Fasano–Franceschini Test: an Implementation of a 2-Dimensional Kolmogorov–Smirnov test in R

by Elan Ness-Cohn, Rosemary Braun

Abstract The univariate Kolmogorov-Smirnov (KS) test is a non–parametric statistical test designed to assess whether two samples come from the same underlying distribution. The versatility of the KS test has made it a cornerstone of statistical analysis across the scientific disciplines. However, the test proposed by Kolmogorov and Smirnov does not naturally extend to multidimensional distributions. Here, we present the **fasano.franceschini.test** package, an R implementation of the 2-D KS two–sample test (Fasano and Franceschini, 1987) and provide multiple use cases across the scientific disciplines. The **fasano.franceschini.test** package provides three improvements over the current 2-D KS test on the Comprehensive R Archive Network (CRAN): (i) the Fasano and Franceschini test has been shown to run in $O(n^2)$ versus the Peacock implementation which runs in $O(n^3)$; (ii) the package implements a procedure for handling ties; and (iii) the package implements a parallelized permutation procedure for improved significance testing. Ultimately, the **fasano.franceschini.test** package presents a robust statistical test for analyzing random samples defined in 2-dimensions.

Introduction

The Kolmogorov–Smirnov (KS) is a non–parametric, univariate statistical test designed to assess whether a set of data is consistent with a given probability distribution (or, in the two-sample case, whether the two samples come from the same underlying distribution). First derived by Kolmogorov and Smirnov in a series of papers (Kolmogorov, 1933a,b; Smirnov, 1936, 1937, 1939, 1944, 1948), the one-sample KS test defines the distribution of the quantity D_{KS}, the maximal absolute difference between the empirical cumulative distribution function (CDF) of a set of values and a reference probability distribution. Kolmogorov and Smirnov's key insight was proving the distribution of D_{KS} was independent of the CDFs being tested. Thus, the test can effectively be used to compare any univariate empirical data distribution to any continuous univariate reference distribution. The two-sample KS test could further be used to compare any two univariate empirical data distributions against each other to determine if they are drawn from the same underlying univariate distribution.

The nonparametric versatility of the univariate KS test has made it a cornerstone of statistical analysis and is commonly used across the scientific disciplines (Atasoy et al., 2017; Chiang et al., 2018; Hahne et al., 2018; Hargreaves et al., 2020; Wong and Collins, 2020; Kaczanowska et al., 2021). However, the KS test as proposed by Kolmogorov and Smirnov does not naturally extend to distributions in more than one dimension. Fortunately, a solution to the dimensionality issue was articulated by Peacock (Peacock, 1983) and later extended by Fasano and Franceschini (Fasano and Franceschini, 1987).

Currently, only the Peacock implementation of the 2-D two-sample KS test is available in R (R Core Team, 2016) with the **Peacock.test** package via the `peacock2` function, but this has been shown to be markedly slower than the Fasano and Franceschini algorithm (Lopes et al., 2007). A C implementation of the Fasano–Franceschini test is available in Press et al. (2007); however, arguments have been made to the validity of the implementation of the test not being distribution-free (Babu and Feigelson, 2006). Furthermore, in the C implementation, statistical testing is based on a fit to Monte Carlo simulation that is only valid for significance levels $\alpha \lessapprox 0.20$.

Here we present the **fasano.franceschini.test** package as an R implementation of the 2-D two-sample KS test described by Fasano and Franceschini (Fasano and Franceschini, 1987). The **fasano.franceschini.test** package provides two improvements over the current 2-D KS test available on the Comprehensive R Archive Network (CRAN): (i) the Fasano and Franceschini test has been shown to run in $O(n^2)$ versus the Peacock implementation which runs in $O(n^3)$; and (ii) the package implements a permutation procedure for improved significance testing and mitigates the limitations of the test brought noted by Babu and Feigelson (2006).

Models and software

1-D Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov (KS) test is a non–parametric method for determining whether a sample is consistent with a given probability distribution (Stephens, 1992). In one dimension, the Kolmogorov-Smirnov statistic (D_{KS}) is the defined by the maximum absolute difference between the cumulative density functions of the data and model (one–sample), or between the two data sets (two–sample), as illustrated in **Figure 1**.

