Statistics course and applied workshop for life sciences postgraduates

3-days basic statistics course for postgraduate students

28-30.1.12

Organisation team: Postgraduate Society of Biology & Biochemistry (Anne Gesell, Weihao Zhong, Martin Hughes, Alberto Lapedrizia, Dr. Matthew Wills)

Aim: provide an understanding of the main principles and rationales of common statistical analyses needed to plan experiments and interpret scientific data

Course outline:

Date: 28th – 30th January 2013

1. Day (28.1.13): 10:15 – 12:05 lecture theatre 6W 1.2
   - Descriptive statistics (measures of central tendency e.g. mean, median, etc.)
   - Testing normality and data transformation
   - Binomial and sign tests
   - Chi square tests

   14:00 – 17:00 R practical 1 in 3 South Computer teaching room
   - Introduction to R, R as a calculator and a graph maker
   - Self-run exercises of material covered in morning lectures

2. Day (29.1.13): 10:15 – 12:05 lecture theatre 6W 1.2
   - T-test (and tests for homogeneity of variance)
- Wilcoxon test (non-parametric paired t-test equivalent)
- Mann Whitney U test (non-parametric unpaired t-test)
- Kruskal Wallis test
- Friedman test (non-parametric two-way ANOVA)

14:00 – 17:00 R practical 2 in 3 South Computer teaching room
- Self-run exercises of material covered in morning lectures

3. Day (30.1.13): 10:15 – 12:05 lecture theatre 6W 1.2
- One-way ANOVA and post hoc tests
- Two-way ANOVA
- Correlation: parametric/non-parametric
- Simple linear regression and multiple regression

14:00 – 17:00 R practical 3 in 3 South Computer teaching room
- Self-run exercises of material covered in morning lectures
- Bring your own data R practice (try using knowledge gained during the course on your own data!)

Background:

The ability to make predictions, analyse data and draw conclusions is central to many disciplines in science, engineering and medicine. This course was organised to provide practical training in data analysis and statistical methods for postgraduates, Master students and staff within the Department of Biology/Biochemistry and beyond. The content was drafted using the comments and suggestions received from postgraduate students and academic staff, who clearly identified a need for a statistics course.

We divided this course into morning lectures and afternoon workshops. The morning lectures covered the statistics most commonly used in experimental life sciences including examples involving real data collected by previous students. The lectures were located in 6W2.1, which can facilitate 60-80 students (64 students attended).

The afternoon workshops provided an introduction to the statistical programming environment “R” and how to use it for applying the theory taught in morning sessions. This workshop was held within the department of Biology/Biochemistry computer laboratory. In general, R is one of the most powerful and flexible statistical packages available today, rivalling other popular statistics software in functionality and power. The afternoon sessions took place in the 3 South Biology/Biochemistry Computer teaching rooms and were
attended by 29 students (maximum capacity 30). This part of the course was based on course material for teaching R provided by Dr Matthew Wills and was taught by two very experienced and enthusiastic postgraduates (Weihao and Martin).

To improve the attendance level we sent out e-mails within our department, advertised it on the University PGSkills training homepage and placed flyers within different departments. Attendees came from a wide variety of departments, including Biology & Biochemistry, Health, Pharmacy & Pharmacology, Chemistry and Engineering.

Project outcomes and benefits:

- The participants will be more confident and capable with the statistical analysis of their data. This is a highly valuable generic skill in biological sciences.
- The participants will have obtained an excellent foundation in the use of the powerful and versatile program R. This will encourage and facilitate them to apply more advanced analysis techniques that R has to offer in the future.
- The participants will have the opportunity to network with other researchers and to share different analytical approaches in a wide variety of research problems.

The course organisers will gain valuable experience in managing a focussed event to a specific audience, which can easily be transferred outside academic research setting.

Feedback from the attendees included positive as well as negative outcomes. The table below shows the percentage in each category.

**Feedback for morning lecture:**

<table>
<thead>
<tr>
<th></th>
<th>Strongly agree</th>
<th>Agree</th>
<th>disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfied with this course</td>
<td>51%</td>
<td>42%</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>Would recommend this course</td>
<td>51%</td>
<td>44%</td>
<td>5%</td>
<td>0%</td>
</tr>
<tr>
<td>Was relevant to my needs</td>
<td>42%</td>
<td>47%</td>
<td>12%</td>
<td>0%</td>
</tr>
<tr>
<td>Content was appropriate</td>
<td>42%</td>
<td>49%</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>Methods used were effective</td>
<td>49%</td>
<td>44%</td>
<td>7%</td>
<td>0%</td>
</tr>
<tr>
<td>Facilitator had appropriate skills/knowledge</td>
<td>72%</td>
<td>26%</td>
<td>2%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Much of the positive feedback indicated that the lecture material was easy to follow and included many worked examples. Some participants stated that the course was a good refresher for their statistic knowledge. Among the negative comments, some participants reported that the lecturer didn’t follow the hand-outs exactly and that is was too basic.

**Feedback for afternoon R workshop:**

<table>
<thead>
<tr>
<th></th>
<th>Strongly agree</th>
<th>Agree</th>
<th>disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfied with this course</td>
<td>50%</td>
<td>50%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Would recommend this course</td>
<td>55%</td>
<td>45%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Was relevant to my needs</td>
<td>40%</td>
<td>60%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Content was appropriate</td>
<td>60%</td>
<td>35%</td>
<td>5%</td>
<td>0%</td>
</tr>
<tr>
<td>Methods used were effective</td>
<td>35%</td>
<td>65%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Facilitator had appropriate skills/knowledge</td>
<td>55%</td>
<td>45%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

The majority of attendees enjoyed attending this part of the course and could apply tests in their own research. Twelve participants found the R course to be very useful and expressed interest in using the program to analyse their own data. This is a good outcome and demonstrates that we need to provide basic statistic courses for postgraduate researchers and staff.

**Future Reference:**

All material is available electronically on Moodle under "Basic Statistics Course for Life Sciences" (free access for everyone with an active University of Bath username).
Photos of the day are included below:

Morning lecture:

Afternoon practical: