Working Paper 01-3
Statistics and Econometrics Series 2
January 2001

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

EXPLICIT NONPARAMETRIC CONFIDENCE INTERVALS FOR THE VARIANCE WITH GUARANTEED COVERAGE

Joseph P. Romano and Michael Wolf*

Abstract

In this paper, we provide a method for constructing confidence intervals for the variance that exhibit guaranteed coverage probability for any sample size, uniformly over a wide class of probability distributions. In contrast, standard methods achieve guaranteed coverage only in the limit for a fixed distribution or for any sample size over a very restrictive (parametric) class of probability distributions. Of course, it is impossible to construct effective confidence intervals for the variance without some restriction, due to a result of Bahadur and Savage (1956). However, it is possible if the observations lie in a fixed compact set. We also consider the case of lower confidence bounds without any support restriction. Our method is based on the behavior of the variance over distributions that lie within a Kolmogorov-Smirnov confidence band for the underlying distribution. The method is a generalization of an idea of Anderson (1967), who considered only the case of the mean; it applies to very general parameters, and particularly the variance.
While typically it is not clear how to compute these intervals explicitly, for the special case of the variance we provide an algorithm to do so. Asymptotically, the length of the intervals is of order $n^{-1/2}$ in probability), so that, while providing guaranteed coverage, they are not overly conservative. A small simulation study examines the finite sample behavior of the proposed intervals.

Keywords: Confidence interval; empirical distribution function; standard deviation.

*Romano, Department of Statistics, Stanford University, Stanford, CA 94305, U.S.A., Phone: 650-7236326, e-mail: romano@stat.stanford.edu. Wolf, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-91-624.98.63, Fax: 34-91-624.98.49, e-mail: wolf@est-econ.uc3m.es; Research of the first author supported by the National Science Foundation, DMS 9704487. Research of the second author partly funded by the Spanish "Dirección General de Enseñanza Superior" (DGES), reference number PB98-0025.

EXPLICIT NONPARAMETRIC CONFIDENCE INTERVALS FOR THE VARIANCE WITH GUARANTEED COVERAGE

Joseph P. Romano *
Department of Statistics
Stanford University
Stanford, CA 94305

Michael Wolf
Dpto. de Estadística y Econometría
Universidad Carlos III de Madrid
28903 Getafe, SPAIN

Working Paper 01-3, Statistics and Econometrics Series 2,
January 2001

Key Words: confidence interval, empirical distribution function;

---

*Joseph P. Romano, Phone: (650) 723-6326, Fax: (650) 725-8977,
E-mail: romano@stat.stanford.edu.

Research of the first author supported by the National Science Foundation DMS 9704487, Research of the second author partly funded by the Spanish "Dirección General de Enseñanza Superior" (DGES), reference number PB98-0025.

standard deviation.

Abstract

In this paper, we provide a method for constructing confidence intervals for the variance that exhibit guaranteed coverage probability for any sample size, uniformly over a wide class of probability distributions. In contrast, standard methods achieve guaranteed coverage only in the limit for a fixed distribution or for any sample size over a very restrictive (parametric) class of probability distributions. Of course, it is impossible to construct effective confidence intervals for the variance without some restriction, due to a result of Bahadur and Savage (1956). However, it is possible if the observations lie in a fixed compact set. We also consider the case of lower confidence bounds without any support restriction. Our method is based on the behavior of the variance over distributions that lie within a Kolmogorov-Smirnov confidence band for the underlying distribution. The method is a generalization of an idea of Anderson (1967), who considered only the case of the mean; it applies to very general parameters, and particularly the variance. While typically it is not clear how to compute these intervals explicitly, for the special case of the variance we provide an algorithm to do so. Asymptotically, the length of the intervals is of order $n^{-1/2}$ (in probability), so that, while providing guaranteed coverage, they are not overly conservative. A small simulation study examines the finite sample behavior of the proposed intervals.

1 Introduction

Suppose $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) according to a cumulative distribution function (c.d.f) F on the line. Consider the problem of constructing a level $1-\alpha$ confidence interval for $\sigma^2(F)$, the variance of F. The distribution F is assumed to belong to a large class $\mathbf{F}$ of distributions. Clearly, $\mathbf{F}$ must be restricted somewhat since we are assuming that $\sigma^2(F)$ exists and is finite. For a general parameter $\theta(F)$, if I_n is a random interval, define its coverage level over $\mathbf{F}$ to be

$$\inf\{Prob_F\{\theta(F) \in I_n\} : F \in \mathbf{F}\}.$$

In fact, even if we assume $\mathbf{F}$ consists of all distributions F having finite moments of all orders, Bahadur and Savage (1956) proved the negative result that it is impossible to construct an effective nonparametric confidence interval for $\mu(F)$, the mean of F. That is, if I_n is a random interval (depending on $X_1, \ldots, X_n$) such that, even for one F, the probability under F that I_n is a bounded set is one, the the coverage level for $\theta(F)$ over F is zero. Thus, one cannot find an interval that is bounded (with probability one) whose coverage is at least the nominal level $1 - \alpha$. Similar arguments can be constructed to show that it is impossible to construct a conservative, yet bounded interval for $\sigma^2(F)$.

It is well-known that there are many inference methods that yield valid inference for *fixed* F, such as Efron's (1979) bootstrap, methods based on Edgeworth expansion, likelihood, and other resampling refinements. Typically, these methods yield intervals I_n of nominal level $1 - \alpha$ satisfying, for *fixed* F,

$$|Prob_F\{\sigma^2(F) \in I_n\} - (1 - \alpha)| = O(n^{-p}), \tag{1}$$

for some $p > 0$. In fact, $p = 1$ for intervals that are first-order accurate (that is, whose coverage error is of the same order as that provided by the normal approximation if the parameter of interest is $\mu(F)$), $p = 2$ for higher-order accurate intervals such as the symmetric bootstrap-t interval, and p can even be bigger by bootstrap iteration (under assumptions to ensure the validity of Edgeworth expansions). Unfortunately, all these intervals have the property that their coverage level over a nonparametric class $\mathbf{F}$ is zero.

