

Exploring the Bootstrap Discrepancy

Russell Davidson^{1,2}

¹ Department of Economics,
McGill University,
855 Sherbrooke Street West,
Montreal, Quebec, Canada H3A 2T7 *russell.davidson@mcgill.ca*

² GREQAM,
Centre de la Vieille Charité,
13236 Marseille cedex 02, France

Abstract. Many simulation experiments have shown that, in a variety of circumstances, bootstrap tests perform better than current asymptotic theory predicts. Specifically, the discrepancy between the actual rejection probability of a bootstrap test under the null and the nominal level of the test appears to be smaller than suggested by theory, which in any case often yields only a rate of convergence of this discrepancy to zero. Here it is argued that the Edgeworth expansions on which much theory is based provide a quite inaccurate account of the finite-sample distributions of even quite basic statistics. Other methods are investigated in the hope that they may give better agreement with simulation evidence. They also suggest ways in which bootstrap procedures can be improved so as to yield more accurate inference.

Keywords: bootstrap discrepancy, bootstrap test, Edgeworth expansion

1 Introduction

Since the bootstrap was introduced by Efron (1979), its use by statisticians and econometricians has grown enormously; see for instance Horowitz (2001) for a useful survey. Asymptotic theory for the bootstrap has not been in short supply; after Bickel and Freedman (1981), landmark contributions have been Beran (1987) and (1988), and especially Hall (1992), in which a profound connection is established between bootstrap inference and Edgeworth expansions.

Although current asymptotic theory for the bootstrap accounts for many of the properties of bootstrap inference as discovered by simulation experiments, a recurrent phenomenon is that the bootstrap performs better than theory indicates. In this paper, I argue that the approximations provided by Edgeworth expansions are quite inadequate to describe the behaviour of bootstrap tests, and look at other methods which, while still inadequate, give quite different results. My hope is that the approach outlined here can give better insights into the properties of the bootstrap. One suggestion developed in this paper leads to the possibility of designing improved bootstrap schemes.

2 The Bootstrap Discrepancy

Suppose that a test statistic τ is designed to test a particular null hypothesis. The set of all DGPs that satisfy that hypothesis is denoted as \mathbb{M}_0 ; this set constitutes what we may call the null model. A bootstrap test based on the statistic τ approximates the distribution of τ under a DGP $\mu \in \mathbb{M}_0$ by its distribution under a bootstrap DGP that also belongs to \mathbb{M}_0 and can be thought of as an estimate of the true DGP μ .

We define the *bootstrap discrepancy* as the difference, as a function of the true DGP and the nominal level, between the actual rejection probability and the nominal level. In order to study it, we suppose, without loss of generality, that the test statistic is already in approximate P value form, so that the rejection region is to the left of a critical value.

The *rejection probability function*, or RPF, depends both on the nominal level α and the DGP μ . It is defined as

$$R(\alpha, \mu) \equiv \Pr_{\mu}(\tau < \alpha). \quad (1)$$

We assume that, for all $\mu \in \mathbb{M}_0$, the distribution of τ has support $[0, 1]$ and is absolutely continuous with respect to the uniform distribution on that interval. For given μ , $R(\alpha, \mu)$ is just the CDF of τ evaluated at α . The inverse of the RPF is the *critical value function*, or CVF, which is defined implicitly by the equation

$$\Pr_{\mu}(\tau < Q(\alpha, \mu)) = \alpha. \quad (2)$$

It is clear from (2) that $Q(\alpha, \mu)$ is the α -quantile of the distribution of τ under μ . In addition, the definitions (1) and (2) imply that

$$R(Q(\alpha, \mu), \mu) = Q(R(\alpha, \mu), \mu) = \alpha \quad (3)$$

for all α and μ .

In what follows, we assume that the distribution of τ under the bootstrap DGP, which we denote by μ^* , is known exactly. The bootstrap critical value for τ at level α is then $Q(\alpha, \mu^*)$. If τ is approximately (for example, asymptotically) pivotal relative to the model \mathbb{M}_0 , realisations of $Q(\alpha, \mu^*)$ under DGPs in \mathbb{M}_0 should be close to α . This is true whether or not the true DGP belongs to the null model, since the bootstrap DGP μ^* does so. The bootstrap discrepancy arises from the fact that, in a finite sample, $Q(\alpha, \mu^*) \neq Q(\alpha, \mu)$.

