THE ROYAL STATISTICAL SOCIETY

Incorporated by Royal Charter: 1887

Professor A T A Wood
School of Mathematical Sciences
University of Nottingham
University Park
Nottingham NG7 2RD
United Kingdom
*Tel*: +44 (0)115 9514983 *Fax*: +44 (0)115 9514951
*E-mail*: andy.wood@nottingham.ac.uk

from the Joint Editor
*Journal of the Royal Statistical Society*
Series B, *Statistical Methodology*

17 April 2007

Professor Robert Stine
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6340
USA

Dear Professor Stine

**JRSSB submission B6215R: "Alpha-investing: a new multiple hypothesis testing procedure that controls mFDR" (with Foster)**

Thank you for sending the revision of your paper to JRSSB. I am sorry for the length of time it has taken to review B6215R. Unfortunately, Referee 2 has not provided a report and I have decided not to wait any longer. Reports from the Associate Editor and Referee 1 are enclosed; I have also provided a brief report.

As you can see, both reviewers believe the paper has improved substantially. However, there are still a number of points that require attention. Referee 1 feels that the case for using mFDR in preference to FDR needs clarification, and also does not find the material in Section 7.2 all that convincing. The main concern of the Associate Editor is the length of the paper (the current version is around one third longer than the previous version). I share this concern.

I am inviting you to prepare a revision which addresses all the points raised in the reports. When it is ready, please email the new version in pdf form directly to me, along with point-by-point responses to the reviewers' comments and a list of the changes you have made to the paper. Please follow, as closely as possible, the suggestions in my report for shortening the paper.

I look forward to receiving the next revision of your paper in due course.

Yours sincerely

Andy Wood
Joint Editor, JRSSB

**Editor's report on B6215R**

I share the opinions of the two reviewers that the current version is a major improvement and that the paper is now not too far from being suitable for JRSSB.

However, I agree with the Associate Editor that the current version is too long. I would like to see a reduction to somewhere between 28-30 pages in the current format and line spacing (i.e. I am requesting a bit more of a reduction than the AE). The AE has suggested removing section 7.3 which would account for about 3 pages. I would also like to see the writing made more concise throughout the paper; I have in mind a space reduction of a further 1-3 pages.

**Typos and minor corrections**

Page 2, line –11: a existing → an existing

Page 4, line 2: do well → does well

Page 7, line 6:

...test statistics Efron, 2005a,b → ...test statistics; see Efron, 2005a,b

Page 7, line 15: the wording is not quite right here. Should it be "This step-down procedure first compares..."

Page 7, line –6: "This theorem...". Which theorem?

Page 14, line –1: should the superscript $\theta$ in $T^{\theta}_{R=r}$ be removed? In any case, this notation does not appear to have been defined in the case of $T$.

Page 19, line 1: capture → captures

Page 33, lines 3 and 4: the reference 2004b should be presented in the same way as 2004a.

AE-Report on MS # B6215R

The revised version of the paper has much improved. My main criticism is about the length of the paper.

**Major Comment**
I am suggesting a couple of places where the paper could be shortened. Overall, I would like to encourage a paper which is about 3 pages shorter.

1. Section 5.1: could this be deleted or much shortened? See also the second minor comment by the referee.

2. Section 7.2 and 7.3: could one or both of these sections be deleted? (The referee probably wants to refer to Section 7.2 and not to Section 7.3). I would think that Section 7.3 is less interesting than Section 7.2 (but the choice should be done by the authors).

3. Section 8: the two last paragraphs which refer to the own works from the authors could be deleted.

**Minor Comments:**

- page 2, l.4: "... in a variety of ..."
- page 2, l.-11: "... an existing criterion..."
- Efron (2005b): has the paper appeared? Is this an early version of Size, Power, and False Discovery Rates (Efron) which is to appear in the Annals of Statistics?
- Meinshausen and Bühlmann has appeared (and the correct spelling and title of the paper should be used):
  Meinshausen and Bühlmann (2005). Lower bounds for the number of false null hypotheses for multiple testing of associations under general dependence structures. Biometrika 92, 893-907.

Ref 1

# Report on B6215R, "Multiple Hypothesis testing using the excess discovery count and alpha-investing rules"

The perhaps biggest single change in the revised version of manuscript B6215 is the omission of the EDC as a criterion for multiple testing. Now, the authors propose to control mFDR, a variation of the False Discovery Rate, with alpha-investing rules. This change will undoubtely help the reader to understand the proposed method. Also due to many other improvements, the manuscipt is now easier to read and comprehend. Overall, I think that the paper in its current form would be an interesting contribution to JRSSB.

Nevertheless, I still have some comments I want to share.

- The justification of mFDR over FDR on the bottom of page 3 is a bit weak in my opinion. In essence, the last sentence of page 3 says thats one would like to control neither FDR nor mFDR, but something entirely different. There is no really persuasive argument why mFDR is preferable to FDR. The only reason I can see is that mFDR is easier to control than FDR. Is this correct? If so, this could be stated frankly without diminishing the contribution.

- The example in section 7.2 is maybe not all that convincing. In my opinion, one of the key strengths of the proposed methodology is the ability to handle a stream of hypotheses (in the sense that it is unknown in advance how many hypotheses are being tested eventually). The example considers, however, only the case of a fixed set of hypotheses. In the setting of the considered example, the proposed method is very similar to the mentioned weighted hypotheses testing of Genovese, Roeder and Wasserman. It would be interesting to (a) compare weighted hypothesis testing and alpha-investing more closely and/or (b) provide an example where the unique advantage of alpha-investing is clearer.

Some minor comments:

- The attributes "uniform" and "universal" seem to be used interchangeably in section 5? A consistenct nomenclature would be helpful.

- I find section 5.1 a bit long for simply stating the obvious (that most variations of FDR are identical if $R = r$ if almost surely constant).

- Regarding the paragraph on page 18: "There is, however, ...". This paragraph is an interesting thought. In my opinion, however, this is not a particular danger with the proposed method but seems an inherent property of any method controlling FDR or variations thereof. (The more tests of "gravity does not exists" are placed into a multiple testing setup, the more of the other non-obvious hypotheses can be rejected when controlling FDR.) Could this paragraph be moved somewhere more appropriate?