The technical reason why these methods can misbehave so badly yet still satisfy (1) is that the convergence result holds for each fixed F rather than uniform over $\mathbf{F}$. Moreover, the result is of asymptotic nature ensuring, in general, correct coverage in the limit only. How soon the asymptotics 'kick in' is unknown in practice, since the answer not only depends on the parameter of interest and the interval type used, but also on the unknown distribution F.

The practical consequence is that, even with moderate sample sizes, methods that satisfy (1) with $p = 1$ or bigger can yield intervals that undercover by quite a bit; for example, see Section 4.4 of Shao and Tu (1995) and Section 4 of this paper.

Alternatively, there exist methods to construct confidence intervals with guaranteed coverage probability when the class of distributions is restricted to certain parametric families. A standard example is the normal theory interval (based on a chi-squared distribution of the scaled sample variance) which is exact in case the underlying distribution is normal. However, the disadvantage of these methods is that they lack robustness of validity in the sense that the coverage is not near the nominal level, unless the parametric assumptions are met. For example, the normal theory interval can undercover even in the limit if the underlying distribution is not normal.

The objective of this paper is to derive an interval with guaranteed finite sample coverage, satisfying

$$\sup_n \inf\{Prob_F\{\theta(F) \in I_n\} : F \in \mathbf{F}\} \geq 1 - \alpha, \tag{2}$$

as well as being not too big in terms of its length. Because of the Bahadur and Savage result, some restriction has to be made, at least when two-sided confidence intervals are desired. The assumption imposed then is that the unknown F has support in a fixed compact set, which we take to be $[0, 1]$ without loss of generality; otherwise, F is arbitrary. In the special case of lower one-sided intervals, we also consider F having support $(-\infty, \infty)$, so that no restrictions need to be made.

The remainder of the paper is organized as follows. Section 2 provides a general result for constructing confidence intervals for an arbitrary parameter $\theta(F)$. The problem is that it is not clear how to compute these intervals in general. In Section 3, this result is applied to the special case $\theta(F) = \sigma^2(F)$ and it is shown that an explicit computation of the intervals is indeed possible. The intervals not only have guaranteed coverage, but their length is of order $n^{-1/2}$ in probability. This order is of course the smallest possible because, even for the subfamily of normal distributions, the optimal intervals have lengths of this same order. Section 4 provides some small sample considerations via a simulation study. Finally, some conclusions are given in Section 5. All tables appear at the end of the paper.

2 General Confidence Intervals with Guaranteed Coverage

The goal of this section is to present a general method to construct conservative confidence intervals, that is, confidence intervals with guaranteed coverage. As will be seen shortly, while the method is valid in theory for an arbitrary parameter $\theta(P)$, it typically does not lead to intervals that can be computed explicitly. However, the explicit computation is feasible for the variance $\sigma^2(F)$ (and it is actually trivial for the mean $\mu(F)$). The method is based on restricting attention to distributions lying within a level $1-\alpha$ Kolmogorov-Smirnov (KS) confidence band for the c.d.f F. For c.d.f.s F and G, define the sup distance

$$d_{KS}(F, G) = \sup_{x \in \mathbf{R}} |F(x) - G(x)|.$$

Let $\hat{F}_n$ be the empirical c.d.f., that is, the discrete distribution which places mass $1/n$ at each of the X_i. The statistic $d_{KS}(\hat{F}_n, F)$ was introduced in the fundamental paper of Kolmogorov (1933) who also obtained the limiting distribution of $n^{1/2} d_{KS}(\hat{F}_n, F)$. Note that the sampling distribution of $n^{1/2} d_{KS}(\hat{F}_n, F)$ and its limiting distribution do not depend on F as long as F is continuous; for example, see Csáki (1984). Denote by $c_n(1-\alpha)$ the $1-\alpha$ quantile of the distribution of $n^{1/2} d_{KS}(\hat{F}_n, F)$ under F when F is any continuous distribution. This leads to the following KS uniform confidence bands for F of nominal level $1-\alpha$.

$$\hat{R}_{n,1-\alpha} = \{F \in \mathbf{F} : n^{1/2} d_{KS}(\hat{F}_n, F) \leq c_{n,1-\alpha}\}.$$

Note that for any F (continuous or otherwise)

$$Prob_F\{F \in \hat{R}_{n,1-\alpha}\} \geq 1 - \alpha,$$

where the inequality is an equality if and only if F is continuous.

This confidence band for F leads to the following construction of a conservative confidence set $I_{n,1-\alpha}$ for a general real-valued parameter $\theta(F)$. In words, the value θ is contained in $I_{n,1-\alpha}$ if and only if there is some distribution G in $R_{n,1-\alpha}$ for which $\theta = \theta(G)$. In general, this prescription gives a confidence set which need not be a confidence interval. For well-behaved parameters $\theta(F)$, on the other hand, this construction reduces to an interval given by

$$I_{n,1-\alpha} = [\inf_{G \in R_{n,1-\alpha}} \theta(G), \sup_{G \in R_{n,1-\alpha}} \theta(G)]. \tag{3}$$

The proof of the following proposition is immediate.

Proposition 2.1 $I_{n,1-\alpha}$ *is conservative in the sense that it satisfies (2).*

Another nice property this construction possesses is the following. If $g(\theta)$ is a monotone function of θ, then the interval for $g(\theta)$ is the interval for θ transformed by $g(\cdot)$.

The idea of constructing conservative confidence intervals in this way is not a new one and dates back to Anderson (1967) who considered intervals for the mean $\mu(F)$. However, for a general parameter $\theta(P)$ it is not clear how to compute the interval, since formula (3) involves finding the infimum and the supremum over an infinitely dimensional set. On the other hand, for certain parameters the explicit computations become feasible, and one of these parameters will be seen to be $\sigma^2(F)$.

3 Explicit Computation for the Variance

As mentioned before, the Bahadur and Savage (1956) result essentially implies that when two-sided confidence intervals for $\sigma^2(F)$ are desired, one has to restrict the support of F to be bounded. Hence, we take the support to be $[0, 1]$ without loss of generality. Later, we shall consider one-sided lower confidence intervals for $\sigma^2(F)$ which allows the distribution to have unbounded support.