Rejection by the bootstrap test is the event $\tau < Q(\alpha, \mu^*)$. Applying the increasing transformation $R(\cdot, \mu^*)$ to both sides and using (3), we see that the bootstrap test rejects whenever

$$R(\tau, \mu^*) < R(Q(\alpha, \mu^*), \mu^*) = \alpha. \quad (4)$$

Thus the bootstrap P value is just $R(\tau, \mu^*)$, which can therefore be interpreted as a bootstrap test statistic.

We define two random variables that are deterministic functions of the two random elements, τ and μ^* , needed for computing the bootstrap P value $R(\tau, \mu^*)$. The first of these random variables is distributed as $U(0, 1)$ under μ ; it is

$$p \equiv R(\tau, \mu). \quad (5)$$

The uniform distribution of p follows from the fact that $R(\cdot, \mu)$ is the CDF of τ under μ and the assumption that the distribution of τ is absolutely continuous on the unit interval for all $\mu \in \mathbb{M}_0$. The second random variable is

$$q \equiv R(Q(\alpha, \mu^*), \mu) - \alpha = R(Q(\alpha, \mu^*), \mu) - R(Q(\alpha, \mu), \mu). \quad (6)$$

Let the CDF of q under μ conditional on the random variable p be denoted as $F(q | p)$. Then it is shown in Davidson and MacKinnon (2006) that the bootstrap discrepancy can be expressed as

$$\int_{-\alpha}^{1-\alpha} x \, dF(x | \alpha + x). \quad (7)$$

that is, the expectation of q conditional on p being at the margin of rejection at level α .

The random variable $q + \alpha$ is the probability that a statistic generated by the DGP μ is less than the α -quantile of the bootstrap distribution, conditional on that distribution. The expectation of q can thus be interpreted as the bias in rejection probability when the latter is estimated by the bootstrap. The actual bootstrap discrepancy, which is a nonrandom quantity, is the expectation of q conditional on being at the margin of rejection.

2.1 An asymptotically normal statistic

In some approaches to approximating the bootstrap discrepancy, it is assumed that the statistic is in asymptotically $N(0,1)$ rather than approximately $U(0,1)$ form. This is the case for the Edgeworth expansion approach considered in the next section. It is useful to define the random variables p and q in terms of new functions R_N and Q_N that respectively express the CDF and quantile function of the approximately normal statistic. Thus $R_N(x, \mu)$ is the CDF of the statistic under DGP μ , while $Q_N(\alpha, \mu)$ is the α -quantile. It is easy to see that $R_N(x, \mu) = R(\Phi(x), \mu)$ and $Q_N(\alpha, \mu) = \Phi^{-1}(Q(\alpha, \mu))$, where Φ is the CDF of the $N(0,1)$ distribution. If now we denote the approximately normal statistic by τ_N , we see that $q = R_N(Q_N(\alpha, \mu^*), \mu) - \alpha$ and $p = R_N(\tau_N, \mu)$; compare (6) and (5). Here we assume that the rejection region is to the left, as it would be for a statistic in P value form. Straightforward modifications can handle two-tailed tests or tests that reject to the right.

3 Approximations to the Bootstrap Discrepancy

3.1 Edgeworth expansion

Suppose that the statistic τ_N is computed using data generated by a DGP μ . Under the null hypothesis that τ_N is designed to test, we suppose that its distribution admits a valid Edgeworth expansion; see Hall (1992) for a complete treatment of Edgeworth expansions in connection with the bootstrap. The expansion takes the form

$$R_N(x, \mu) = \Phi(x) - n^{-1/2}\phi(x) \sum_{i=1}^{\infty} e_i(\mu) \text{He}_{i-1}(x). \quad (8)$$

Here ϕ is the density of the $N(0,1)$ distribution, $\text{He}_i(\cdot)$ is the Hermite polynomial of degree i (see for instance Abramowitz and Stegun (1965), Chapter 22, for details of these polynomials), and the $e_i(\mu)$ are coefficients that are at most of order 1 as the sample size n tends to infinity. The Edgeworth expansion up to order n^{-1} then truncates everything in (8) of order lower than n^{-1} .

