

Re: Bootstrapping for confidence interval of effect size

Source: <http://sci.tech-archive.net/Archive/sci.stat.edu/2007-07/msg00054.html>

- *From:* Richard Ulrich <Rich.Ulrich@xxxxxxxxxxx>
 - *Date:* Mon, 30 Jul 2007 20:33:29 -0400
-

On Mon, 30 Jul 2007 10:05:58 -0000, amorphia <spam.ontoast@xxxxxxxxxxx> wrote:

Hi all,

Thanks for your reply Rich. I take your point about how you can't compare a confidence interval to a non-constant mean. However, I have seen it written in many places that comparing a confidence interval to a constant is equivalent to a hypothesis test.

- Right. You can create a CI by inverting the test.

So it seems to me that logically there ought to be some way to compare two confidence distributions (which is what you get with a bootstrap).

Well, yes, but not efficiently, and only if they are independent.

For example, you could compare every point (A_i) in bootstrap distribution A with every point (B_i) in bootstrap distribution B, and count how often $A_i > B_i$. That would give your your confidence that A is greater than B, right? It seems to me that it is exactly the same logic as comparing against a fixed constant.

Do it that way. That gives a better limit. Complications that arise are most likely with highly unequal Ns and variances, and I really don't know what is assured for that sort of case.

Your point about how the control groups are not equal doesn't apply -

Re: Bootstrapping for confidence interval of effect size

I think you missed a point in my original point – I am comparing the effect sizes of the two comparisons, not the means. There may be something wrong with that aspect of what I did but it isn't what you said! ;)

Sorry, but I still think you are missing a basic point of experimental design. There are three potential criteria for what group ends up better — which has the better final score, which has the better change score, or which has the better *regressed* change score. These are not assured to be in agreement, if the groups are not comparable at their starting scores,

These especially can arise when there are scaling problems, such as 'basement' and 'ceiling' effects. It is nice to assure the audience that such problems won't effect the analysis or conclusions.

There are separate problems that can arise in deciding what is the appropriate regression for the change, in special cases.

By the way, I have always in the past transformed my data too, but this data is just untransformable – you would see what I mean if you saw it.

[snip, previous]

—

Rich Ulrich, wpilib@xxxxxxxx
<http://www.pitt.edu/~wpilib/index.html>

.