3.1 Distributions with Compact Support

The family $\mathbf{F_0}$ is now taken to be the set of all distribution functions having support $[0,1]$. We now need to be more specific about the KS uniform confidence band $\hat{R}_{n,1-\alpha}$. The empirical c.d.f. based on the observations $X_1, \ldots, X_n$ is defined by

$$\hat{F}_n(t) = \frac{\#\{X_i \leq t\}}{n}.$$

Since $\mathbf{F_0}$ has compact support, the upper and lower bounds of the KS band can be taken to be proper distribution function themselves denoted by $\hat{F}_{n,1-\alpha,up}$ and $\hat{F}_{n,1-\alpha,low}$, say. Let $Y_1, \ldots, Y_m$ be the distinct values of $\{X_1, \ldots, X_n, 0, 1\}$, ordered from smallest to largest. Note that if F is a continuous distribution, then $m = n+2$ with probability one. On the other hand, if F is discrete, m can be smaller than $n+2$. Then, $\hat{F}_{n,1-\alpha,up}$ and $\hat{F}_{n,1-\alpha,low}$ are step functions with jumps at the Y_j only and

$$\hat{F}_{n,1-\alpha,U}(Y_j) = \min\{\hat{F}_n(Y_j) + c_n(1-\alpha), 1\} \quad j = 1, \ldots, m \tag{4}$$

and

$$\begin{aligned} \hat{F}_{n,1-\alpha,L}(Y_j) &= \max\{\hat{F}_n(Y_j) - c_n(1-\alpha), 0\} \quad j = 1, \ldots, m-1, \\ \hat{F}_{n,1-\alpha,L}(1) &= 1. \end{aligned} \tag{5}$$

For convenience of notation, in the remainder of the paper we shall suppress the subscript $1-\alpha$ and write $\hat{F}_{n,L}$ and $\hat{F}_{n,U}$.

According to the general Proposition 2.1, we can construct a conservative two-sided confidence interval for $\sigma^2(F)$ along the lines of (3), with $\theta(F)$ replaced by $\sigma^2(F)$, provided that the calculations can be carried out explicitly. We will now demonstrate that this indeed is possible. First of all note that, due to the assumption of the support being the compact set $[0,1]$, one can replace the infimum by a minimum and the supremum by a maximum in formula (3), so the interval of interest becomes

$$\begin{aligned} I_{n,1-\alpha} &= [\min_{G \in \hat{R}_{n,1-\alpha}} \sigma^2(G), \max_{G \in \hat{R}_{n,1-\alpha}} \sigma^2(G)] \\ &\equiv [\hat{\sigma}^2_{n,Min}, \hat{\sigma}^2_{n,Max}]. \end{aligned} \tag{6}$$

Next, we state a general result that will be needed shortly.

Proposition 3.1 *If F is in $\mathbf{F_0}$, then*

(i) $\mu(F) = 1 - \int_0^1 F(x)dx$.

(ii) For any constant a, $E_F((X_i - a)^2) = (1-a)^2 - 2\int_0^1 (x-a)F(x)dx$.

Proof of Proposition 3.1. Let $G(x)$ be a nonnegative function satisfying $G(x) = c + \int_0^x g(t)dt$. Then, by the general integration by parts formula (18.15) of Billingsley (1986) we get

$$\int_{(0,1]} G(x)dF(x) = F(1)G(1) - F(0)G(0) - \int_0^1 g(x)F(x)dx. \tag{7}$$

To prove (ii), choose $G(x) = (x-a)^2 = a^2 + \int_0^x 2(t-a)dt$. Equation (7) then implies

$$\begin{aligned} E_F((X_i - a)^2) &= a^2F(0) + \int_{(0,1]} (x-a)^2 dF(x) \\ &= a^2F(0) + (1-a)^2F(1) - a^2F(0) - \int_0^1 2(x-a)F(x)dx \\ &= (1-a)^2 - 2\int_0^1 (x-a)F(x)dx. \end{aligned}$$

The proof of (i) is analogous, choosing $G(x) = x = \int_0^x dt$. ■

We now consider the minimum variance $\hat{\sigma}^2_{n,Min}$. Intuition suggests that $\hat{\sigma}^2_{n,Min}$ will be the variance of a c.d.f. that among all c.d.f.s $F \in \mathbf{F_0}$ distributes as much mass as possible at a single point. Hence, it should be sufficient to restrict our attention to the class $\hat{\mathbf{F}}_{\mathbf{n,jump}} = \{\hat{F}_{n,jump,t} : 0 \le t \le 1\}$, where

$$\hat{F}_{n,jump,t}(x) = \begin{cases} \hat{F}_{n,L}(x) & \text{for } x < t \\ \hat{F}_{n,U}(t) - \hat{F}_{n,L}(t) & \text{for } x = t \\ \hat{F}_{n,U}(x) & \text{for } x > t \end{cases} \tag{8}$$

We note the following elementary fact.

Fact 3.1 *Consider the function $g_n(t) = \mu(\hat{F}_{n,jump,t})$, $0 \le t \le 1$. Then, $g_n(\cdot)$ is strictly increasing and continuous. In addition, $g_n(0) = \mu(\hat{F}_{n,U})$ and $g_n(1) = \mu(\hat{F}_{n,L})$.*

The next proposition shows that we can indeed restrict attention to the class $\hat{\mathbf{F}}_{\mathbf{n,jump}}$ in finding $\hat{\sigma}^2_{n,Min}$.