The $e_i(\mu)$ can be related to the moments or cumulants of the statistic τ_N as generated by μ by means of the equation

$$n^{-1/2}e_i(\mu) = \frac{1}{i!} \mathbf{E}_{\mu}(\text{He}_i(\tau_N)). \quad (9)$$

The bootstrap DGP, μ^* , is realised jointly with τ_N , as a function of the same data. We suppose that this CDF can be expanded as in (8), with the $e_i(\mu)$ replaced by $e_i(\mu^*)$, and so the CDF of the bootstrap statistics is $R_N(x, \mu^*)$. We consider a one-tailed test based on τ_N that rejects to the left. Then, from (8), the random variable $p = R_N(\tau_N, \mu)$ is approximated by the expression

$$\Phi(\tau_N) - n^{-1/2}\phi(\tau_N) \sum_{i=1}^{\infty} e_i(\mu) \text{He}_{i-1}(\tau_N) \quad (10)$$

truncated so as to remove all terms of order lower than n^{-1} . Similarly, the variable q of (6) is approximated by $R'_N(Q_N(\alpha, \mu), \mu)(Q_N(\alpha, \mu^*) - Q_N(\alpha, \mu))$, using a Taylor expansion where R'_N is the derivative of R_N with respect to its first argument.

It is convenient to replace μ and μ^* as arguments of R_N and Q_N by the sequences \mathbf{e} and \mathbf{e}^* of which the elements are the $e_i(\mu)$ and $e_i(\mu^*)$ respectively. Denote by $\mathbf{D}_e R_N(x, \mathbf{e})$ the sequence of partial derivatives of R_N with respect to the components of \mathbf{e} , and similarly for $\mathbf{D}_e Q_N(\alpha, \mathbf{e})$. Then, on differentiating the identity $R_N(Q_N(\alpha, \mathbf{e}), \mathbf{e}) = \alpha$, we find that

$$R'_N(Q_N(\alpha, \mathbf{e}), \mathbf{e}) \mathbf{D}_e Q_N(\alpha, \mathbf{e}) = -\mathbf{D}_e R_N(Q_N(\alpha, \mathbf{e}), \mathbf{e}). \quad (11)$$

To leading order, $Q_N(\alpha, \mathbf{e}^*) - Q_N(\alpha, \mathbf{e})$ is $\mathbf{D}_e Q_N(\alpha, \mathbf{e})(\mathbf{e}^* - \mathbf{e})$, where the notation implies a sum over the components of the sequences. Thus the variable q can be approximated by

$$-\mathbf{D}_e R_N(Q_N(\alpha, \mathbf{e}), \mathbf{e})(\mathbf{e}^* - \mathbf{e}). \quad (12)$$

The Taylor expansion above is limited to first order, because, in the cases we study here, $Q_N(\alpha, \mu^*) - Q_N(\alpha, \mu)$ is of order n^{-1} . This is true if, as we expect, the $e_i(\mu^*)$ are root- n consistent estimators of the $e_i(\mu)$. From (8) we see that component i of $\mathbf{D}_e R_N(x, \mathbf{e})$ is $-n^{-1/2}\phi(x)\text{He}_{i-1}(x)$. To leading order, $Q_N(\alpha, \mathbf{e})$ is just z_α , the α -quantile of the $N(0,1)$ distribution. Let $l_i = n^{1/2}(e_i(\mu^*) - e_i(\mu))$. In regular cases, the l_i are of order 1 and are asymptotically normal. Further, let $\gamma_i(\alpha) = E(l_i | p = \alpha)$. Then the bootstrap discrepancy (7) at level α is a truncation of

$$n^{-1}\phi(z_\alpha) \sum_{i=1}^{\infty} \text{He}_{i-1}(z_\alpha)\gamma_i(\alpha). \quad (13)$$

3.2 Approximation based on asymptotic normality

If the distribution of a statistic τ_N has an Edgeworth expansion like (8), then it is often the case that τ_N itself can be expressed as a deterministic function of a set of asymptotically jointly normal variables of expectation 0; the special case of the next section provides an explicit example. If so, then the distribution of τ_N can be approximated by that of the same function of variables that are truly, and not just asymptotically, normal. This distribution depends only on the covariance matrix of these variables, and so can be studied at moderate cost by simulation.

