined by specifying a number of sets of parameter values with a weight given to each set. For instance, the example has two parameter sets. Set 1 is (0, 1) with weight 0.4, and set 2 is (2, 2) with weight 0.6.

6.3 Defining a distribution

The form shows just one parameter set and associated weight at a time. To view and/or edit a different parameter set, type its number in to the Parameter set edit field, or click on the up or down buttons beside it to move to the next or previous set.

When you have defined parameter values for the first parameter set, a weight of 1 is automatically inserted. A valid distribution is now defined, and the left-hand region becomes active. If you wish to define a mixture distribution, you will now continue with the second parameter set. Note that the weights you define need not add up to one. First Bayes automatically scales the weights internally so that they sum to one (and this is shown in the name of the distribution as displayed on the left). So for instance to give equal weights to the parameter sets you can just give weight 1 to each.

Further details about defining parameter sets will become clear in describing the action of the four buttons below the Weight field.

- Reset. Pressing this button clears all the parameter sets and gives you a blank form to begin again with. (See also Section 7.2 for the action of this button when a Distribution form is opened to specify a prior distribution.)

- Delete. This button deletes the parameter set currently showing.

- Next. This is the default button when you are operating in the parameter and Weight fields, and so is effectively pressed by pressing the Enter key. Its action is to move to the next parameter weight field that is empty (or has an invalid value). Note that if there is any empty (or invalid) field the Parameter set field and its up and down buttons are inactive. You cannot move to another parameter set if the current one is not complete (including Weight). You must first either complete it or delete it with either of the preceding buttons. If the parameter set (and Weight) is complete, the Next button accepts the parameter set, updates the data in the left-hand region, activates the Parameter set field and buttons, and places the cursor on the up button (so that pressing Enter again will move you to the next parameter set).

- Quit. This button closes the window.

6.4 Particular distributions

The operation of all kinds of distribution is as described above. They differ only in the number of parameter fields and their names. None of the distributions should be unfamiliar to a Bayesian statistician. A classical statistician, on the other hand, might find the Beta, Beta-Binomial, Gamma and Inverse Chi-Square distributions new or unusual. Their definitions could be found in many standard textbooks.

Again, a classical statistician might be very familiar with the Chi-Square, *F* and *t* distributions, but might be surprised by the numbers of parameters given to them in First Bayes. In classical statistics these distributions are invariably met only in standardised forms. In Bayesian statistics it is convenient to give them all an extra scale parameter, and to add a location parameter (median, equal to the mean if the degrees of freedom exceed 1) to the *t* distribution. The role of these parameters will be obvious, or should become obvious if you experiment with them.

CHAPTER SEVEN. ONE PARAMETER ANALYSES

7.1 The three one-parameter analyses

Selecting any of "Binomial sample", "Poisson sample" or "Normal sample, known variance", from the "Analysis" menu on the menubar, opens an Analysis window that is basically the same for each of these analyses. Each allows a prior distribution from the conjugate family (including mixtures) to be combined with data of the appropriate type.

- Binomial sample. The data in this case are independent trials with probability θ of success in each. They are therefore a series of ones (denoting success) and zeros (denoting failure). The distribution of the number of successes is binomial. The conjugate prior family for θ is the Beta family.

- Poisson sample. The data in this case are iid observations from a Poisson distribution with mean θ. They are therefore a series of non-negative integers. The conjugate prior family for θ is the Gamma family.

- Normal sample, known variance. The data in this case are iid observations from a Normal distribution with mean θ and known variance. The conjugate prior family for θ is the Normal family.

7.2 Defining the prior

When you open an analysis window for one of these analyses, another window is automatically opened for you to define the prior distribution. This is a distribution window (see Chapter Six) of the conjugate type.

The behaviour of First Bayes when this prior distribution window is open is different from the usual. You can *only* work with the prior distribution window, and all the others are inactive. Even if you open a graphics window from the prior distribution window (to plot the prior density) it is inactive, and you cannot close it, or bring it to the front when it is obscured. You must complete the definition of a prior distribution before you can continue with any other operations. When you close the prior distribution window, the distribution that was defined there becomes the prior distribution for the Analysis window, and its name is displayed there.

Pressing the Edit button reopens the prior distribution window for you to edit the prior. The currently defined prior distribution is already set up in the Distribution window for you to edit (and the Reset button will now restore that setting, rather than clearing all parameter sets).

If you close the prior distribution window without a proper distribution being defined, an error message box will appear. Pressing the OK button will reopen the prior distribution window for you to try again. Pressing the cancel button aborts the whole process and *closes* the Analysis window.

7.3 The variance field

In the case of a "Normal sample, known variance" only, an edit field appears for you to enter the known value of the population variance (initially set to 1).