Proposition 3.2 *Let G be a c.d.f. in $\hat{R}_{n,1-\alpha}$. Then, $\sigma^2(\hat{F}_{n,jump,\mu(G)}) \leq \sigma^2(G)$.*

Proof: By construction, we have

$$G(x) \geq \hat{F}_{n,jump,\mu(G)}(x) \text{ for } x < \mu(G) \text{ and } G(x) \leq \hat{F}_{n,jump,\mu(G)}(x) \text{ for } x > \mu(G). \tag{9}$$

Using part (ii) of Proposition 3.1 twice yields

$$\begin{aligned} & E_{\hat{F}_{n,jump,\mu(G)}}[(X_i - \mu(G))^2] - \sigma^2(G) \\ &\quad = 2\int_0^1 (x-\mu(G))(G(x) - \hat{F}_{n,jump,\mu(G)}(x))dx \\ &\quad = 2\int_0^{\mu(G)} (x-\mu(G))(G(x) - \hat{F}_{n,jump,\mu(G)}(x))dx + \\ &\quad\quad\; 2\int_{\mu(G)}^1 (x-\mu(G))(G(x) - \hat{F}_{n,jump,\mu(G)}(x))dx. \end{aligned}$$

Relation (9) implies that both summands in the last equation are less than or equal to zero. The proof is completed by noting that $\sigma^2(\hat{F}_{n,jump,\mu(G)})$ is less than or equal to $E_{\hat{F}_{n,jump,\mu(G)}}((X_i - \mu(G))^2)$. ■

Therefore, $\hat{\sigma}^2_{n,Min} = \min\{\sigma^2(\hat{F}_{n,jump,t}) : 0 \leq t \leq 1\}$. At first sight, it is not clear how to compute this minimum over an infinite set. However, the solution turns out to be quite simple and insightful at the same time. The first step is to minimize over the jump functions that only jump at the Y_j. Recall that $Y_1, \ldots, Y_m$ are the distinct values of $\{X_1, \ldots, X_n, 0, 1\}$ in increasing order. So consider

$$\hat{\sigma}^2_{n,Min,approx} = \min\{\sigma^2(\hat{F}_{n,jump,Y_j}) : 1 \leq j \leq m\}. \tag{10}$$

Also, denote the minimizing index by j^*, so $\hat{\sigma}^2_{n,Min,approx} = \sigma^2(\hat{F}_{n,jump,Y_{j^*}})$. We ask whether we can further reduce the variance by moving the jump point to the left or to the right of Y_{j^*}.

Denote the mass of $\hat{F}_{n,jump,Y_{j^*}}$ at Y_{j^*} by m, that is, $m = \hat{F}_{n,jump,Y_{j^*}}(Y_j^*) - \hat{F}_{n,jump,Y_{j^*}}(Y_{j^*-1})$. If, starting with $\hat{F}_{n,jump,Y_{j^*}}$, we shift mass m from Y_{j^*} to

$Y_{j^*}+\epsilon$, then the mean will increase by $m\,\epsilon$ and the (uncentered) second moment will increase by $m(2\epsilon\, Y_{j^*}+\epsilon^2)$. Therefore, the variance will increase by

$$m(2\epsilon\, Y_{j^*}+\epsilon^2) - 2m\,\epsilon\,\hat{\mu}_{j^*} - m^2\epsilon^2, \tag{11}$$

where $\hat{\mu}_{j^*}$ denotes the mean of $\hat{F}_{n,jump,Y_{j^*}}$. Differentiating with respect to ϵ and equating to zero yields

$$\epsilon = \frac{\hat{\mu}_{j^*} - Y_{j^*}}{1-m}. \tag{12}$$

Taking the second derivative verifies this as a minimum. What have we learned? If $\hat{\mu}_{j^*} = Y_{j^*}$, the minimum variance is already achieved. Otherwise, we jump at $Y_{j^*}+\epsilon$ rather than Y_{j^*}. However, it is now easy to see that the overall minimum variance is given by $\sigma^2(\hat{F}_{n,jump,t^*})$, where t^* is the unique solution of $t = \mu(\hat{F}_{n,jump,t})$; the existence and uniqueness of this solution follows from Fact 3.1. Indeed, repeating the minimization exercise, starting with $\sigma^2(\hat{F}_{n,jump,t^*})$ instead of $\hat{F}_{n,jump,Y_{j^*}}$, will obviously yield a shift of zero. Thus, we have the nice feature that the c.d.f. with the minimizing variance has the jump point at its mean. In summary, we have proven the following proposition.

Proposition 3.3 *Among all c.d.f.s in $\hat{R}_{n,1-\alpha}$, the minimum variance is given by*

$$\hat{\sigma}^2_{n,Min} = \sigma^2(\hat{F}_{n,jump,t^*}),$$

where t^* is the unique solution of $t = \mu(\hat{F}_{n,jump,t})$.

To summarize, one would use the following algorithm to compute $\hat{\sigma}^2_{n,Min}$ in practice.

Algorithm 3.1 (Computation of $\hat{\sigma}^2_{n,Min}$)

1. Denote by $Y_1,\ldots,Y_m$ the distinct elements of $\{X_1,\ldots,X_n,0,1\}$ in increasing order.

2. Compute $\hat{\sigma}_{n,Min,approx} = \min\{\sigma(\hat{F}_{n,jump,Y_j}) : 1 \le j \le m\}$ Let j^* be the corresponding maximizing index and compute $t^* = Y_{j^*}+\epsilon$ with ϵ given by (12).

3. We have $\hat{\sigma}^2_{n,Min} = \sigma^2(\hat{F}_{n,jump,t^*})$.

We now turn our attention to $\hat{\sigma}^2_{n,Max}$. Intuitively, we should be able to restrict ourselves to c.d.f.s that allocate as much mass as possible in both tails. After all, without any restrictions we could maximize the standard deviation by placing mass 0.5 at 0 and mass 0.5 at 1. In this spirit, we will define a class of "cross functions" that start out as $\hat{F}_{n,U}$, then stay flat (or "cross over") at level p until hitting $\hat{F}_{n,L}$, and finish as $\hat{F}_{n,L}$. More formally, let $\hat{\mathbf{F}}_{\mathbf{n,cross}} = \{\hat{F}_{n,cross,p} : 0 \leq p \leq 1\}$, where

$$\hat{F}_{n,cross,p}(x) = \max\{\min\{p, \hat{F}_{n,U}(x)\}, \hat{F}_{n,L}(x)\}, \quad 0 \leq x \leq 1. \tag{13}$$

We note the following elementary fact.

