Introduction to Computing in Statistics

Susan Hutchinson

Department of Statistics, University of Oxford

October 2017

Outline

Who we are

We are a team of four who manage IT within the Department.

Computing manager:	Stuart McRobert
Computing support specialist:	Susan Hutchinson
Computing officers:	Mark Feasey
	Simon Patchett

How to get help

- The best way is to email ithelp@stats.ox.ac.uk .
- No access to email? Phone 70869, 82852, 72588, or 72862.
- We are here between 9am and 5pm (at least).
- If there are no phones and email isn't working then there's probably a serious problem that we know about!
- If all else fails, come to our offices on the ground floor.

Contacting IT help

When you email ithelp@stats.ox.ac.uk please include the following information:

- Your username.
- Your computer's name.
- The operating system (Windows or Linux).
- Your office number.
- A brief description of the request or problem, including any error messages.

Specification

All new research students based full time in the Statistics Department have their own computer.

Hardware
Intel Core i7 3.40 GHz processor
32GB RAM
256GB Solid state disk
24" Dell flat screen monitor

Operating system
Windows 7 or Fedora 26

Software

For both Windows and Linux:

- R version 3.4.1.
- Matlab R2016a.
- LaTeX (MiKTeX 2.9 or TeXLive 2017).
- Office (Microsoft Office 2016 or Libre Office).

and lots more!

The application I want isn't there

Mail ithelp@stats.ox.ac.uk with:

- The name of the application you want, and a link to it if applicable.
- A brief description of the application.
- Windows or Linux.
- The name of your desktop computer.

Your department ID

You will be given a Statistics account - a username and password.

- It gives you access to Departmental services.
- It should be used by you only and never shared.
- It is separate from your University single sign-on account.
- Passwords cannot be changed except by coming to us.
- The first time you log on, it should be to your designated desktop computer.

Your files #1

Most of your files and folders are stored on a central file server.

- On a Windows computer this is your P: drive. Never store data on the 'Windows Desktop' or on the C: drive.
- On a Linux computer this is your home directory.
- PRS - quota of 10GB; CDT - 5GB. Reasonable requests for an increase will usually be met.
- Backups are taken every day.

On Linux computers, some configuration and cache files are stored on your local, faster, computer.

Your files #2

- Files stored on the central server should be for academic use only.
- Don't load music, films or other copyrighted material.
- Traffic is monitored both by us and by the central computing services.
- Illegal downloads are noticed and may result in at least the suspension of your account.

The central file server space should be used with care. Applications such as Dropbox should be configured to save files on the fast local desktop ***not*** on the central file server.

Large datasets

Store large work-related datasets on your fast local computer's disk.

Linux
There should be a directory /data/*host*/*user* which should be used – for example /data/parrot/flint.

Windows
Data should be stored in D:*user*.

- Data stored on the local disk is backed up weekly.
- If the directory or folder where you should store your data doesn't yet exist locally, email ithelp@stats.ox.ac.uk .

Where to run jobs #1

Your desktop computer is not meant for running compute-intensive jobs. Use it to:

- Develop and test software. If you are testing jobs please store data and output on your local fast directory ***not*** on the central server.
- Read email.
- Browse the web.
- Do word-processing (LaTeX / Microsoft Word / LibreOffice).

Where to run jobs #2

Use the Departmental compute servers for long and/or compute-intensive jobs.

- See the list of servers and guidelines for use here: http://tinyurl.com/OxStats-compute.
- Contact your supervisor — some groups have their own compute servers.
- All data and output you are using must be stored on the compute server.
- Run jobs at a low priority. See http://tinyurl.com/OxStats-priority for details.
- Use checkpointing software to allow jobs to restart. We recommend dmtcp.
- If you are not checkpointing then save the output and data regularly in case of power failure or emergency reboot.

Rules for running jobs

Departmental compute servers are a shared resource, so we ask users to keep to some simple rules to allow everyone to get time to run jobs.

- No more than one third of any available resources per server.
- Each job should last no more than than 24 hours. Mail ithelp@stats.ox.ac.uk if the job will be longer.
- Use the screen command to detach jobs from your desktop login.

Checkpointing jobs with dmtcp #1

We strongly encourage users to checkpoint their jobs so that they can be restarted in the event of a reboot due to:

- Power failure
- Software failure
- Monthly system security update

The first two are very rare, the third may be necessary on the Wednesday morning after the second Tuesday of each month.

