Subject: Re: RD update
From: Joseph Toscano <joseph.toscano@villanova.edu>
Date: 12/15/2017 10:04 AM
To: Jie Lu <lujie12345@gmail.com>
CC: Jont Allen <jont1@mimosaacoustics.com>

An updated SCO script is attached. This now tests for interactions between phonType x phon, and phonType x subject. Results are similar to what we saw before. First, for the fixed effects, the results are the same as the ANOVA. Significant main effects for RD and phonPosition; no interactions. You can use the coefficients from model8a in the script to compare effect sizes for RD and phonPosition if you'd like.

For the random effects (i.e., effects involving subject and the specific phoneme), we see the following significant effects:

Subject main effect (i.e. individual subjects vary in their error rate; p<.001)
Subject x phonPosition interaction (p<.001)
Subject x phonType interaction (p<.001)
Subject x phonPosition x phonType interaction (p<.001)

All of this indicates that individual subjects vary in their overall error rate, whether they make more errors in initial vs. final position, and whether they make more errors with consonants or vowels. The subject x phoneme interaction is still not significant. However, I don't think this is as critical as the main effect of subject (i.e. subjects are idiosyncratic in their error rates), and the interactions suggest that they do vary in which types of errors they make.

Jie: I can help you write out how to report these effects if you'd like. It is a bit different from the way the ANOVA is reported. Let me know if all of this makes sense first.

Joe

From: Jie Lu <lujie12345@gmail.com>
Date: Thursday, December 14, 2017 at 9:00 PM
To: Joseph Toscano <joseph.toscano@villanova.edu>
Cc: Jont Allen <jont1@mimosaacoustics.com>
Subject: Re: RD update

Hi Professor Toscano,

Thanks for helping me to write it correctly. I will copy your example down in my thesis.

Right, the F statistics and the p values do not change. Maybe for ANOVA, the transform just makes the sum of squares larger.

Another question is that I didn't find the coefficients as of value 1.37 and 0.26 for RD and Initial in the regression model output.
I tried your suggestion to calculate effect size, the eta squared, from the ANOVA output for trial number=1.

RD: 0.6738/0.5097=1.321954
Initial: 0.04597/0.06644=0.6919025

It seems that the subject group factor has about 1.910607 times effect size of the position factor. But I am not sure whether I have calculated it wrong.

Thanks,
Jie

On Thu, Dec 14, 2017 at 2:16 PM, Joseph Toscano <joseph.toscano@villanova.edu> wrote:

Hi Jie,

I think sticking with minN=10 is fine for now. Yes, you lose a couple subjects, but it doesn't change the pattern of results.

Here is how I would describe the pattern of results, with the standard method of reporting ANOVA statistics in text:

"The ANOVA revealed a main effect of reading disability ($F(1,13)=17.19$, $p=.001$), with more errors for the reading disabled listeners than for the control listeners. There was also a main effect of phoneme position ($F(1,13)=8.99$, $p=.01$), with more errors for final position than for initial position. None of the other effects or interactions were significant."

Also, two comments about the text in 2.6 you sent:

1. The arcsine transform you give there isn't in the script I sent. You would have to add it back in. This would change the sums of squares values in the ANOVA, but I don't think it would change the F statistics or p-values.
2. "we also fitted the data". It's the other way around: we fit a statistical model to the data. So, I would say we "fit a model".

Joe

From: Jie Lu <lujie12345@gmail.com>
Date: Thursday, December 14, 2017 at 2:43 PM
To: Joseph Toscano <joseph.toscano@villanova.edu>
Cc: Jont Allen <jont1@mimosaacoustics.com>
Subject: Re: RD update

As we discussed, I need to add in some sections about the correct ANOVA and regression and all of the explanation. These would appear in chapter 3.

Best,
Jie

On Thu, Dec 14, 2017 at 1:39 PM, Jie Lu <lujie12345@gmail.com> wrote:

I updated section 2.6 to briefly summarize the situation of stats testing.

Jie

On Thu, Dec 14, 2017 at 1:36 PM, Jie Lu <lujie12345@gmail.com> wrote:

Does it mean:
for individual subject, RD did make more errors; for interactions between individual subject and phone position, position was a significant effect but not interaction between subject group and phone position? Everything else was not significant?