Fact 3.2 *Consider the function $h_n(p) = \mu(\hat{F}_{n,cross,p})$, $0 \leq p \leq 1$. Then $h_n(\cdot)$ is strictly decreasing and continuous. In addition, $h_n(0) = \mu(\hat{F}_{n,L})$ and $h_n(1) = \mu(\hat{F}_{n,U})$.*

Before we can verify our intuition, one more result is needed. Denote by $[\hat{L}_{n,cross,p}, \hat{U}_{n,cross,p})$ the interval where $\hat{F}_{n,cross,p}$ is equal to p. More formally,

$$[\hat{L}_{n,cross,p}, \hat{U}_{n,cross,p}) = \{0 \leq x \leq 1 : \hat{F}_{n,cross,p}(x) = p\}.$$

Note that for $p = 1$ this would be a closed interval.

Proposition 3.4 *Let G be a c.d.f. in $\hat{R}_{n,1-\alpha}$. Then there exists a $\hat{F}_{n,cross,p} \in \hat{\mathbf{F}}_{\mathbf{n,cross}}$ such that*

$$\begin{aligned} G(x) &\leq F_{n,cross,p}(x) \quad \textit{for } x \leq \mu(\hat{F}_{n,cross,p}), \\ G(x) &\geq F_{n,cross,p}(x) \quad \textit{for } x \geq \mu(\hat{F}_{n,cross,p}). \end{aligned} \tag{14}$$

Proof: By construction, for any $F_{n,cross,p} \in \hat{\mathbf{F}}_{\mathbf{n,cross}}$, we have

$$\begin{aligned} G(x) &\leq F_{n,cross,p}(x) \quad \text{for } x \leq \hat{U}_{n,cross,p}, \\ G(x) &\geq F_{n,cross,p}(x) \quad \text{for } x \geq \hat{L}_{n,cross,p}. \end{aligned}$$

Hence, we just need to show that one can find a $F_{n,cross,p} \in \hat{\mathbf{F}}_{\mathbf{n,cross}}$ for which

$$\mu(\hat{F}_{n,cross,p}) \in [\hat{L}_{n,cross,p}, \hat{U}_{n,cross,p}]. \tag{15}$$

We trivially have that for $p_1 < p_2$, $\hat{L}_{n,cross,p_1} \leq \hat{L}_{n,cross,p_2}$ and $\hat{U}_{n,cross,p_1} \leq \hat{U}_{n,cross,p_2}$. Also, as p ranges from 0 to 1, the union of all $[\hat{L}_{n,cross,p}, \hat{U}_{n,cross,p})$ is $[0,1]$. Together with Fact 3.2, this implies that at least one $\hat{F}_{n,cross,p}$ satisfying relation (15) must exist. ∎

The next proposition shows that we can indeed restrict our attention to the class $\hat{\mathbf{F}}_{\mathbf{n,cross}}$ in finding $\hat{\sigma}^2_{n,Max}$.

Proposition 3.5 *Let G be a c.d.f. in $\hat{R}_{n,1-\alpha}$. Then, there exists a $\hat{F}_{n,cross,p} \in \hat{\mathbf{F}}_{\mathbf{n,cross}}$ such that $\sigma^2(\hat{F}_{n,cross,p}) \geq \sigma^2(G)$.*

Proof: By Proposition 3.4 one can find a $\hat{F}_{n,cross,p}$ such that relation (14) is satisfied; denote its mean by $\hat{\mu}_p$ for notational convenience. For this $\hat{F}_{n,cross,p}$ we then have by part (ii) of Proposition 3.1

$$\begin{aligned}
\sigma^2(\hat{F}_{n,cross,p}) - E_G((X_i - \hat{\mu}_p)^2) &= 2\int_0^1 (x-\hat{\mu}_p)(G(x) - \hat{F}_{n,cross,p}(x))dx \\
&= 2\int_0^{\hat{\mu}_p} (x-\hat{\mu}_p)(G(x) - \hat{F}_{n,cross,p}(x))dx + \\
&\quad 2\int_{\hat{\mu}_p}^1 (x-\hat{\mu}_p)(G(x) - \hat{F}_{n,cross,p}(x))dx.
\end{aligned}$$

Since relation (14) is satisfied, it follows that both summands in the last equation are bigger than or equal to zero. The proof is completed by noting that $E_G((X_i - \hat{\mu}_p)^2)$ is bigger than or equal to $\sigma^2(G)$. ∎

Therefore, $\hat{\sigma}^2_{n,Max} = \max\{\sigma^2(\hat{F}_{n,cross,p}) : 0 \leq p \leq 1\}$. At first sight, it is not clear how to compute this maximum over an infinite set. While the explicit maximization can be done, it is, unfortunately, not as simple and insightful as in the case of $\hat{\sigma}^2_{n,Min}$. To see why, start with a particular $\hat{F}_{n,cross,p}$ and ask whether one can increase the variance by changing p to $p+\epsilon$. Let us assume that the resulting change only means that the mass at $\hat{L}_{n,cross,p}$ increases by ϵ while the mass at $\hat{U}_{n,cross,p}$ decreases by ϵ (this will in general not be true;

e.g., decreasing a mass by ϵ could result in a negative mass). For notational convenience, let $u = \hat{U}_{n,cross,p}$, $l = \hat{L}_{n,cross,p}$, and $\hat{\mu}_p = \mu(\hat{F}_{n,cross,p})$. Then the mean will increase by $\epsilon(l-u)$ and the (uncentered) second moment will increase by $\epsilon(l^2-u^2)$. Therefore, the variance will increase by

$$\epsilon(l-u) - 2\,\hat{\mu}_p\,\epsilon(l-u) - \epsilon^2(l-u)^2.$$

Differentiating with respect to ϵ and equating to zero yields

$$\epsilon = \frac{\hat{\mu}_p - (l+u)/2}{u-l}.$$

Taking the second derivative verifies this as a maximum. Hence, if $\hat{\mu}_p = (l+u)/2$, the maximum is already achieved. The first thought is therefore that the maximum variance is given by $\sigma^2(\hat{F}_{n,cross,p^*})$, where p^* is the unique solution of $\mu(\hat{F}_{n,cross,p}) = (\hat{L}_{n,cross,p} + \hat{U}_{n,cross,p})/2$. However, such a solution does not always exist, although it is unique if it does. The reason is that $\mu(\hat{F}_{n,cross,p})$ decreases continuously as p increases from 0 to 1 while $(\hat{L}_{n,cross,p} + \hat{U}_{n,cross,p})/2$ is an increasing jump function; both $\hat{L}_{n,cross,p}$ and $\hat{U}_{n,cross,p}$ are increasing and can only be 0, one of the X_i, or 1. Even though not presented here, a counterexample can easily be constructed, that is, a case where $\mu(\hat{F}_{n,cross,p}) = (\hat{L}_{n,cross,p} + F_{n,cross,p})/2$ does not have a solution. If the solution exists, it has the nice interpretation that in order to maximize the variance we have to distribute the total mass in the left and right tail only, in such a way that the left and right masses start equally far away from the mean.

