The jackknife and the bootstrap

1. Jackknifing bias and variance
Consider estimating a parameter \( \theta(F) \) from a sample \( X_1, \ldots, X_n \) of iid random variables from \( F \). The jackknife is a way of correcting the bias and estimating the variance of an estimator \( T_n \). Define a pseudovalue \( T^j = nT(x_1, \ldots, x_n) - (n-1)T(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \). We then estimate the bias of \( T_n \) by 
\[
\hat{\text{BIAS}} = \frac{1}{n} \sum_j (T^j - \bar{T}^*)
\]
and the variance by 
\[
\hat{\text{VAR}} = \frac{1}{n^2} \sum (T^j - \bar{T}^*)^2.
\]

2. The bootstrap
A bootstrap sample \( X^* \) is drawn iid from the empirical \( F_n \). Assuming the quantity of interest is some kind of centered quantity \( R(X, F) \), we estimate its distribution by the (known) distribution of \( R^* = R(X^*, F_n) \). This distribution can be computed exactly, approximated using small or large sample asymptotics, or simulated by drawing \( N \) samples from \( F_n \).

3. Expansions
For quantities \( R \) of interest that are invariant under permutation of the data we can represent the bootstrap distribution by a multinomial distribution. Let \( P_i^* = N_i^* / n \) where 
\[
N_i^* = \# \{ j : X_i^* = x_j \}.
\]
Then \( P^* = (P_1^*, \ldots, P_n^*) \sim \text{Mult}_n(n, \frac{1}{n}I) \), and we can write 
\[
R^* = S(P^*)
\]
for some \( S \). Let \( P_0 = \frac{1}{n}I \). Then 
\[
S(P^*) = S(P_0) + (P^* - P_0)^T U(P_0) + \frac{1}{2}(P^* - P_0)^T V(P^* - P_0)
\]
where \( U \) and \( V \) are the vector of first and matrix of second derivatives, respectively. We can choose \( P_0^T 1 = 0 \) and \( P_0^T V P_0 = 0 \). It follows that 
\[
E_s S(P^*) = S(P_0) + \frac{1}{n} \sum V_{ii}
\]
and 
\[
\text{Var}_s S(P^*) = \frac{1}{n^2} \sum U_{ii}^2.
\]

4. Bootstrap confidence intervals
(a) The bias-corrected percentile interval
Since 
\[
0.95 \hat{=} P(T^*_{(0.025N)} - T(F_n) \leq T(F_n) - T(F) \leq T^*_{(0.975N)})
\]
where \( T_{(k)} \) are ordered computations of \( T \) from the \( k \)th bootstrap sample, we get the corrected percentile interval for \( T(F) \) as 
\[
2T(F_n) - T^*_{(0.975N)}, 2T(F_n) - T^*_{(0.025N)}
\]
(b) Studentizing the root \( T(F_n) - T(F) \) yields better coverage probability.
(c) Consider a root (studentized or not) $R_n(x,t)$, and let $H_n(\cdot,F)$ be its cdf. The bootstrap estimates $H_n$ by $\hat{H}_n = H_n(\cdot,F_n)$. The corresponding confidence set is
\[ \left\{ t : \hat{H}_n(R_n(x,t)) \leq 1 - \alpha \right\}. \]
Let $R_{n,1}(\theta) = \hat{H}_n(R_n(x,\theta))$ and call its distribution $H_{n,1}$. A prepivoted confidence set results from estimating this distribution and constructing the set
\[ \left\{ t : \hat{H}_{n,1}(R_{n,1}(t)) \leq 1 - \alpha \right\} = \left\{ t : R_{n,1}(t) \leq \hat{H}_n^{-1}(\hat{H}_{n,1}^{-1}(1 - \alpha)) \right\}. \]
This has better coverage probability (at least asymptotically, but usually also in finite samples) than the intervals in (a) and (b).

If there is an exact pivot it tends to be produced by (possibly repeated) prepivoting. Generally prepivoting a non-studentized root yields a studentized root.

The prepivoted confidence interval can always be computed using a double bootstrap (i.e., bootstrapping the bootstrap samples).