Jie

On Thu, Dec 14, 2017 at 1:30 PM, Jie Lu <lujie12345@gmail.com> wrote:

I think I would still use trial number as 10 to update all the numbers. I saw that if trial =10 then

4 subjects are already excluded. If I exclude more subjects maybe it is not that good for the statistics? Is it better that I leave out table 3.2?

I ran your code and got the ANOVA result for 10-trial, but not sure how to explain it. Would you might helping me with that?
>aovModel <- aov(pAsin ~ rd*initial*phonType + Error(subject/(initial*phonType)), data=dataAvg)
>fixedEffectsANOVA
Error: subject
          Df Sum Sq Mean Sq F value Pr(>F)
rd   1 0.6738 0.6738   17.19 0.00115 **
Residuals 13 0.5097 0.0392

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: subject:initial
          Df Sum Sq Mean Sq F value Pr(>F)
initial   1 0.04597 0.04597   8.994 0.0103 *
rd:initial 1 0.00229 0.00229   0.447 0.5153
Residuals 13 0.06644 0.00511

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: subject:phonType
           Df Sum Sq Mean Sq F value Pr(>F)
phonType   1 0.00562 0.005616   1.027 0.329
rd:phonType 1 0.00001 0.000006   0.001 0.974
Residuals 13 0.07111 0.005470

Error: subject:initial:phonType
                 Df Sum Sq Mean Sq F value Pr(>F)
initial:phonType   1 0.00292 0.002917   0.720 0.411
rd:initial:phonType 1 0.00132 0.001322   0.327 0.577
Residuals   13 0.05263 0.004049

Thanks,
Jie

On Thu, Dec 14, 2017 at 6:56 AM, Joseph Toscano <joseph.toscano@villanova.edu> wrote:

> Great. I suggest updating the tables with the correct numbers from the ANOVAs in that script. Note that I did not run the specific analysis that's reported in Table 3.2. Can you do that given the script I sent?

> Joe

> **From:** Jie Lu <lujie12345@gmail.com>
> **Date:** Wednesday, December 13, 2017 at 10:56 PM
> **To:** Joseph Toscano <joseph.toscano@villanova.edu>
> **Cc:** Jont Allen <jont1@mimosaacoustics.com>
> **Subject:** Re: RD update

> Hi Professor Toscano,

> Thank you so much for helping me with all the stats and also correct my mistakes in the codes. I saw your email earlier but the network broke down again for a long time.

I wanted to say all your suggestions are great. If those were what we need to do to deposit my thesis then everything is fine by me.

Best,

Jie

On Wed, Dec 13, 2017 at 10:59 AM, Joseph Toscano <joseph.toscano@villanova.edu> wrote:

Jont,

Jie and I had a productive conversation today about getting all the stats correct for her thesis.

I've updated the R script to include an ANOVA for comparison to the regression results (Jie: The Error term in the ANOVA should be correct now). Note the following:

1. ANOVA is really not the correct model for binomial data (as mentioned before), but with the arcsine transform, it at least gives us something we can compare to the regression.
2. Because ANOVA requires a perfectly balanced dataset, I had to drop three additional subjects who did not have any C_fnl data at minN=10.
3. ANOVA does not report significance for random effects (subject and consonant), so it only allows us to test the significance of RD, consonant position, and their interaction.

With all that in mind, the ANOVA and regression yield the same pattern of effects: A significant main effect of RD (RD subjects make more errors than non-RD subjects), and no other effects. There is a marginal effect ($.05 < p < .1$) of consonant position in both analyses, but that's all.

I suggest reporting the results of the regression in Jie's thesis and in the paper we write up. ANOVA just requires so many assumptions that we would have to explain. Jie, how does this sound to you? I will look at the relevant part of your thesis and help with reporting the stats.

Also, in talking to Jie, I wanted to clear up something about how to interpret p-values: The only information we should take away from the p-value in terms of hypothesis testing is whether $p < .05$. The p-value itself doesn't tell you about the effect size. In an ANOVA framework, there is an additional statistic that can be computed, partial-η^2, which gives the effect size. With regression, however, the effect size is simply given by the coefficients in the model. So, for the current consonant data with minN=1, these are the effect sizes of the two fixed effects:

RD: 1.37
Position: 0.26

In other words, the effect of RD is about 5x the size of the effect of position. Note that those numbers are "unstandardized" coefficients in log-odds space, so they are not very intuitive, except for comparing them to each other. I could convert them into proportions/probabilities; let me know if that would be useful.

Joe

Re: RD update