7.4 Defining the data

The data are defined from a drop-down menu labelled Data. Initially, the data are set to Nul, i.e. no data at all. Clicking on this field drops a menu of all data items currently available in memory (see Chapter Five), from which the desired item can be selected.

7.5 Examining the posterior distribution

The left-hand region of the Analysis form is a display of distribution details, just as in a Distribution window (see Section 6.1). Initially, it shows the posterior distribution of θ, but may be switched to show a predictive distribution—see Section 7.6 below. Once a prior distribution has been defined, this region becomes active. The data will then be Nul, and so the posterior distribution will just be the prior distribution. So at this stage the display is effectively repeating information that was available in the prior distribution window. Once some actual data have been defined, however, it will display the new posterior distribution. It may be used in the same way as described in Section 6.1 to obtain appropriate posterior inferences.

7.6 Predictive distributions

A small region on the right of the window deals with predictive analysis. An edit field allows you to specify the number of observations whose sum or mean (depending on the analysis) is of interest. Pressing the "Show predictive" button then switches the left-hand region to display the predictive distribution of that sum or mean. This button actually toggles the left-hand display, and will therefore switch the display back to showing the posterior distribution of θ.

The precise meaning of the predictive distribution for each analysis is a follows.

- In the Binomial sample analysis, the prediction is for a number of future independent trials, also with probability θ of success. These will be zero or one, as with the data for this analysis, so their sum is the number of successes in those future trials. The predictive distribution will be Beta-Binomial.

- In the Poisson sample analysis, the prediction is for the sum of a number of future iid Poisson observations with mean θ. The predictive distribution will be Negative Binomial.

- In the Normal sample, known variance analysis, the prediction is for the mean of a number of future iid Normal observations with mean θ and the same known variance. The predictive distribution will be Normal.

7.7 Other controls

There are three buttons at the bottom of the right-hand region of the window.

- The Triplot button opens a graphics window showing the prior density, likelihood (normalised to integrate to 1) and posterior density on a single plot. This is extremely useful for gaining understanding of how Bayes' Theorem works. It is of course updated automatically if you change either the prior distribution or the data.

- The Update button resets the prior distribution to the current posterior distribution, and the data to Nul. The analysis is then ready to receive more data, to compute a new posterior distribution. The original prior distribution and data can no longer be edited, but this new prior distribution (the old posterior) can be edited if desired.

- The Quit button closes the Analysis window.

CHAPTER EIGHT. LINEAR MODELS

8.1 Two linear models

Through the two items "Normal sample(s), unknown variance" and "Regression Data" in the "Analysis" menu, First Bayes offers analysis of the two simplest kinds of linear model. They provide opportunities to learn about the Bayesian approach in multiparameter problems.

- In the Normal sample(s), unknown variance, analysis the data are one or more samples, each consisting of iid Normal observations with unknown mean and variance. Each sample is a separate data item in First Bayes. The variances of the various samples are assumed equal, so the unknown parameters comprise an unknown mean for each sample plus a single unknown variance. In the case of a single sample, the mean is called "Mu" by First Bayes, and for more than one sample the means are "Mu1", "Mu2",... . The variance is called "Sig".

- In the Regression data analysis, two data items of the same length are required. One is the *X*-data, the values of the "independent" or "regressor" variable. The other is the *Y*-data, the corresponding observed values of the "dependent" or "response variable. The model is the usual simple regression model

$$y_i = \alpha + \beta\, x_i + e_i,$$

  where the e_is are Normally distributed with zero mean and unknown variance. First Bayes calls α and β "Alpha" and "Beta", and the variance "Sig".

The Analysis window has basically the same features in both cases, although there are a number of differences. The left-hand region of the window is, as usual, used to display data about a distribution, and may be used as described in Section 6.1. In these analyses, the distribution being displayed can be changed in a variety of ways, see Sections 8.4 and 8.5.

8.2 The prior distribution

In these analyses you cannot specify genuine prior informtion. The initial assumption is shown as "Weak" prior information, corresponding to a conventional non-informative prior. As with a one-parameter analysis, the Update button (see Section 8.7) makes the current posterior distribution into the prior distribution. It is therefore technically possible to create a proper conjugate prior distribution by first inputting some artificial data and then using Update. A later version of First Bayes may allow the creation and editing of proper prior information, but this will inevitably be a complex task.

8.3 Data

The two analyses require different kinds of data.

The Normal sample(s), unknown variance, analysis first has an edit field at the top of the right-hand region, to enter the number of samples. [There is a small bug in the current version - you should define the first data set as described in the next paragraph *before* changing the number of samples.] The Data area then shows just one sample at a time. The number of the sample being currently defined may be set by typing in the Sample number edit field or by clicking on the up and down buttons beside it. Then the data item providing the observations for that sample is set using the drop-down menu as in Section 7.4. The data are initially set to Nul for all samples. Once genuine data have been entered for all samples, the posterior distribution is defined and the rest of the window becomes active.