In order to study the bootstrap discrepancy, one looks at the covariance matrix under the bootstrap DGP. This is normally an estimate of the true covariance matrix, and can often be expressed as a function of asymptotically normal variables, including those of which τ_N is a function. The joint distribution of the approximate p and q can then be used to approximate the bootstrap discrepancy, in what is of course a very computationally intensive procedure.

3.3 Approximation by matching moments

The Edgeworth expansion (8) is determined by the coefficients $e_i(\mu)$. These coefficients are enough to determine the first four moments of a statistic τ_N up to the order of some specified negative power of n . Various families of distributions exist for which at least the first four moments can be specified arbitrarily subject to the condition that there exists a distribution with those moments. An example is the Pearson family of distributions, of which more

later. A distribution which matches the moments given by the $e_i(\mu)$, truncated at some chosen order, can then be used to approximate the function $R_N(\tau_N, \mu)$ for both the DGP μ and its bootstrap counterpart μ^* . An approximation to the bootstrap discrepancy can then be formed in the same way as (13), with a different expression for $\mathbf{D}_e R_N(z_\alpha, \mathbf{e})$.

4 A Special Case: I. The Distribution

One of the simplest tests imaginable is a test that the expectation of a distribution is zero, based on an IID sample of n drawings, u_t , $t = 1, \dots, n$, from that distribution. We suppose that the expectation is indeed zero, and that the variance exists. The sample mean is $\hat{a} = n^{-1} \sum_t u_t$, and the sample variance, under the null that the expectation is zero, is $\hat{\sigma}^2 = n^{-1} \sum_t u_t^2$. A statistic that is asymptotically standard normal under the null is then $n^{1/2} \hat{a} / \hat{\sigma}$. Since this is homogeneous of degree 0 in the u_t , we may without loss of generality suppose that their true variance is 1. If we define the asymptotically normal variables $w_i = n^{-1/2} \sum_{t=1}^n (\text{He}_i(u_t) - \text{E}(\text{He}_i(u_t)))$, $i = 1, 2, \dots$, then the statistic can be written as

$$w_1 / (1 + n^{-1/2} w_2)^{1/2}. \quad (14)$$

On expanding the denominator by use of the binomial theorem, and truncating everything of order lower than n^{-1} , we can study the approximate test statistic

$$\tau_N = w_1 - \frac{1}{2} n^{-1/2} w_1 w_2 + \frac{3}{8} n^{-1} w_1 w_2^2. \quad (15)$$

4.1 The Edgeworth expansion

In order to apply the methodologies of Section 3.1 or Section 3.3, we have first to compute the expectations of the Hermite polynomials evaluated at τ_N . The quantities $e_i(\mu)$ can then be computed using (9) – here μ is the DGP that generates samples of n IID drawings from the given distribution. Working always only to order n^{-1} means that we need the $e_i(\mu)$ only to order $n^{-1/2}$. We see that

$$\begin{aligned} e_1(\mu) &= -\frac{1}{2} \kappa_3, & e_2(\mu) &= n^{-1/2} \kappa_3^2, \\ e_3(\mu) &= -\frac{1}{3} \kappa_3, & e_4(\mu) &= \frac{1}{12} n^{-1/2} (8\kappa_3^2 - 3 - \kappa_4) \\ e_5(\mu) &= 0, & e_6(\mu) &= \frac{1}{144} n^{-1/2} (9 + 8\kappa_3^2 - 3\kappa_4), \end{aligned} \quad (16)$$

where κ_3 and κ_4 are the third and fourth cumulants respectively of the distribution from which the u_t are drawn. All $e_i(\mu)$ for $i > 6$ are zero to order $n^{-1/2}$. The Edgeworth expansion of the distribution of τ_N is then, from (8),

$$\begin{aligned} R_N(x, \mu) &= \Phi(x) + \phi(x) \left(\frac{1}{6} n^{-1/2} \kappa_3 (1 + 2x^2) + n^{-1} \left(\frac{1}{48} x (8\kappa_3^2 + 3\kappa_4 - 81) \right. \right. \\ &\quad \left. \left. + \frac{1}{72} x^3 (63 - 8\kappa_3^2 - 9\kappa_4) - \frac{1}{144} x^5 (9 + 8\kappa_3^2 - 3\kappa_4) \right) \right) \end{aligned} \quad (17)$$