In practice, one can compute $\hat{\sigma}^2_{n,Max}$ using the following algorithm.

Algorithm 3.2 (Computation of $\hat{\sigma}^2_{n,Max}$)

1. As before, denote by $Y_1, \ldots, Y_m$ the distinct elements of $\{X_1, \ldots, X_n, 0, 1\}$ in increasing order. Define $p_{j,L} = \hat{F}_{n,L}(Y_j)$ and $p_{j,U} = \hat{F}_{n,U}(Y_j)$ for $1 \leq j \leq m$.

2. Compute $\hat{\sigma}_{n,Max,approx} = \max\{\sigma(\hat{F}_{n,cross,p_{j,E}}) : 1 \leq j \leq m, E = L, U\}$ Let $p_{j,E}$ be the corresponding maximizing index and $\hat{\mu}_{p_{j,E}} = \mu(\hat{F}_{n,cross,p_{j,E}})$. Find j_1 and j_2 satisfying $Y_{j_1} = \hat{L}_{n,cross,p_{j,E}}$ and $Y_{j_2} = \hat{U}_{n,cross,p_{j,E}}$.

3. In case $E = U$,

   - if $(Y_{j_1} + Y_{j_2})/2 \leq \hat{\mu}_{p_{j,E}} \leq (Y_{j_1+1} + Y_{j_2})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}})$.
   - if $\hat{\mu}_{p_{j,E}} < (Y_{j_1} + Y_{j_2})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}-\epsilon})$, where
   $$\epsilon = \frac{(Y_{j_1} + Y_{j_2})/2 - \hat{\mu}_{p_{j,E}}}{Y_{j_2} - Y_{j_1}}.$$
   - if $\hat{\mu}_{p_{j,E}} > (Y_{j_1+1} + Y_{j_2})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}+\epsilon})$, where
   $$\epsilon = \frac{\hat{\mu}_{p_{j,E}} - (Y_{j_1+1} + Y_{j_2})/2}{Y_{j_2} - Y_{j_1+1}}.$$

4. In case $E = L$,

   - if $(Y_{j_1} + Y_{j_2-1})/2 \leq \hat{\mu}_{p_{j,E}} \leq (Y_{j_1} + Y_{j_2})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}})$.
   - if $\hat{\mu}_{p_{j,E}} < (Y_{j_1} + Y_{j_2-1})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}-\epsilon})$, where
   $$\epsilon = \frac{(Y_{j_1} + Y_{j_2-1})/2 - \hat{\mu}_{p_{j,E}}}{Y_{j_2-1} - Y_{j_1}}.$$
   - if $\hat{\mu}_{p_{j,E}} > (Y_{j_1} + Y_{j_2})/2$, we have $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p_{j,E}+\epsilon})$, where
   $$\epsilon = \frac{\hat{\mu}_{p_{j,E}} - (Y_{j_1} + Y_{j_2})/2}{Y_{j_2} - Y_{j_1}}.$$

The rationale behind this somewhat complicated algorithm is as follows. Let p^* be the maximizing level, that is, $\hat{\sigma}^2_{n,Max} = \sigma^2(\hat{F}_{n,cross,p^*})$. From the previous discussion we know that p^* satisfies $\mu(\hat{F}_{n,cross,p^*}) = (\hat{L}_{n,cross,p^*} + \hat{U}_{n,cross,p^*})/2$, provided that such a solution exists. When such a solution does not exist, we already found the maximum standard deviation by $\sigma^2(\hat{F}_{n,cross,p_{j,E}})$. However, when the solution does exist, we can find it by updating $p_{j,E}$ properly, as outlined in the algorithm.

Proposition 3.6 *Among all c.d.f.s in $\hat{R}_{n,1-\alpha}$, the maximum variance is $\hat{\sigma}^2_{n,Max}$ as computed in Algorithm 3.2.*

Proof: The proof is analogous to the proof of Proposition 3.3, though somewhat more lengthy, and it is thus omitted. ■

3.2 Length of the Interval

We now show that the confidence interval for $\sigma^2(F)$ has length that is of order $n^{-1/2}$ in probability.

Theorem 3.1 *For any F in $\mathbf{F_0}$, $n^{1/2}(\hat{\sigma}^2_{n,Max} - \hat{\sigma}^2_{n,Min})$ is bounded in probability.*

To prove the theorem, the following lemma is needed. Let $m_k(F) = E_F(X_i^k)$ be the kth moment of F.

Lemma 3.1 *Suppose F and G are in $\mathbf{F_0}$ and $d_{KS}(F, G) \leq \epsilon$.*

(i) Then, $|m_k(F) - m_k(G)| \leq \epsilon$ for any $k \geq 0$.

(ii) Hence, $|\sigma^2(F) - \sigma^2(G)| \leq 3\,\epsilon$.

The proof of (i) follows by integration by parts (see Proposition 3.1 of Romano and Wolf, 2000). The proof of (ii) follows by writing $\sigma^2(F)$ as a function of moments and applying the triangle inequality.

Proof of Theorem 3.1: For any $F \in \hat{R}_{n,1-\alpha}$, $d_{KS}(\hat{F}_n, F) \leq n^{-1/2} c_{n,1-\alpha}$ so that by part (ii) of Lemma 3.1 for such an F,

$$|\sigma^2(F) - \sigma^2(\hat{F}_n)| \leq 3\, n^{-1/2} c_{n,1-\alpha}.$$

Hence, the triangle inequality implies $n^{1/2}(\hat{\sigma}^2_{n,Max} - \hat{\sigma}^2_{n,Min}) \leq 6\, c_{n,1-\alpha}$. But $c_{n,1-\alpha}$ is bounded because $c_{n,1-\alpha} \to c(1-\alpha)$, the $1-\alpha$ quantile of the limiting distribution of the Kolmogorov-Smirnov statistic. ∎

3.3 Distributions with Infinite Support

The family $\mathbf{F_1}$ is now taken to be the set of all distribution functions having support $(-\infty, \infty)$ and finite variance. For the reasons mentioned before, it is no longer possible to construct two-sided conservative confidence intervals for the variance with finite length. However, it will still be possible to construct one-sided lower confidence intervals with guaranteed coverage, also based on Proposition 2.1.