The Data area for the Regression data analysis just has two drop-down menus, which are used as in Section 7.4 to define the *X* and *Y* data items. Both are initially set to Nul. Once genuine data have been entered for both, the posterior distribution is defined and the rest of the window becomes active.

8.4 Marginal distributions

The left-hand region is initially set to display facts about a posterior marginal distribution. It can be set to display the posterior marginal distribution of any individual parameter. It may also display the posterior marginal distribution of any linear combination of the location parameters (i.e. all parameters except "Sig"). Which marginal distribution is displayed is set and shown in the area in the middle of the right-hand region labelled "Margin".

The Margin area contains a field with the name of one of the parameters in it, with an up button beside it, and also an edit field labelled Weight, with another up button beside it. It is important to recognise the different effects of using the two buttons. They are superficially similar, because pressing either will cause the parameter name showing to change to the next parameter. But they affect the Weights differently.

Using the left-hand up button causes the marginal posterior distribution of the parameter now showing to be displayed in the left-hand region. This is then the distribution of that parameter alone, not a linear combination. Or more precisely, it is the distribution of the linear combination which gives weight 1 to the parameter whose name is showing, and 0 to all those not showing. So this button not only changes the name showing but also changes the weights so that this parameter gets weight 1 and all the others 0.

The right-hand up button changes the parameter whose name is showing, and displays the weight currently given to that parameter, but does not change any weights. Consequently, it does not change the distribution being displayed in the left-hand region. It is this button that must be used if you want to examine the posterior distribution of a linear combination. Use it to cycle through the parameters, and type the weights you require in the Weight field for each one.

Notice that if you give a non-zero weight to "Sig", then weights for all the others are automatically set to zero. Likewise, giving a non-zero weight to any other parameter causes the weight for "Sig" to be set to zero. You cannot define a combination that includes both types of parameter.

In the case of the Normal sample(s), unknown variance, you can use the linear combinations feature to find posterior distributions of arbitrary contrasts between the sample means. In this case, there is another way to change quickly the marginal distribution being displayed. If you change the Sample number in the Data area, the posterior marginal distribution of that sample's mean is shown. It has the effect of using the left-hand up button in the Margin area to cycle through the parameters. (So it erases any linear combination you may have set up.)

The main use for linear combinations in a Regression data analysis is to examine the fitted regression line at a particular value of the X variable. The linear combination with weight 1 for "Alpha" and x for "Beta" is $\alpha + \beta x$, the expectation of Y when $X = x$.

8.5 Predictive distributions

The area in the window for predictive analysis is similar to that in the one-parameter analysis, Section 7.6, and is used in the same way. However, its meaning is linked to the current marginal distribution as defined in the Margin area. The predictive distribution calculated is for the mean of a specified number of iid observations with mean equal to the parameter or linear combination of parameters currently defined in the Margin area, and with variance "Sig". (The predictive analysis is inactive if the current marginal distribution is that of "Sig".)

In a Normal sample(s), unknown variance, analysis, if the current marginal distribution is for the population mean parameter of sample k, then the predictive distribution is for the sample mean of a future sample from the *k*-th population.

In a Regression data analysis, if the current marginal distribution is that of $\alpha + \beta x$, then the predictive distribution is that for the mean of a future sample of iid observations, all with $X = x$.

As in the one-parameter analysis, Section 7.6, the button "Show predictive" toggles the left-hand display between showing the posterior marginal distribution and the corresponding predictive distribution (and the text on the button changes accordingly).

8.6 Bayesian F tests

Examining the appropriate marginal posterior distributions allows a wide range of posterior inferences to be made in a natural and direct way. The area of the window labelled "Bayesian F test" supplements these with further inferences that can be useful.

Three alternatives are offered, indicated by the "radio buttons" marked "All 0", "All =" and "Marg 0". Clicking on any button toggles it on (spot in the middle) or off (empty circle). Switching one on switches all the others off. If one button is on, the Test probability field shows the result of calculating a "Bayesian F test".

The explanation of what this test is, is easiest in the case of "Marg 0", which tests whether the parameter (or linear combination of parameters) defined as the current marginal distribution could be zero. The probability reported is the posterior probability that this parameter is outside the smallest HDI (highest density interval) containing the point zero. So this is the posterior probability that the parameter in question should be as far or further from its mean as the value zero. If this is small, the parameter is unlikely to be so far from its posterior mean as zero, and so it is implausible in some sense that it could really be zero. The "Marg 0" test could be used to test whether a particular contrast in the group means in a Normal sample(s), unknown variance, analysis would be zero.

The "Alt 0" test is similar, but tests all the parameters (except "Sig") simultaneously. The reported probability is the probability that these parameters could jointly lie as far or further from their joint posterior mean as the origin.