We recommend the dmtcp package for checkpointing jobs.

Checkpointing jobs with dmtcp #2

dmtcp will checkpoint most binary programs including

- R
- Java
- Python
- Perl
- Matlab
- ... and many more

No additional compilation is needed.

There is a FAQ at http://dmtcp.sourceforge.net/FAQ.html which will get you started.

Rules and regulations

The University expects all users to abide by the rules governing computers used on the network.

- Please read the information at https://www.it.ox.ac.uk/welcometoit.
- Some restrictions are placed on 'Peer-to-Peer' (P2P) applications to avoid excessive consumption of limited shared resources such as bandwidth.

Finally, remember that sharing copyrighted material is illegal.

Central IT services

In addition to those provided by the Statistics IT team, some services are provided by central IT services (ITS).

- At your college address you should have received a separate username in the form ***abcd1234*** and an activation code.
- This is your Single Sign-ON account (SSO) which is used to access several centrally provided services, **especially email**.
- Please activate this account as soon as possible.

Nexus Email

- Your Nexus email account is the only University email mailbox available to you.
- All email in Statistics is forwarded automatically to your Nexus email address.
- For more about Nexus email see http://www.it.ox.ac.uk/nexus.
- I will be contacting everyone on their Nexus email address before the end of the week. Please reply to confirm that you got my message.

Configuring Nexus email

Once you have activated your Single Sign-On account, use the Nexus account setting page https://register.it.ox.ac.uk/self/nexus to:

- Set spam filter level — choose between Off, Low, Medium or High and make sure you activate it. See http://help.it.ox.ac.uk/email/filter.
- Set a preferred email address if you have more than one Oxford address.
- Set an email forwarding address if you want all email delivered to a different address, eg gmail or hotmail.
- Show email usage and quota to see the size of each folder.

A final email configuration

As well as the standard set-up, we strongly advise you to make one final change. This only applies if you use Outlook 2010 or Outlook 2007 to read email.

- If you choose to configure Outlook to read your mail, please make sure you *never* use the Auto-Archive function.
- This function moves older messages to the local hard drive of the machine you are using and so won't be available if you move computer and will probably not be backed up.

Reading email

The quickest way to read email is via a browser. The address is http://nexus.ox.ac.uk.

Email readers such as:

- Outlook 2016 (Windows)
- Outlook 2015 (Mac)
- Thunderbird (Linux, Mac and Windows)

can make make managing large amounts of mail easier.

There are instructions on how to set up these clients to read Nexus email here: http://help.it.ox.ac.uk/nexus/email/.

Remote access account

What about WiFi?

Once you have set up your SSO account, you should generate a remote access account which will allow you to connect to the University's WiFi network: Eduroam.

There are more details here: http://www.it.ox.ac.uk/welcome/connect.

Preparing for the worst

Bad things happen! How would you manage if your personal laptop or desktop was stolen or stopped working?

Backups

- IT services run a free backup service which allows you to take regular backups of the data on your computer. See http://www.it.ox.ac.uk/hfs.
- Register for the service using your SSO. You can then download and set up the software. This will take only a few minutes. We strongly recommend that you do this.

Hardware repairs
Central IT services also run an Apple hardware repair service http://www.it.ox.ac.uk/breakdown/) which is available for all.
They also recommend a third-party repair company for all computers: Equinox Maintenance Ltd. The details are on the same page.

Software and courses

Software
The online shop provides software for personal use at a reduced or no cost. Applications include Endnote and Corel products. See http://www.it.ox.ac.uk/shop/ for details.

It is also possible to buy Microsoft Office products via Dreamspark. See https://www.it.ox.ac.uk/sls/fulllist/ for more details. You will need your Single Sign-On account.

Courses
Central IT services run a programme of IT courses which cover a range of software and skills, usually at little or no cost. See https://www.it.ox.ac.uk/itlp/ for more information.

OxFile: Large file exchange service

Sometimes you want to transfer a file to someone but it is too large to attach to an email or to put into your personal web space.

- OxFile is the University's "DropBox" service
- Visit it on https://oxfile.ox.ac.uk/
- Limits: (file) 25Gb, (folder) 40Gb
- Access using Oxford username and password
- Can be used for receiving files as well as for passing them to others
- Data can be received from/passed to people outside the University

Using your laptop

You may have a laptop that you will be using within the Department or University.

