**Thanks to the associate editor for these comments, which improved the chapter. Our responses are below.**

This is a pleasure to read. I learned quite a bit reading this chapter. I have only a few minor comments and questions.

**Thanks!**

On p. 4, my version of Word shows the bootstrapped statistics with the results as subscripts, instead of (e.g.) = 10 or = 50. It could be because my version of Word is so old. This happens in other places too. If yours looks right, it probably is my version.

**Corrected**

On p. 5, again my version of Word, the asterisk is too small and placed too high, so it’s hard (for my old eyes) to see that it’s an asterisk.

**Our eyes are old enough to have this same problem, but we’er not sure where asterisks are relevant for this. We hope that we can handle this at the proof stage.**

You don’t talk much about the bootstrapped median for the first example. The sampling distribution is very discrete because of the small sample size. Any hesitancy about using bootstrap in this situation?

**Small sample size is a problem, in all circumstances. The bootstrap is demonstrably better than other resampling methods (because the sampling frame is larger, due to sampling with replacement), as well as most parametric methods (unless normality is exactly correct, of course). So we have hesitancy, but less than for using any other available method.**

On p 8, I think you need another set of brackets in the p value following eq (4) to include the 1+ in the numerator.

**Corrected (I had to create a new equation –Will)**

P 9: “static” 🡪 “statistic”

**Corrected**

P 9 Another way to get people to understand is to remind them that the theory is that the sampling distribution involves sampling many times from F, but we don’t know F, so we use F(hat) as an approximation to F, which with reasonable sample size does look a lot like F, and from which we can sample.

**We have borrowed some of this writing, thank you.**

p. 10 at the top. I’m confused by the reference to “…an empirical distribution representing a null hypothesis.” Isn’t it a hypothetical distribution?

**\*\*\*\*\* Will and Patrick, I’m not sure I understand this comment. Any thoughts?**

**I changed   
“**An inferential procedure mimics *F* with an empirical distribution, , to assess the accuracy of  (or any other plug-in statistic). Conceptually,  stands in for *F* because we don’t know *F*, but we do know .**”**

**to**

**“**An inferential procedure mimics *F* with a theoretical distribution called , toassess the accuracy of  (or any other plug-in statistic). Conceptually,  stands in for *F* because we don’t know *F*, but we do know . In the world of resampling, is more specifically called an empirical distribution.**”**

Also p. 10; could you give a few more phrases or sentences about Smith and Kimball; without context it’s hard to understand the example.

**This is now slightly enhanced**

p. 12 Maybe this is picky, but Fisher didn’t use resampling; as you note he used a permutation test, which involves constructing the population of all possible samples. As you later state, resampling is useful when complete enumeration isn’t possible.

**“Resampling” in the modern usage refers to all resampling procedures (see Rodgers, 1999). Fisher did use resampling, other resampling methods followed.**

p. 13 Another picky point: For testing the null, it’s not a CI that’s constructed, but a region of (non)rejection.

**Adjusted our language in relation to this comment.**

On p. 15 will we get different results if we simplify and ignore the means and standard deviations (pretend means are 0 and sd are 1), and just have the correlation as the only unknown that we care about?

**\*\*\*\*\*\*\* I think so – Will and/or Patrick? I think we standardize here as part of the procedure (though it’s not noted in the text – adjust?)**

**--do you mean (a) standardize the sampling (answer: we don’t) or (b) standardize the bootstrap sample (answer: we implicitly do it calculating *r*\*obs)**

On p.20 you mention small sample sizes, and go straight to outliers, which lost me. And about small sample sizes: Does that violate the rationale for the bootstrap that F(hat) is a good approximation to F?

**Language is slightly adjusted, but is most cases the bootstrap is simply the best we have. Nothing solves the problem of a small sample size creating non-representative samples because of high likelihood of anomalies.**

p. 23 I’ve usually seen CDF in all caps, pdf lower case, but maybe that’s not universal.

**You’re correct, this is now corrected.**

A comment on Bayes: It may not be very important but Bayes use of MCMC is somewhat different than bootstrap methods. Bootstrap methods resample data; Bayesian methods resample parameter values. It’s true that MCMC was originally developed for sampling data, but Bayesians use it differently.

**We take this point, though it’s not relevant anywhere in our writing that we could find.**

I love your election analogy, with candidates and incumbents and elections!

**Thanks.**

On p 30, is the histogram in the left panel instead of the right? You might try darker shading to make it more prominent.

**Reference is now clarified. \*\*\*\*\* Will, do we want to make this darker? I’m find if it’s not.**

p. 32, Example 6: Gibbs sampler.

**Corrected**

P 36, you might also mention brms which is a passthrough to Stan and set up to run many standard models.

**That’s a good one. Added.**

P 36 typos:

“…allow that researcher to never….”

**Changed “**allof the researchwer to never**” to “**allows the research to never**” (I don’t know what happened here –I don’t have a cat)**

“…MCMCpack that handle common details…” **Now corrected**