Statistics & Probability Letters
Volume 111, April 2016, Pages 67-71

A general parametric Stein characterization

Christophe Ley [a], Yvik Swan [b]

https://doi.org/10.1016/j.spl.2016.01.008

Abstract

We present a general characterization theorem for parametric probability distributions in terms of a differential operator akin to the so-called Stein operators from the literature on Stein's method.

Introduction

Stein's method is a collection of tools allowing to compute sharp rates of convergence in stochastic approximation problems. In short, the method consists of transforming the problem of bounding a suitable distance between two probability measures (specifically distances which can be represented as integral probability metrics) into that of bounding a suitable linear operator over a well-chosen class of functions. The foundations of the method were laid down in Charles Stein's classic paper (Stein, 1972) and book (Stein, 1986). For a more modern overview of the method we refer the reader to the monograph (Chen etal., 2011) or the review article (Ross, 2011).

Contributions to the literature on Stein's method are generally of three types:

- Applications of the method to new settings (e.g.in network analysis Franceschetti and Meester, 2006, random matrices Mackey etal., 2014 or even quantum physics McKeague and Levin, inpress).

- Development of new versions of the method to tackle new specific targets or dependency structures (e.g. Pearson statistics and χ^2 approximation Gaunt etal., 2015, spin glasses and distributions from statistical physics Chatterjee and Shao, 2011, Pólya urns and beta distribution Goldstein and Reinert, 2013).

- Extensions of various aspects of the method to more abstract and general settings (e.g. for Pearson distributions Afendras etal., 2011, general densities Ley and Swan, 2013, invariant measures of diffusions Kusuoka and Tudor, 2013).

The present paper falls within the last category.

Let P be some probability distribution with respect to which one aims to perform Stein's method (that is, P is the target distribution). A Stein characterization for P is a result of the form

$X \sim P$ if and only if $E\left[\mathscr{A}_P f(X)\right] = 0$ for all $f \in \mathscr{F}(\mathscr{A}_P)$ with $\mathscr{A}_P$ a suitable (usually linear) operator (called Stein operator for P) and $\mathscr{F}(\mathscr{A}_P)$ a well chosen class of functions (called Stein class). There exist many different ways of obtaining such operators, see Ley et al. (2014) for a general approach.

In Ley and Swan (in press) we propose a tool for deriving identities such as (1), this time with the added advantage of also providing a statistical interpretation to the resulting quantities. The main result of that paper, however, is only valid in a one-parameter setting under the strong assumption that the target has support which does not depend on the parameter of interest. As will be shown in Section 2, such restrictions are not necessary and the difficulties entailed by a general setting can, if handled with care, be dealt with in a uniform and elegant manner. The main result of this paper, Theorem 2.1, therefore contributes a new tool to the literature on Stein's method, in the form of a most general parametric Stein characterization.

Section snippets

Notations and definitions

Throughout, we let $k, p \in \mathbb{N}_0$ and consider two measure spaces $(\mathscr{X}, \mathscr{B}_{\mathscr{X}}, m_{\mathscr{X}})$ and $(\Theta, \mathscr{B}_\Theta, m_\Theta)$, where $\mathscr{X}$ is either $\mathbb{R}^k$ or $\mathbb{Z}^k$, where Θ is a subset of $\mathbb{R}^p$ whose interior is non-empty, where $m_{\mathscr{X}}$ is either the Lebesgue measure or the counting measure, depending on the nature of $\mathscr{X}$, where m_Θ is the Lebesgue measure, and where $\mathscr{B}_{\mathscr{X}}$ and $\mathscr{B}_\Theta$ are the corresponding σ-algebras. In this setup we disregard the case of discrete parameter spaces (as in, e.g., the discrete uniform).

Consider a couple $(\mathscr{X}, \Theta)$ equipped with the…

Proof of equality (4)

First note that $\partial_t \left(f_A(x;t)\, g(x;t)\right)\big|_{t=\theta} = \partial_t \left(\int_{\theta_0}^t l_A(x;u,t)\, g(x;u)\, dm_\Theta(u)\right)\Big|_{t=\theta}$ Now we have

$$= l_A(x;\theta,\theta)\, g(x;\theta) + \int_{\theta_0}^\theta \partial_t \left(l_A(x;u,t)\right)\big|_{t=\theta} g(x;u)\, dm_\Theta(u).$$

$\partial_t \left(l_A(x;u,t)\right)\big|_{t=\theta} = \partial_t \left(\mathbb{I}_{S_t}(x)\right)\big|_{t=\theta} \left(\mathbb{I}_A(x) - \mathrm{P}\left(Z_u \in A \mid Z_u \in S_\theta\right)\right)$ On the one hand, setting

$- \partial_t \left(\mathrm{P}\left(Z_u \in A \mid Z_u \in S_t\right)\right)\big|_{t=\theta} \mathbb{I}_{S_\theta}(x).$

$H_1(x;\theta) := \partial_t \left(\mathbb{I}_{S_t}(x)\right)\big|_{t=\theta}$ we immediately get

$\int_{\theta_0}^\theta \left(\mathbb{I}_A(x) - \mathrm{P}\left(Z_u \in A \mid Z_u \in S_\theta\right)\right) g(x;u)\, dm_\Theta(u)$

$\left|H_1(x;\theta)\right| \leq 2h(x) \left|\partial_t \left(\mathbb{I}_{S_t}(x)\right)\big|_{t=\theta}\right| m_\Theta\left([\theta_0, \theta]\right)$ and is thus bounded uniformly in θ over Θ_0 by a $m_{\mathscr{X}}$-integrable function and satisfies $H_1(x;\theta_0) = 0$. On the other hand, we have (thanks to…

Acknowledgments

Yvik Swan gratefully acknowledges support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy).

We thank the referee for his/her careful reading of our manuscript and his/her accurate remarks which led us to clarify certain important parts of the paper.…

Recommended articles

References (15)

G. Afendras *et al.*

An extended Stein-type covariance identity for the Pearson family with applications to lower variance bounds

Bernoulli (2011)

S. Chatterjee *et al.*

Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model

Ann. Appl. Probab. (2011)

L.H.Y. Chen *et al.*

M. Franceschetti *et al.*

Critical node lifetimes in random networks via the Chen–Stein method

IEEE/ACM Trans. Netw. (TON) (2006)

Gaunt, R.E., Pickett, A., Reinert, G., 2015. Chi-square approximation by Stein's method with application to Pearson's...

L. Goldstein *et al.*

Stein's method for the Beta distribution and the Pólya–Eggenberger urn

J. Appl. Probab. (2013)

Kusuoka, S., Tudor, C.A., 2013. Extension of the fourth moment theorem to invariant measures of diffusions. Preprint...

There are more references available in the full text version of this article.

Cited by (4)

When the score function is the identity function - A tale of characterizations of the normal distribution

2023, Econometrics and Statistics

> *Citation Excerpt :*
>
> …Cacoullos (1982) General parametric Stein characterizations, based on parameters of interest of other natures than location and scale, have been studied in Ley and Swan (2016b) and Ley and Swan (2016a). The latter paper also develops a link between typical operators from the literature, such as those from Chatterjee et al. (2011), and the operators obtained by adopting the (till then not considered) parametric viewpoint.…

Show abstract

Gauss and the identity function - a tale of characterizations of the normal distribution ↗
2020, arXiv

Stein operators, kernels and discrepancies for multivariate continuous distributions ↗
2018, arXiv

Stein's method for comparison of univariate distributions ↗
2017, Probability Surveys

View full text