Journal of Econometrics 204 (2018) 119–130

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Uniform confidence bands: Characterization and optimality☆

Joachim Freyberger *, Yoshiyasu Rai

Department of Economics, University of Wisconsin, Madison, United States

ARTICLE INFO

Article history:
Received 11 August 2017
Received in revised form 11 August 2017
Accepted 26 January 2018
Available online 5 February 2018

JEL classification:
C13

Keywords:
Uniform confidence bands
Simultaneous inference
Projections
Optimality

ABSTRACT

This paper studies optimal uniform confidence bands for functions $g(x, \beta_0)$, where β_0 is an unknown parameter vector. We provide a simple characterization of a general class of taut $1-\alpha$ uniform confidence bands, allowing for both nonlinear functions and nonparametrically estimated functions. Specifically, we show that all taut bands can be obtained from projections on confidence sets for β_0 and we characterize the class of sets which yield taut bands. Using these results, we then present a computational method for selecting an approximately optimal confidence band for a given objective function. We illustrate the applicability of these results in numerical applications.

1. Introduction

Uniform confidence bands for functions are useful to summarize statistical uncertainty in both parametric and nonparametric models. They allow the reader to easily assess statistical accuracy and perform various hypothesis tests about the function without access to the data. While there are many different $1-\alpha$ confidence bands for the same function, so far there is little guidance in the literature on which one to choose in practice.

A uniform confidence band for a function $g(x, \beta_0)$, where β_0 is an unknown parameter vector, consists of upper and lower bound functions $\hat{g}_u(x)$ and $\hat{g}_l(x)$, such that $g(x, \beta_0)$ is contained in $[\hat{g}_l(x), \hat{g}_u(x)]$ for all x with probability $1-\alpha$. Many different $1-\alpha$ confidence bands could be reported in a given application. The choice is important because not all $1-\alpha$ confidence bands are *taut* in the sense that it might be possible to weakly decrease the width of the interval for all x and to strictly decrease it for some x while keeping the same coverage probability (see Section 2 for a formal definition). Moreover, even two taut confidence bands for the same function can have very different shapes and properties. In this paper, we provide a simple characterization of a general class of taut $1-\alpha$ confidence bands, allowing for both nonlinear functions and nonparametrically estimated functions. Specifically, we show that, under certain restrictions, all taut bands can be obtained from projections on confidence sets for β_0 and we characterize the class of confidence sets which yield taut bands. We provide a second characterization of taut bands in terms of inversions of suprema of weighted t-statistics.

Using our simple and constructive characterization of taut uniform confidence bands, we then present a computational method for selecting approximately optimal bands for different objective functions. Our leading example is the band which minimizes a weighted area. For this example we provide low level conditions for the selected band to be approximately optimal and asymptotically valid. The general results in the paper also apply to a variety of other objective functions, such as minimizing average marginal coverage probabilities.

As a starting point we consider confidence bands for functions of the form $g(x, \beta_0) = p(x)'\beta_0$. We also assume that we have an estimator $\hat{\beta}$ of β_0, where $\hat{\beta} \sim N(\beta_0, \Sigma)$. Due to the normality assumption, the first set of results are exact finite sample results. We then discuss extensions of these results to asymptotic approximations using $\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{d} N(0, \Sigma)$, nonlinear functions satisfying $g(x, \beta) \approx g(x, \beta_0) + \nabla_\beta g(x, \beta_0)'(\beta - \beta_0)$ in a neighborhood of β_0, and nonparametric estimators using a finite dimensional approximation $g(x) \approx p_K(x)'\beta_K$.

We illustrate the wide applicability of these results in two numerical applications. First, we consider a regression model with heteroskedasticity and simulated data. Second, we use data from Berry et al. (1995) and construct confidence bands for price

☆ We thank a co-editor, an associate editor, and a referee for valuable comments which have helped to improve the paper. We also thank Jean-Marie Dufour, Amit Gandhi, Bruce Hansen, Matt Masten, Pepe Montiel Olea, Jack Porter, Andres Santos, Xiaoxia Shi as well as seminar and conference participants at UW Madison, the 2016 CEME conference at Duke, and the 2017 ASSA meetings for helpful comments and discussions.

* Corresponding author.
E-mail addresses: jfreyberger@ssc.wisc.edu (J. Freyberger), yrai@wisc.edu (Y. Rai).

https://doi.org/10.1016/j.jeconom.2018.01.006
0304-4076/© 2018 Elsevier B.V. All rights reserved.

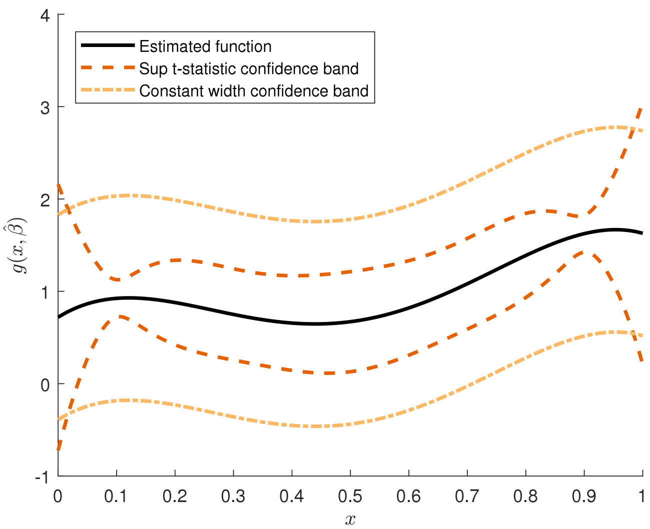

Fig. 1. Illustrative example.

elasticities implied by the estimated parameters of a structural model of demand.