Since $\mathbf{F_1}$ has infinite support, the upper and lower bounds of the KS band—again denoted by $\hat{F}_{n,up}$ and $\hat{F}_{n,low}$, respectively—are no longer proper distribution functions themselves. More specifically, they are step functions with jumps at the data points X_i only and

$$\hat{F}_{n,U}(X_i) = \min\{\hat{F}_n(X_i) + c_{n,1-\alpha}, 1\} \quad i = 1, \ldots, n \tag{16}$$

and

$$\hat{F}_{n,L}(X_i) \quad = \quad \max\{\hat{F}_n(X_i) - c_{n,1-\alpha}, 0\} \quad i = 1, \ldots, n. \tag{17}$$

It is obvious that the interval of Proposition 2.1 (applied to the variance) now yields infinity as the upper bound resulting, effectively, in a one-sided lower confidence interval. As in the case of bounded distributions, the exact computation of the lower bound, that is, the minimum variance $\hat{\sigma}^2_{n,Min}$ within the KS band is possible.

As before, it is intuitively clear that $\hat{\sigma}^2_{n,Min}$ will be the variance of a c.d.f. that among all c.d.f.s $F \in \mathbf{F_1}$ distributes as much mass as possible at a single point. Hence, it should be sufficient again to restrict our attention to the class $\hat{\mathbf{F}}_{\mathbf{n,jump}} = \{\hat{F}_{n,jump,t} : -\infty < t < \infty\}$, where

$$\hat{F}_{n,jump,t}(x) = \begin{cases} \hat{F}_{n,L}(x) & \text{for } x < t \\ \hat{F}_{n,U}(t) - \hat{F}_{n,L}(t) & \text{for } x = t \\ \hat{F}_{n,U}(x) & \text{for } x > t \end{cases} \tag{18}$$

Not surprisingly, it turns out that the minimum variance can be found in basically the same way as in the case of distributions with bounded support. The proof of the following proposition is analogous to the proof of Proposition 3.3 and it is thus omitted.

Proposition 3.7 *Among all c.d.f.s in $\hat{R}_{n,1-\alpha}$, the minimum variance is given by*

$$\hat{\sigma}^2_{n,Min} = \sigma^2(\hat{F}_{n,jump,t^*}),$$

where t^* is the unique solution of $t = \mu(\hat{F}_{n,jump,t})$.

It is obvious that $X_{(1)} \leq t^* \leq X_{(n)}$ and therefore essentially the same algorithm as in the case of bounded distributions can be used to find t^* in practice. The only difference is that we do not add the values 0 and 1 to the observed data points, that is, now $Y_1, \ldots, Y_m$ are the distinct values of $X_1, \ldots, X_n$ only, arranged in increasing order.

4 Simulation Study

The goal of this section is to shed some light on the small sample properties of confidence intervals for the variance by means of simulations. We consider estimated coverage probability of intervals with nominal confidence levels 90% and 95% for samples of sizes $n = 10$, $n = 30$, and $n = 60$; we also look at estimated mean length of two-sided intervals. In addition to the conservative intervals proposed in this paper, we include the well-known normal theory interval—based on a chi-square distribution of the (scaled) sample variance in case of normal data—and the percentile and hybrid bootstrap intervals (e.g., Hall, 1992). The corresponding intervals are denoted by CONS, NORM, BOOT_P, and BOOT_H, respectively. Note that the bootstrap intervals are based on $B = 1,000$ resamples and that estimated coverage probabilities are based on 1,000 simulations for each scenario (the various confidence interval types are computed from the same simulated data).

We start by considering distributions with support $[0, 1]$ and two-sided confidence intervals for the variance. The distributions included in the study are Uniform on $[0, 1]$, the triangle distribution on $[0, 1]$ (which is the distribution of the average of two i.i.d. Uniform on $[0, 1]$ random variables), and a two-point distribution placing mass 0.95 at 0 and mass 0.05 at 1. The corresponding variances are easily seen to be 1/12, 1/24, and 19/400. Note that the normal and the hybrid bootstrap intervals are truncated at 0 and at 0.25 if necessary. The results are presented in Table 1. It is seen that our interval is the only which is conservative, that is, which always meets or exceeds the nominal coverage level. The bootstrap intervals undercover consistently while the normal theory interval is conservative for the first two distributions but undercovers for the two-point distribution. The price that our interval pays in achieving guaranteed coverage is that the intervals are quite wide in the sense that the estimated coverage probability is always equal to 1.

Note that we also considered the mean length of the intervals and the corresponding results are given in Table 2. Our interval is, of course, wider than the normal and the bootstrap intervals. However, apart from the two-point distribution where larger sample sizes seem to be needed, the length is seen to decrease with the sample size (according to the asymptotic theory). Therefore, our interval clearly improves upon the trivial interval with guaranteed coverage given by [0, 0.25].

Next, we consider distributions with infinite support and one-sided lower confidence intervals for the variance. The distributions included in the study are Normal(0, 1), Exponential(1), and a three point distribution placing mass 0.1 at 0 and mass 0.4 at both -1 and 1. The corresponding variances are easily seen to be 1, 1, and 0.9. The results are presented in Table 3. Now, our interval and the percentile bootstrap interval are the only two which are conservative. As with bounded distributions, the price that our interval pays lies in the fact that the estimated coverage probability is always equal to 1.

5 Conclusions

In this paper, we have provided a method for constructing confidence intervals for the variance which exhibit guaranteed coverage probability for any (finite) sample size, uniformly over a large class of probability distributions. This is in contrast to standard methods that provide correct coverage only asymptotically for fixed distributions ('pointwise asymptotics'), such as bootstrap intervals. In addition, standard classical methods provide correct coverage for any (finite) sample size only over a very restrictive class of probability distributions, such as the normal theory interval for the class of normal distributions.

