

EDUARDO FERNANDEZ PASCUAL

From: Christian Ritz <ritz@nexs.ku.dk>
Sent: martes, 6 de junio de 2017 14:20
To: EDUARDO FERNANDEZ PASCUAL
Subject: Re: drc package & Invitation of coauthorship: seed germination R tool

Dear Eduardo,

in this case you need to use LL.2() (or any other model with an upper limit=1) as an upper limit of 1 corresponds to all seeds used in the experiment were germinating, whereas an upper limit $d < 1$ corresponds to all seeds that germinated, being less than 100% and then T50 becomes relative to a variable upper limit (which is not the case if the upper limit is fixed at 1).

I hope this makes sense?

Best wishes Christian

On 06-06-2017 12:39, EDUARDO FERNANDEZ PASCUAL wrote:

Dear Christian,

Thank you very much for your help.

We see now that to use the model you propose in the paper (with variable upper limit d) we should use LL.3, as LL.2 fixes the upper limit at 1.

But we are interested in a model in which t_{50} is fixed to represent all used seeds, because this produces estimates of the germination rate that can be compared across temperature treatments.

So do you think that using LL.2 (and W2.2) with type = event is appropriate?

Best regards,
Eduardo

From: Christian Ritz [<mailto:ritz@nexs.ku.dk>]
Sent: martes, 6 de junio de 2017 5:32
To: Emma Ladouceur <emmala@gmail.com>
Cc: EDUARDO FERNANDEZ PASCUAL <eduardofp.indurot@uniovi.es>
Subject: Re: drc package & Invitation of coauthorship: seed germination R tool

Dear Emma,

here are my comments:

1) Not sure I see/understand the difference between wordings: If you choose LL.2() then T50 is the time until 50% of all seeds used germinate. For your data: it took 41 days(?) until 50% of 25 seeds (used) had germinated.

If you all the time use LL.2() then T50 will all the time have the same (relative) interpretation.

However, it sounds like your research questions are about T50 relative to the maximum number of seeds *that will germinate*? Then you need to use a different model (the one defined in the paper I attached to the last e-mail): LL.3().

2) Models that are defined for negative time are meaningless in my opinion and hence not useful for describing germination. The logistic model is one such model. It's a common "mistake" to use such models where they don't really make sense. I would suggest omitting it entirely from the analyses.

Best wishes Christian

On 06-06-2017 10:23, Emma Ladouceur wrote:

Dear Cristian,

Thank you for your email and for attaching your helpful paper.

We have tried our analysis with the time-to-event method and we have the following questions that perhaps you may be able to comment on;

1. In your paper, you say *"In particular, this means that t_{50} only corresponds to the median germination time (the time when 50% of the seeds have germinated, i.e., $F(t_{50}) = 0.5$). In all other cases t_{50} corresponds to the 100d/2 percentile in the distribution of germination times. However, in any case t_{50} may be interpreted as the median germination time of seeds that germinated during the experiment. For instance, if d was estimated to be 0.2 it means that only 20% of the seeds used in the experiment germinated and if the resulting t_{50} was estimated to be 200 days it means that 10% of the seeds used in the experiment had germinated after 200 days."*

We interpret this to mean that the method calculates the T50 to be the time needed for the germination of 50% of the seeds germinated within the experiment, and not 50% of the seeds used in the experiment.

In step 3 of our analysis, we compare the $1/t_{50}$ of different temperature treatments. For this we understand that we need all the t_{50} s to refer to the same amount of seeds, i.e. 50% of the total sample and not 50% of the seeds germinated in each case.

To test this, we have created an example dataset (attached) in which final germination is 52%. Half of this (26%) is reached between day 17 and day 18. 50% of the total sample germinated between day 40 and day 41. We have fitted a binomial and a time-event model to the data (script attached), and the estimated t_{50} is 41.936 for the former and 41.2607 for the latter. So, both models seem to be calculating t_{50} as 50% of the total sample.

Could you please clarify this?

2. Your paper also states *"Note that the restriction rules out models such as the Gompertz, logistic and normal models even though the former two were proposed by Torres and Frutos (1989, 1990) and the latter even more recently by Onofri et al. (in press)."*

Some colleagues had given us feedback that we should add the logistic / Boltzmann model as one of the potential models in the model comparison, as this model seems

to be popular with some seed researchers. We have been able to fit this model using the time-event method, but we should remove it in any case, correct?

Thank you again for your comments.

Best wishes,

Emma

Emma Ladouceur
PhD Candidate
Museo delle Scienze, Trento, Italy
m +39 380 121 7641 | e emma.ladouceur@muse.it
w <https://emma-ladouceur.github.io/conservation/>

On Jun 2, 2017, at 12:07 AM, Christian Ritz <ritz@nexs.ku.dk> wrote:

Dear Emma,

thank you for the e-mail.

I looked into the manuscript and could only find very little description of the example data you used (I think more is needed in the manuscript). However, from looking at the data file it seems to that your data are event times, which should not be analyzed using a binomial model. Please see this paper for more details.

So I think you need to rethink/revise this part of your approach entirely using the appropriate statistical model for the germination data. Question 4 from the reviewer will then no longer be relevant.

Best wishes Christian

On 01-06-2017 13:52, Emma Ladouceur wrote:

Dear Dr. Cristian Ritz,

I am a Canadian PhD candidate based in Trento, Italy involved in the Marie Curie Native Seed Science Technology and Conservation (NASSTEC.eu) Initial Training Network. I am working with Eduardo Fernández-Pascual, one of our post-docs, on an R script that aims to standardize the calculation seed germination traits as one aspect of my work. In this analysis, and in other seed germination related analysis we have been using your helpful R package, *drc*.

We submitted this work on seed germination traits as an [Application](#) to the *Journal of Methods in Ecology & Evolution*, and while the work was rejected—it overall received relatively positive

comments, but reviewers requested we fix a few statistical issues and provide more detail on the specifics behind the statistics we are applying. It was pointed out we should try to engage a statistician to help us revise the manuscript as it was clear from the manuscript we are not statisticians.

We have also written to Vito Muggeo, the author of the other main package we use- *segmented*. Vito has responded he cannot join co-authorship but he will try to answer some of our *segmented* specific questions.

We revised all aspects of the manuscript we were able to without a statistician, and we summarised the reviewer comments we were not able to address into 5 short questions in a document, 'questions_from_review'.

Here in this email, i attach a compressed folder containing the revised manuscript, R script, example dataset, and the summarised questions. Eduardo and I would like to ask you if you would be interested or able to join us in co-authorship in this manuscript, and/or if you might be able to help us with any of the statistical questions we have compiled here. The questions document contains questions relevant to both *segmented* and *drc*, and we don't expect you to be able to respond to the *segmented* related questions but please do not hesitate to make any comments on this if you have them. Questions #4-5 specifically relate to *drc*.

Best wishes,

Emma Ladouceur

Emma Ladouceur
PhD Candidate
Museo delle Scienze, Trento, Italy
m +39 380 121 7641 | e
<https://emma-ladouceur.github.io/conservation/emma.ladouceur@muse.it>

<ritz_pipper_streibig2013.pdf>