Illustrative example: The following example illustrates that the choice of the uniform confidence band can be important. In this example $g(x, \beta) = \beta_1 + x\beta_2 + x^2\beta_3 + x^3\beta_4 + x^4\beta_5$ and $g(x, \beta_0) = E[Y \mid X = x]$. Section 4.1 explains the DGP, the estimator, and the confidence bands in detail. Fig. 1 shows the estimated function as well as two different 90% confidence bands. The dashed band is based on inversion of a standard sup t-statistic and the dotted–dashed band has a constant width for all x. Both of these bands are taut, have the same coverage probability, and are of the form $g(x, \hat{\beta}) \pm c(x)$. However, depending on how much importance a researcher places on different values of x, one might have clear preferences for one over the other. Moreover, hypothesis tests can lead to different outcomes depending on the band reported. For example, the null hypothesis that $g(x, \beta_0)$ is constant can be rejected with the sup t-statistic band, but not with the constant width band.

Related literature: The literature on uniform confidence bands goes back to Working and Hotelling (1929) who introduced hyperbolic bands in a simple linear regression model with normal errors. They showed that such a band can be obtained from a projection on an ellipse shaped confidence region, although this construction usually leads to conservative bands. These bands are often referred to as Scheffé bands due to his seminal work on multiple hypothesis testing (Scheffé, 1953). The width of the band based on the sup t-statistic is suitably smaller than but proportional to that of the Scheffé band and is thus also sometimes referred to as the Scheffé band. A variety of other bands, such as two or three segment bands or constant width bands, have been proposed in the literature (see Liu (2010) for an excellent overview). The first definition of taut bands we are aware of has been provided by Wynn and Bloomfield (1971), in a less general framework, who also showed in a linear regression with homoskedastic errors that all taut bands can be obtained by a projection (see also Khorasani and Milliken (1979) and Naiman (1984a) for characterization results in linear models). Our characterization of projection bands is more constructive which allows us, among others, to select (approximately) optimal bands using this result. Moreover, we start with more primitive assumptions, relate projection bands to bands obtained by t-statistic inversion, and extend the characterization to more general settings including nonlinear models. Gsteiger et al. (2011) discussed a confidence band for nonlinear regression models. Our results provide a formal justification for this band as well as various other ones. More recently, Belloni et al. (2015) constructed hyperbolic confidence bands for nonparametric regression functions using a series estimator and relying on high dimensional normal approximations. Other recent work on uniform confidence bands in nonparametric models includes Horowitz and Lee (2012, 2017), Chen and Christensen (2015), and Tao (2016). Projection based confidence regions have also been used in various other settings, such as Dufour (1990), Lütkepohl et al. (2015), Gafarov et al. (2016), and Kaido et al. (2016), but these papers do not consider our characterization and optimality results.

Related work on optimality properties of uniform confidence bands includes Bohrer (1973) who proved that Scheffé bands have the smallest average width with respect to the Lebesgue measure when the support is an ellipse (hence there cannot be an intercept in a linear model). Other papers extended this result to show that for certain confidence bands in a regression framework, there exist weight functions such that the bands minimize a weighted area (see Naiman (1984a) and Piegorsch (1985)). Our results imply the reverse, namely that one can find the optimal band for a given weight function. Naiman (1984b) considered optimality of bands which satisfy a bound on an expected coverage measure instead of having a certain coverage probability. Naiman (1987) characterized certain minimax regret bands. More recent papers such as Liu and Hayter (2007), Liu et al. (2009), and Liu and Ah-Kine (2010) are concerned with confidence bands which have optimality properties in terms of the implied confidence set for the parameter. Recently, Montiel Olea and Plagborg-Møller (2017) showed that a band obtained from inverting a sup t-statistic minimizes worst-case regret among a class of confidence bands. Relatedly, as we discuss further in Section 5, this band is also optimal in a minimax sense, but other bands are optimal with different criterion functions. To the best of our knowledge, we are the first to provide a constructive method to obtain optimal uniform confidence bands in a general class of models and general criterion functions.

Finally, note that many of these characterization and optimality results are obtained in a regression setting with normal and homoskedastic errors and rely on certain algebraic features of that model. Thus, not all of these results carry over to a general setting. For example, in a simple linear regression model, the width of the hyperbolic band is minimized at the mean value of the regressor, upper bound functions of taut bands are convex, and the constant width band is taut. These features do not hold more generally and hence optimality considerations can be more important in other settings (see e.g. Corollary 2).

2. Finite sample results

In this section we consider uniform confidence bands for a function $g(x, \beta_0) = p(x)'\beta_0$, where $p(x) \in \mathbb{R}^K$ is a vector of transformations of a vector $x \in \mathbb{R}^{d_x}$, and we have an estimator $\hat{\beta} \sim N(\beta_0, \Sigma)$. Bands are defined over a set $\mathcal{X} \subseteq \mathbb{R}^{d_x}$. The next section extends these results to asymptotic approximations, nonlinear functions, and nonparametric estimators.

2.1. Definitions and assumptions

In our setting the only information in the data about $g(x, \beta_0)$ is the estimator $\hat{\beta} \sim N(\beta_0, \Sigma)$. Thus, we restrict ourselves to confidence bands of the form $[g_l(x, \hat{\beta}), g_u(x, \hat{\beta})]$. Furthermore, we impose a regularity condition on the bands and only consider bands in the class $\mathcal{C}$ described in the following definition, where $\alpha \in (0, 1)$.