For many numerical illustrations, we will use the Pearson family of distributions. By adjusting four parameters, the first and second moments can be set to 0 and 1 respectively, and the third and fourth cumulants can be chosen from a wide range. In Table 1 are shown the maximum differences between the true CDF of a statistic of the form (14), as estimated using 100,000 simulations, and the asymptotically normal approximation (d_0), the approximation given by (17) through the $n^{-1/2}$ term (d_1), and through the n^{-1} term (d_2), for a range of sample sizes, and values of κ_3 and κ_4 . It can be seen that for large values of κ_3 and κ_4 , the Edgeworth approximations are not close to the true distribution until the standard normal approximation is also fairly close. What the table does not show is that the Edgeworth approximations are not necessarily bounded between 0 and 1, and are not necessarily increasing.

4.2 The asymptotic normality approximation

The distributions of the statistics (14) and (15), both functions of the asymptotically normal w_1 and w_2 , can be approximated by those of the same functions of two genuinely normal variables z_1 and z_2 , with the same first and second moments as those of w_1 and w_2 . We have $\text{var}(w_1) = 1$, $\text{var}(w_2) = 2 + \kappa_4$, and $\text{cov}(w_1, w_2) = \kappa_3$. Measures of the maximum differences between the true CDF and the approximation based on (14) are shown as d_3 in Table 1. The d_3 are smaller than the d_1 , especially for large values of κ_3 and κ_4 , and are of similar magnitude to the d_2 . Of course, the approximations are themselves true distributions, unlike the Edgeworth expansions.

4.3 Matching moments

The first four moments of the statistic (14) are implicitly given to order n^{-1} by (16). They are as follows:

$$\begin{aligned} \text{E}(\tau) &= -\frac{1}{2}n^{-1/2}\kappa_3 & \text{E}(\tau^2) &= 1 + 2n^{-1}\kappa_3^2 \\ \text{E}(\tau^3) &= -\frac{7}{2}n^{-1/2}\kappa_3 & \text{E}(\tau^4) &= 3 + 2n^{-1}(14\kappa_3^2 - \kappa_4 - 3). \end{aligned} \quad (18)$$

The distribution of (14) can be approximated by a Pearson distribution with those cumulants. Again, this is a true distribution. The maximum differences between the true CDF and this approximation are given as d_4 in Table 1.

5 A Special Case: II. The Bootstrap Discrepancy

5.1 The Edgeworth approximation

In order to compute the approximate bootstrap discrepancy (13), we make use of the differences $e_i(\mu^*) - e_i(\mu)$ between the coefficients of the Edgeworth expansion of the bootstrap statistic and that of the statistic itself. The

$e_i(\mu^*)$ are given by the expressions in (16) with κ_3 and κ_4 replaced by the estimates $\hat{\kappa}_3$ and $\hat{\kappa}_4$ used, explicitly or implicitly, in the bootstrap DGP. The most obvious bootstrap DGP is a resampling DGP in which the elements of a bootstrap sample are drawn at random, with replacement, from the u_t after centring. Since the statistic is scale invariant, the distribution of the bootstrap statistics would be the same if we resampled the $(u_t - \hat{a})/\hat{\sigma}$. The third cumulant of the bootstrap distribution is then the third moment of the rescaled quantities, and the fourth cumulant is their fourth moment minus 3. Some algebra then shows that

$$n^{1/2}(\hat{\kappa}_3 - \kappa_3) = w_3 - \frac{3}{2}\kappa_3 w_2 - \frac{3}{8}n^{-1/2}(8w_1 w_2 + 4w_2 w_3 - 5\kappa_3 w_2^2) \quad (19)$$

$$\begin{aligned} n^{1/2}(\hat{\kappa}_4 - \kappa_4) &= w_4 - 4\kappa_3 w_1 - 2\kappa_4 w_2 \quad (20) \\ &\quad - n^{-1/2}(6w_1^2 - 8\kappa_3 w_1 w_2 + 4w_1 w_3 + 3(1 - \kappa_4)w_2^2 + 2w_2 w_4) \end{aligned}$$