The "Alt =" test in a similar way tests whether all the parameters (except "Sig") could be equal, effectively by looking at whether a set of contasts might be zero. In a Normal sample(s), unknown variance, analysis it is equivalent to the classical one-way analysis of variance test.

In general, with weak prior information, these tests are equivalent to classical hypothesis tests. The probability being calculated is a tail area probability of an F distribution. (Note the remark on accuracy in Section 1.4—when the data quite strongly suggest that the parameters are not zero or equal, the reported probability is likely to be rounded to zero, meaning only that it is below the level of accuracy of the calculations.)

The tests are all inactive if the current marginal distribution is that of "Sig".

8.7 The remaining Buttons

At the bottom of the right-hand region of the window are two or four buttons.

- The Update button acts as in one-parameter analyses (see Section 7.7). It turns the current posterior distribution to the prior, and sets all data items to Nul. The analysis is then ready to receive further data.

- In a Regression data analysis only, the Resid button creates a new window showing a plot of residuals against the values of the *X* variable.

- In a Regression data analysis only, the Scatter button creates a new window with a scatter plot. The fitted line is shown together with two standard deviation bounds on either side (a) for the line itself (derived from the posterior distribution of $\alpha + \beta x$) and (b) for a future observation (derived from the predictive distribution of a single observation with mean $\alpha + \beta x$).

- The Quit button closes the Analysis window.

CHAPTER NINE. EPILOGUE

9.1 The Bayesian computer package debate

There has been considerable debate amongst Bayesian statisticians about the wisdom of creating Bayesian computer packages. One view is that the Bayesian philosophy emphasises the need to think—to think hard about the model, to think hard about the prior information, and to think about what computations on the posterior distribution will yield relevent inferences. A Bayesian computer package should therefore be an immensely complex thing, allowing the user to experiment and build esentially arbitrary prior joint distributions for all the parameters, to combine this with the likelihood to yield the posterior distribution and then perform essentially arbitrary computations on it to yield any inferences or decisions of interest, and also to experiment with varying the prior, likelihood or decision problem in essentially arbitrary ways in order to study the robustness of the conclusions.

The other view is that such a package is an impossible ideal. Not only can we not wait for it to be realised, we cannot wait for even a fraction of that functionality to exist. There are masses of other computer packages already in existence, all applying classical ideas, all tempting the user into believing that thinking is not really necessary. Bayesian statistics stands no chance if it cannot compete in that market *now*.

In offering First Bayes, I am very conscious of this dilemma. I have a great deal of sympathy with the first view, but also see the force of the second. Perhaps those demanding packages now might be mollified by the progress that the Bayesian approach has made in recent years. That has been built on the ability to solve problems that classical statistics cannot handle at all. Highly sophisticated, non-user-friendly software is used. Yes, we are losing the battle at the lower end. The great mass of statistics being done in the world is being done with classical computer packages, by people who do not understand statistics properly and cannot properly interpret the answers (especially because the answers they are getting are of the convoluted classical kind). But the future lies in training the new generations to do and to understand Bayesian statistics, and that will happen if we first win the people who do the training. Bayesian statistics is starting at the top, by winning the people who train (and particularly those who train the trainers) by showing them its power in difficult problems. In time, this must percolate down.

Where does First Bayes stand in this debate? Well, it does not claim or attempt to be for actually doing Bayesian analysis. It is for teaching, and as such is free to pick and choose what it does. The choice is made to give the student some appreciation of how the Bayesian approach works and what it can do. I believe that the one-parameter analysis is a nice example, because it really does allow essentially arbitrary priors. On the other hand, I have abdicated from that in the case of linear models! I would have liked to treat prior specification better for those cases, and perhaps I will in a later version. I did think it important to include some multi-parameter analysis.

I think that the availability of predictive distributions throughout is another good feature, as is the general facility to study distributions and derive appropriate summaries. Of course this is a long way from offering the computation of a complete range of inferences, and I would like to add more, but it is a start.

I hope you feel, like me, that it was worth doing. The trainers will need tools, and we should make a start on developing them.

9.2 Future development

First Bayes has been something of a labour of love. Please give me your suggestions for improvement (or tell me to give up now). I will not promise to put any of them into effect. I will not even promise that any further development will take place—I have already spent more time on it than I can really spare!

Among the obvious ways to improve it are the following.

- Better analysis of linear models, including prior specification.
- More hard copy facilities.
- Ability to read data from text files created outside First Bayes.
- More posterior summary and inference options, like posterior distributions of general functions of parameters, or bivariate contour plots.
- Better or easier data manipulation facilities.
- More accurate probability calculations.
- Support for other graphics resolutions, etc.
- More families of distributions and more models.
- Online help.

Tony O'Hagan
Department of Mathematics
University of Nottingham
University Park
Nottingham   NG7 2RD
UK