There are numerous ways of calculating bootstrap CIs and p-values. The main issue is that it is impossible for the bootstrap to generate data under a null hypothesis. The permutation test is a viable resampling based alternative to this. To use a proper bootstrap you must make some assumptions about the sampling distribution of the test statistic.

A comment about lack of invariance of testing: it is entirely possible to find 95% CIs not inclusive of the null yet a p > 0.05 or vice versa. To have better agreement, the calculation of bootstrap samples under the null *must* be dope as β∗0=β^−β^∗β0∗=β^−β^∗ rather than β∗0=β^∗−β^β0∗=β^∗−β^. That is to say if the density is skewed right in the bootstrap sample, the density must be skewed left in the null. It is not really possible to invert tests for CIs with non-analytical (e.g. resampling) solutions such as this.

**normal bootstrap**

One approach is a **normal bootstrap** where you take the mean and standard deviation of the bootstrap distribution, calculate the sampling distribution under the null by shifting the distribution and using the normal percentiles from the null distribution at the point of the estimate in the original bootstrap sample. This is a reasonable approach when the bootstrap distribution is normal, visual inspection usually suffices here. Results using this approach are usually very close to robust, or sandwich based error estimation which is robust against heteroscedasticity and/or finite sample variance assumptions. The assumption of a normal test statistic is a stronger condition of the assumptions in the next bootstrap test I will discuss.

**percentile bootstrap**

Another approach is the **percentile bootstrap** which is what I think most of us consider when we speak of the bootstrap. Here, the bootstrapped distribution of parameter estimates an empirical distribution of the sample under the alternative hypothesis. This distribution can possibly be non-normal. A 95% CI is easily calculated by taking the empirical quantiles. But one important assumption is that such a **distribution is pivotal**. This means that if the underlying parameter changes, the shape of the distribution is only shifted by a constant, and the scale does not necessarily change. This is a strong assumption! If this holds, you can generate the "distribution of the statistic under the null hypothesis" (DSNH or F∗0F0∗) by subtracting the bootstrap distribution from the estimates, then calculating what percentage of the DSNH is "more extreme" than your estimate by using 2×min(F∗0(β^),1−F∗0(β^))2×min(F0∗(β^),1−F0∗(β^))

**Studentized bootstrap**

The easiest bootstrap solution to calculating pp-values is to use a studentized bootstrap. With each bootstrap iteration, calculate the statistic *and* its standard error and return the student statistic. This gives a bootstrapped student distribution for the hypothesis which can be used to calculate cis and p-values very easily. This also underlies the intuition behind the bias-corrected-accelerated bootstrap. The t-distribution shifts much more easily under the null since outlying results are downweighted by their corresponding high variance.

hypothesis testing - Computing p-value using bootstrap with R - Cross Validated

<https://stats.stackexchange.com/questions/20701/computing-p-value-using-bootstrap-with-r>