Thus, from the formulas (16), we can see that

$$l_1 = n^{1/2}(e_1(\mu^*) - e_1(\mu)) = -\frac{1}{4}(2w_3 - 3\kappa_3 w_2) + O_p(n^{-1/2}); \quad l_3 = \frac{2}{3}l_1$$

while all l_i for $i \neq 1, 3$ are of order lower than unity. By definition, the w_i are (jointly) asymptotically normal. The variance of w_1 is 1, and so $E(w_i | w_1) = w_1 E(w_i w_1)$. Now

$$E(w_1 w_2) = n^{-1} \sum_{t=1}^n E(\text{He}_1(u_t) \text{He}_2(u_t)) = E(u_t^3 - u_t) = \kappa_3.$$

Similarly, $E(w_1 w_3) = \kappa_4$. The $\gamma_i(\alpha)$ used in the approximate expression (13) for the bootstrap discrepancy are the expectations of the l_i conditional on the event $p = \alpha$. By (10), the variable p is approximated to leading order by $\Phi(\tau_N)$, and, from (15), this is $\Phi(w_1)$, again to leading order. Thus the conditioning event can be written as $w_1 = z_\alpha$. It follows that

$$\gamma_1(\alpha) = -\frac{1}{4}z_\alpha(2\kappa_4 - 3\kappa_3^2) \quad \text{and} \quad \gamma_3(\alpha) = \frac{2}{3}\gamma_1(\alpha)$$

with error of order lower than 1, all other $\gamma_i(\alpha)$ being of lower order. For our special case, therefore, the bootstrap discrepancy at level α , as approximated by (13), is

$$\frac{1}{12}n^{-1}\phi(z_\alpha)(3\kappa_3^2 - 2\kappa_4)z_\alpha(1 + 2z_\alpha^2) \quad (21)$$

We see that this expression vanishes if $3\kappa_3^2 - 2\kappa_4 = 0$. This is true for the normal distribution of course, for which all cumulants of order greater than 2 vanish. But it is true as well for many other commonly encountered distributions. Among these, we find the central chi-squared, exponential, Pearson Type III, and Gamma distributions.

Table 2 gives the maximum differences (d_1) between the actual discrepancy, as evaluated using a simulation with 100,000 replications with 399 bootstrap repetitions, and the approximate discrepancy (21), again for various

sample sizes and various cumulant values. When both κ_3 and κ_4 are zero, the data are standard normal. Some other distributions for which $3\kappa_3^2 - 2\kappa_4 = 0$ are also given.

The approximate discrepancy (21) is an odd function of z_α . The discrepancies approximated by simulation often do not seem to share this property even roughly.

5.2 Matching moments

In this approach, the function $R_N(x, \mu)$ is approximated by the CDF of a Pearson distribution, characterised by the four moments (18). Denote this approximation by $R_N(x, \kappa_3, \kappa_4)$. An approximation to the bootstrap discrepancy can be found exactly as in the preceding subsection. Analogously to (12), we approximate q by

$$-\sum_{i=3}^4 \frac{\partial R_N}{\partial \kappa_i}(Q_N(\alpha, \kappa_3, \kappa_4), \kappa_3, \kappa_4)(\kappa_i^* - \kappa_i). \quad (22)$$

But, of the four moments (18), only the fourth depends on κ_4 , with κ_4 multiplied by n^{-1} . From (20), $\hat{\kappa}_4 - \kappa_4 = O_p(n^{-1/2})$, and so only the term with $i = 3$ contributes to (22) to order n^{-1} . To leading order, $Q_N(\alpha, \kappa_3, \kappa_4) = z_\alpha$, and so the approximate bootstrap discrepancy is

$$\frac{1}{2}n^{-1/2} \frac{\partial R_N}{\partial \kappa_3}(z_\alpha, \kappa_3, \kappa_4)(3\kappa_3^2 - 2\kappa_4)z_\alpha, \quad (23)$$

since, from (19), $E(\hat{\kappa}_3 - \kappa_3 | w_1 = z_\alpha) = (2\kappa_4 - 3\kappa_3^2)z_\alpha/2$. Column d_2 in Table 2 gives the maximum differences between the actual discrepancy and (23). Of course, it coincides with column d_1 for all cases with $3\kappa_3^2 - 2\kappa_4 = 0$.

