# Statistical inference via Stein's method

Eduardo García-Portugués • Department of Statistics, UC3M
Bruno Ebner • Institute of Stochastics, KIT

Stein's method is a probabilistic tool for assessing the similarity between two probability distributions, typically the distribution of interest and a reference distribution. Stein's method works by establishing a characterization of two probability distributions $P$ and $Q$ on a given space:

$$P = Q \iff \mathbb{E}_{X\sim Q}[Tg(X)] = 0 \quad \text{for a set of functions } g. \tag{1}$$

The function-to-function operator $T$ is the Stein's operator, whose obtainment is key, and the expectation is taken for a random vector $X \sim Q$. There exist many worked out examples for Q being Gaussian, exponential, etc. As an example of an application, thanks to characterization (1) one could test for Gaussianity or exponentiality in the data: if the expectation is different from zero, then $P \neq Q$. Beyond this simple example, the list of applications of Stein's method in statistical inference is long and has gained a lot of traction in the last years. The array of modern applications of Stein's method is reviewed in the extensive survey by Anastasiou et al. (2023). There are several online seminars summarizing the applications of Stein's method (e.g., https://t.ly/wMC-i and https://t.ly/qhER6).

A particularly useful application of Stein's method is the construction of estimation methods that are alternative to maximum likelihood, such as Stein's method of moments (Ebner et al., 2024). These alternative estimation approaches can be especially relevant for challenging high-dimensional distributions without closed-form normalizing constant, in which maximum likelihood is challenging. However, benefits are also tangible in low-dimensional cases. For example, the maximum likelihood estimators of the two parameters of the gamma distribution are obtained by solving a system of nonlinear equations, which can be challenging in practice. Stein's method of moments allows obtaining closed-form estimators that are asymptotically as good as the maximum likelihood estimators (Ye & Chen, 2017).

The objective of this thesis proposal is to: (1) dissect the probabilistic fundamentals of Stein's method and explain its range of applicability; (2) illustrate the power of Stein's method through the available worked out examples in statistical inference using numerical experiments; (3) review the recent trends in goodness-of-fit testing and/or estimation based on Stein's method; (4) derive a new Stein-based and/or estimator in an unexplored setup in non-Euclidean statistics and apply it to a dataset provided by the director.

## References

Anastasiou, A., Barp, A., Briol, F.-X., Ebner, B., Gaunt, R. E., Ghaderinezhad, F., Gorham, J., Gretton, A., Ley, C., Liu, Q., Mackey, L., Oates, C. J., Reinert, G., & Swan, Y. (2023). Stein's method meets computational statistics: A review of some recent developments. *Statistical Science*, *38*(1), 120–139. https://doi.org/10.1214/22-STS863

Ebner, B., Fischer, A., Gaunt, R. E., Picker, B., & Swan, Y. (2024). Stein's method of moments. *arXiv:2305.19031v4*. https://doi.org/10.48550/arXiv.2305.19031v4

Ye, Z.-S., & Chen, N. (2017). Closed-form estimators for the gamma distribution derived from likelihood equations. *The American Statistician*, *71*(2), 177–181. https://doi.org/10.1080/00031305.2016.1209129