Wireless connection Central IT provide a wifi service

EDUROAM available in many academic institutions world-wide. You will need your remote access SSO account to use this service.

See http://help.it.ox.ac.uk/network/wireless/services/eduroam/index for details.

Wired connection Please mail ithelp@stats.ox.ac.uk with the MAC (or hardware) address of your computer. There are instructions here: http://tinyurl.com/statsconnect describing how to find this.

Statistics VPN

Or Virtual Private Network.

- Both Central IT and Statistics run a VPN service that you can use to connect to local services from your own computer.
- We recommend that you use the Statistics VPN as this gives access to additional services:
  - Print to departmental printers. See http://tinyurl.com/statsprint.
  - Access data on your P: drive (or home directory). See http://tinyurl.com/statspdrive.

Looking after your laptop

Here are some simple steps you can take to make sure that your laptop stays safe.

Software updates
Make sure your operating system software is up-to-date. This applies equally to Linux, Mac and Windows.

Anti-virus software
Make sure you have installed anti-virus software and that it is being updated very regularly. As a member of the University you can download a copy of Sophos Anti-Virus for free by logging on with your SSO to https://www.it.ox.ac.uk/services/security/antivirus-software.

Mobile computing #1

Smart phones and tablets are for many of us now an essential part of how we work and live our lives.

They do however create extra risks:

- Easy to lose or have stolen.
- Open wireless networks and public WiFi do not protect data as it travels over the network, so it can be read by anyone.

Mobile computing #2

What can I do to minimise the risks?

- Do not store confidential or sensitive data on your device.
- Where possible, use VPN to access data.
- Use a secure network when sending or receiving private information.

These two sites provide simple and authoritative advice on protecting yourself.

- https://www.infosec.ox.ac.uk/
- https://www.getsafeonline.org/smartphones-tablets/

Mobile computing #2

What can I do to minimise the risks?

- Do not store confidential or sensitive data on your device.
- Where possible, use VPN to access data.
- Use a secure network when sending or receiving private information.

These two sites provide simple and authoritative advice on protecting yourself.

- https://www.infosec.ox.ac.uk/
- https://www.getsafeonline.org/smartphones-tablets/

Finally, never carry sensitive or confidential data on a memory stick, unless it is encrypted.

Helping yourself

The following few slides suggest some software that you probably use frequently.

If you have your own computer you may want to consider installing the latest version yourself.

In all cases the software is free and open source and available for Linux, Mac and Windows.

Statistical software

R

- The R software environment is extremely useful for both teaching and research.
- The latest version is R 3.4.1.
- It is simple to install on your own computer. Go to http://cran.r-project.org/mirrors.html and find the nearest location where you can download the software.
- You may also want to look at Rstudio which provides an Integrated Graphical Environment (or IDE).

LaTeX

LaTeX is ideal for producing well-structured documents, particularly those containing scientific formulae, mathematical proofs or computer programs.

To use LaTeX you need

- A LaTeX engine such as MiKTeX for Windows or TeXLive for Linux and Mac.
- An integrated development environment (IDE) or editor. Some editors we recommend are:
  - Texmaker (Linux, Mac, Windows)
  - TeXworks (Linux, Mac, Windows)
  - Kile (Linux only)

All are free and open source, as is the LaTeX engine.

Giving presentations

You are very likely to need to give presentations during your time here.

- LaTeX and the Beamer package provide a very good way to create high quality technical presentations, particularly those containing formulae. For example:

$$h_i(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \frac{\Pr(t < T_i \leqslant t + \epsilon)}{\Pr(T_i > t)}.$$

- Microsoft Powerpoint and LibreOffice Impress are fine for non-technical presentations.

Managing your work with version control

Why use version control? And why use our **Subversion** service?

- Logs and saves changes.
- Supports collaboration.
- Web-based so data is available from wherever you work.
- Cross-platform, so will work on Linux, Mac or Windows.

See http://tinyurl.com/statssvn.

Saving the planet

There are various ways you can save energy during your time in the department.

- Apart from WEDNESDAY NIGHT, please shut down computers when they are not in use.
- To switch your computer on simply press the power button on the front of the computer.
- When you have finished log out and then shutdown the computer.

Our web pages

Finally always check our webpages for help, updates and news:

http://www.stats.ox.ac.uk/about_us/it_information