6 Designing a Better Bootstrap DGP

6.1 Theoretical considerations

It can be seen both from (7) and the discussion in the previous section of the Edgeworth approximation of the bootstrap discrepancy that its rate of convergence to 0 as $n \rightarrow \infty$ is faster if the bootstrapped statistic is uncorrelated with the determinants of the bootstrap DGP. This is often easy to realise with a parametric bootstrap, since a statistic that tests a given null hypothesis is often asymptotically independent of parameters estimated under that null; see Davidson and MacKinnon (1999). But with a nonparametric bootstrap like the resampling bootstrap studied in section 4, it is not obvious how to achieve approximate independence of the statistic and the bootstrap DGP, as shown by the fact the cumulant estimates given in (19) and (20) are correlated with the statistic (15), with the result that the discrepancy (21) is of order n^{-1} .

However, the fact that the Edgeworth approximation (17) depends on just two parameters of the DGP, κ_3 and κ_4 , suggests that it might be possible to construct a parametric bootstrap using just these parameters. For instance, the elements of a bootstrap sample could be drawn from the Pearson distribution with expectation 0, variance 1, and third and fourth cumulants given by those estimated using the u_t . The Edgeworth approximation of the bootstrap discrepancy (21) would be unchanged, although the actual bootstrap discrepancy could be smaller or greater than that of the ordinary resampling bootstrap. Another possibility, that would involve no bootstrap simulations at all, would be to use for the bootstrap distribution the Pearson distribution with the moments (18) with the estimated cumulants. We will shortly explore these possibilities by simulation.

We now turn to the questions of why (19) and (20) are correlated with the statistic (15), and whether it is possible to find other cumulant estimates that are approximately uncorrelated with it. First, we look at estimation of the second cumulant, that is, the variance. The sample variance, $\hat{\sigma}^2$, always assuming that the true variance is 1, can be seen to be equal to $1 + n^{-1/2}w_2$, and, since $E(w_1w_2) = \kappa_3$, it too is correlated with τ_N unless $\kappa_3 = 0$. In fact, $\hat{\sigma}^2$, as a variance estimator, is *inefficient*, since it does not take account of the fact that, under the null, the expectation is 0.

An efficient estimator can be found by various means. Let m_k denote the uncentred moment of order k of the u_t . It can be shown that m_2 , m_3 , and m_4 can be estimated efficiently by $\tilde{m}_k \equiv \hat{m}_k - (\hat{m}_1\hat{m}_{k+1})/\hat{m}_2$, $k = 2, 3, 4$. Some algebra then shows that, to leading order,

$$n^{1/2}(\tilde{\kappa}_3 - \kappa_3) = w_3 - \kappa_4w_1 - \frac{3}{2}\kappa_3(w_2 - \kappa_3w_1) \quad (24)$$

$$n^{1/2}(\tilde{\kappa}_4 - \kappa_4) = w_4 - \kappa_5w_1 - 4\kappa_3w_1 - 2\kappa_4(w_2 - \kappa_3w_1). \quad (25)$$

Here κ_5 is the fifth cumulant. It can be shown that $E(w_1w_4) = \kappa_5 + 4\kappa_3$, and that, consequently, (24) and (25) are uncorrelated with w_1 . Since $\tilde{\sigma}^2$ is more efficient than $\hat{\sigma}^2$, it makes sense to bootstrap the statistic $n^{1/2}\hat{a}/\tilde{\sigma}$ rather than $n^{1/2}\hat{a}/\hat{\sigma}$. To leading order, this statistic is also equal to w_1 , and is thus uncorrelated with $\tilde{\kappa}_3$ and $\tilde{\kappa}_4$.

A bootstrap DGP that uses \tilde{m}_3 and \tilde{m}_4 can be constructed by using a Pearson distribution parametrised with first and second moments 0 and 1 respectively, and these estimators as third and fourth moments.

6.2 Simulation evidence

The first set of experiments concerns the bootstrap without simulation, in which the moments (18) are used to set up a Pearson distribution, which is then used to obtain a bootstrap P value. The simulation results show that the properties of this computationally simple bootstrap are very similar to those of the resampling bootstrap.

The next set of experiments again uses a Pearson distribution, but this time for the bootstrap disturbances. The moments of the distribution of the residuals determine a Pearson distribution, and the bootstrap disturbances are drawings from this. In a further set of experiments, the moments were estimated with the zero expectation imposed, as discussed in the previous subsection.

7 Conclusions

We have investigated various types of approximations to the bootstrap discrepancy, including the traditional Edgeworth expansion approximations, but not restricted to them. We find that all approaches that are implicitly or explicitly based on estimates of the moments of the disturbances are quantitatively not at all accurate, although their inaccuracies take on very different forms.