Our method is a simple application of a more general result that allows one to construct conservative confidence intervals for an arbitrary parameter of an unknown distribution by restricting attention to distributions that lie within a Kolmogorov-Smirnov confidence band for the unknown distribution function and computing the infinimum and the supremum of the parameter as function of these distributions. Note that in general it is not clear how to carry out this computation, since it involves maximizations over an infinite-dimensional set. However, in the special case of the variance, the computation can be done explicitly and we provided an algorithm to this end. When the underlying distribution has infinte support, the resulting interval will necessarily be a one-sided lower interval. However, when the underlying distribution has bounded support, the interval turns out be two-sided. In the latter case, the length of the interval is of order $n^{-1/2}$ (in probability), so that the interval is nontrivial.

We examined the finite sample properties of our interval by a simulation study that also included the normal theory interval and two bootstrap intervals. It was seen that our interval is the only one which is conservative, that is, which always meets or exceeds the nominal coverage probability. As to be expected, this achievement comes at the price of the interval being quite wide in general.

References

Anderson, T. (1967). Confidence limits for the expected value of an arbitrary bounded random variable with a continuous distribution function. *Bull. ISI* **43**, 249–251.

Bahadur, R. and Savage, L. (1956). The nonexistence of certain statistical procedures in nonparametric problems. *Annals of Mathematical Statistics* **25**, 1115–1122.

Csáki, E. (1984). Empirical distribution function. In *Handbook of Statistics, Volume 4*, 405–430. Edited by Krishnaiah, P.R. and Sen, P.K., Elsevier Science Publishers B.V., Amsterdam.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York.

Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari* **4**, 83–91.

Romano, J.P., and Wolf, M. (2000). Finite sample nonparametric inference and large sample efficiency. *Annals of Statistics*, **28**, 756–778.

Shao, J. and Tu, D. (1995). *The Jackknife and the Bootstrap.* Springer, New York.

Smirnov, N.V. (1939). Sur les ècarts de la courbe de distribution empirique. *Mat. Sb.* **6**, 3–26 (Russian; French summary).

Table 1: Estimated coverage probabilities of various two-sided confidence intervals with nominal levels 90% and 95%.

Uniform distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	0.99	0.77	0.80
30	0.90	1.00	0.99	0.85	0.87
60	0.90	1.00	0.99	0.88	0.89
10	0.95	1.00	1.00	0.83	0.84
30	0.95	1.00	1.00	0.91	0.93
60	0.95	1.00	1.00	0.93	0.94
Triangle distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	0.95	0.74	0.76
30	0.90	1.00	0.95	0.85	0.84
60	0.90	1.00	0.95	0.87	0.87
10	0.95	1.00	0.98	0.78	0.80
30	0.95	1.00	0.98	0.89	0.89
60	0.95	1.00	0.98	0.93	0.92
Two-point distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	0.00	0.38	0.32
30	0.90	1.00	0.60	0.76	0.71
60	0.90	1.00	0.95	0.79	0.24
10	0.95	1.00	0.31	0.40	0.31
30	0.95	1.00	0.60	0.76	0.75
60	0.95	1.00	0.95	0.79	0.62

Table 2: Estimated mean lengths of various two-sided confidence intervals with nominal levels 90% and 95%. TRIVIAL corresponds to the interval [0, 0.25].

Uniform distribution						
n	Level	TRIVIAL	CONS	NORM	BOOT_P	BOOT_H
10	0.90	0.25	0.23	0.17	0.07	0.07
30	0.90	0.25	0.17	0.08	0.04	0.04
60	0.90	0.25	0.13	0.05	0.03	0.03
10	0.95	0.25	0.25	0.19	0.09	0.09
30	0.95	0.25	0.19	0.10	0.05	0.05
60	0.95	0.25	0.14	0.06	0.04	0.04
Triangle distribution						
n	Level	TRIVIAL	CONS	NORM	BOOT_P	BOOT_H
10	0.90	0.25	0.21	0.09	0.04	0.04
30	0.90	0.25	0.14	0.04	0.03	0.03
60	0.90	0.25	0.11	0.03	0.02	0.02
10	0.95	0.25	0.25	0.11	0.05	0.05
30	0.95	0.25	0.16	0.05	0.03	0.03
60	0.95	0.25	0.12	0.03	0.02	0.02
Two-point distribution						
n	Level	TRIVIAL	CONS	NORM	BOOT_P	BOOT_H
10	0.90	0.25	0.25	0.07	0.08	0.07
30	0.90	0.25	0.25	0.05	0.09	0.08
60	0.90	0.25	0.25	0.02	0.06	0.06
10	0.95	0.25	0.25	0.08	0.09	0.07
30	0.95	0.25	0.25	0.05	0.10	0.08
60	0.95	0.25	0.25	0.03	0.07	0.07

Table 3: Estimated coverage probabilities of various one-sided confidence intervals with nominal levels 90% and 95%.

Normal distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	0.89	0.98	0.92
30	0.90	1.00	0.90	0.97	0.94
60	0.90	1.00	0.90	0.97	0.95
10	0.95	1.00	0.95	1.00	0.98
30	0.95	1.00	0.94	0.98	0.97
60	0.95	1.00	0.95	0.98	0.97
Exponential distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	0.83	0.99	0.93
30	0.90	1.00	0.77	0.97	0.94
60	0.90	1.00	0.80	0.99	0.97
10	0.95	1.00	0.88	0.99	0.94
30	0.95	1.00	0.85	1.00	1.00
60	0.95	1.00	0.82	0.99	0.99
Three point distribution					
n	Level	CONS	NORM	BOOT$_P$	BOOT$_H$
10	0.90	1.00	1.00	1.00	0.77
30	0.90	1.00	1.00	1.00	0.85
60	0.90	1.00	1.00	0.99	0.88
10	0.95	1.00	1.00	1.00	0.78
30	0.95	1.00	1.00	1.00	0.90
60	0.95	1.00	1.00	1.00	0.91