We consider bootstrap DGPs based on both unrestricted and restricted estimates of the first few moments of the disturbances, and find that these essentially parametric bootstraps compete well with the conventional resampling bootstrap. It appears that much remains to be learned about the determinants of the bootstrap discrepancy for any given procedure, as well as about different procedures.

References

- ABRAMOWITZ, M. and I. A. STEGUN (1965): *Handbook of Mathematical Functions*, Dover, New York.
- BERAN, R. (1987): Prepivoting to Reduce Level Error of Confidence Sets. *Biometrika* 74, 457-468.
- BERAN, R. (1988): Prepivoting Test Statistics: a Bootstrap View of Asymptotic Refinements. *Journal of the American Statistical Association* 83, 687-697.
- BICKEL, P. J. and D. A. FREEDMAN (1981): Some Asymptotic Theory for the Bootstrap. *Annals of Statistics* 9, 1196-1217.
- DAVIDSON, R. and MACKINNON, J. G. (1999): The Size Distortion of Bootstrap Tests. *Econometric Theory* 15, 361-376.
- DAVIDSON, R. and MACKINNON, J. G. (2004): *Econometric Theory and Methods*. Oxford, New York.
- DAVIDSON, R. and MACKINNON, J. G. (2006): The Power of Asymptotic and Bootstrap Tests. *Journal of Econometrics* 133, 421-441.
- EFRON, B. (1979): Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics* 7, 1-26.
- HALL, P. (1992): *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.
- HOROWITZ, J. L. (2001): The Bootstrap. In J. J. Heckman and E. E. Leamer (Eds.): *Handbook of Econometrics* Vol. 5, Amsterdam: North-Holland, 3159-3228.

n	κ_3	κ_4	d_0	d_1	d_2	d_3	d_4
50	2	3	0.042	0.011	0.005	0.017	0.012
100	2	3	0.027	0.005	0.004	0.009	0.007
50	6	40	0.181	0.127	0.080	0.059	0.095
100	6	40	0.106	0.061	0.035	0.039	0.053
50	13	175	0.617	0.587	0.376	0.326	0.284
100	13	175	0.437	0.410	0.305	0.210	0.316
1000	13	175	0.064	0.027	0.014	0.029	0.024

Table 1: Maximum differences between true distribution and various approximations: d_0 for $N(0,1)$, d_1 for $n^{-1/2}$ Edgeworth approximation, d_2 for n^{-1} approximation, d_3 for asymptotic normality approximation, d_4 for matching moments.

n	κ_3	κ_4	$3\kappa_3^2 - 2\kappa_4$	distribution	d_1	d_2
20	0	0	0	$N(0,1)$	0.003	0.003
50	0	0	0	$N(0,1)$	0.003	0.003
20	2.828	12	0	χ_1^2	0.022	0.022
50	2.828	12	0	χ_1^2	0.010	0.010
20	2	6	0	exponential	0.017	0.017
50	2	6	0	exponential	0.009	0.009
20	1	1.5	0	Gamma(4)	0.009	0.009
50	1	1.5	0	Gamma(4)	0.004	0.004
20	0	-1.2	-2.4	uniform	0.003	0.005
50	0	-1.2	-2.4	uniform	0.004	0.004
50	2	3	6	Pearson I	0.008	0.008
50	3	11	5	Pearson I	0.013	0.016
50	6	40	28	Pearson I	0.070	0.133
50	9	83	77	Pearson I	0.199	0.434
50	12	175	82	Pearson I	0.228	0.467
100	12	175	82	Pearson I	0.120	0.233
500	12	175	82	Pearson I	0.017	0.025

Table 2: Maximum differences between bootstrap discrepancy and Edgeworth approximation (d_1) and moment-matching approximation (d_2).

n	distribution	d_1	d_2	d_3
20	χ_1^2	0.029	0.033	0.052
50	χ_1^2	0.010	0.011	0.018
20	exponential	0.006	0.024	0.035
50	exponential	0.006	0.008	0.014
20	uniform	0.012	0.006	0.007
50	uniform	0.007	0.003	0.006

Table 3: Maximum P value discrepancies: resampling (d_1), Pearson with inefficient (d_2) and efficient (d_3) moment estimates.