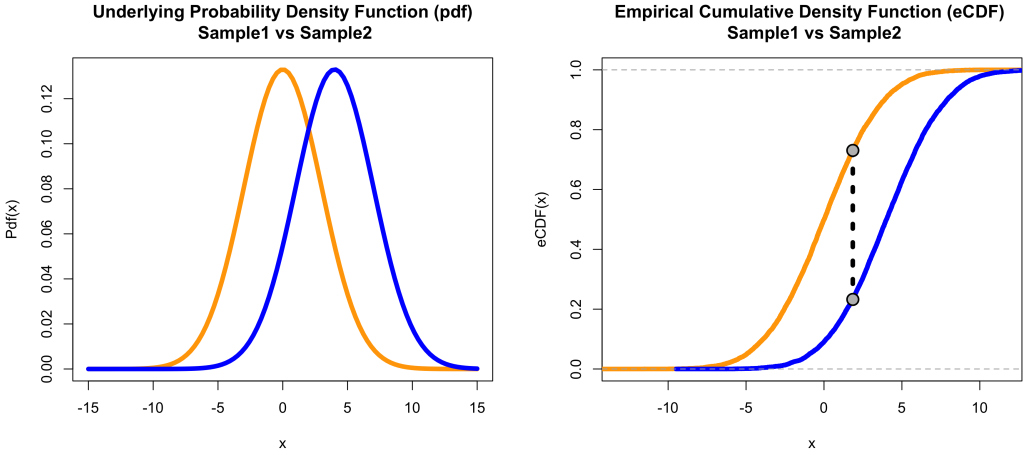

Figure 1: LEFT: Probability density function (PDF) of two normal distributions: orange sample 1, $\mathcal{N}(\mu = 0, \sigma^2 = 1)$; blue sample 2, $\mathcal{N}(\mu = 5, \sigma^2 = 1)$. **RIGHT:** Cumulative density functions (CDF) of the two PDFs; the black dotted line represents the maximal absolute difference between the CDFs (D_{KS}).

In the large–sample limit ($n \geq 80$), it can be shown (Kendall and Stuart, 1946) that D_{KS} converges in distribution to

$$D_{KS} \xrightarrow{d} \Phi(\lambda) = 2 \sum_{k=1}^{\infty} -1^{k-1} e^{-2k^2\lambda^2} . \tag{1}$$

In the one-sample case with a sample of size n, the p value is given by

$$\mathbb{P}(D > observed) = \Phi(D\sqrt{n}) ; \tag{2}$$

in the two-sample case, the p value is given by

$$\mathbb{P}(D > observed) = \Phi\left(D\sqrt{\frac{n_1 n_2}{n_1 + n_2}}\right) . \tag{3}$$

where n_1 and n_2 are the number of observations in the first and second samples respectively.

Higher dimensional variations: Peacock Test (1983) and Fasano–Franceschini Test (1987)

Extending the above to two or higher dimension is complicated by the fact that CDFs are not well-defined in more than one dimension. In 2-D, there are 4 ways (3 independent) of defining the cumulative distribution, since the direction in which we order the x and y points is arbitrary (**Figure 2**); more generally, in k-dimensional space there are $2^k - 1$ independent ways of defining the cumulative distribution function (Peacock, 1983).

Peacock (1983) solved the higher dimensionality issue by defining the 2-D test statistic as the largest difference between the empirical and theoretical cumulative distributions, after taking all possible ordering combinations into account. Peacock's test thus computes the total probability—i.e. fraction of data—in each of the four quadrants around all possible tuples in the data. For example, for n points in a two-dimensional space, the empirical cumulative distribution functions is calculated in the $4n^2$ quadrants of the plane defined by all pairs $(X_i, Y_j) : i, j \in [1, n]$, where X_i and Y_j are any observed value of x and y (whether or not they are observed as a pair). There are n^2 such pairs, each of which can define four quadrants in the 2-D plane; by ranging over all possible pairs of data points

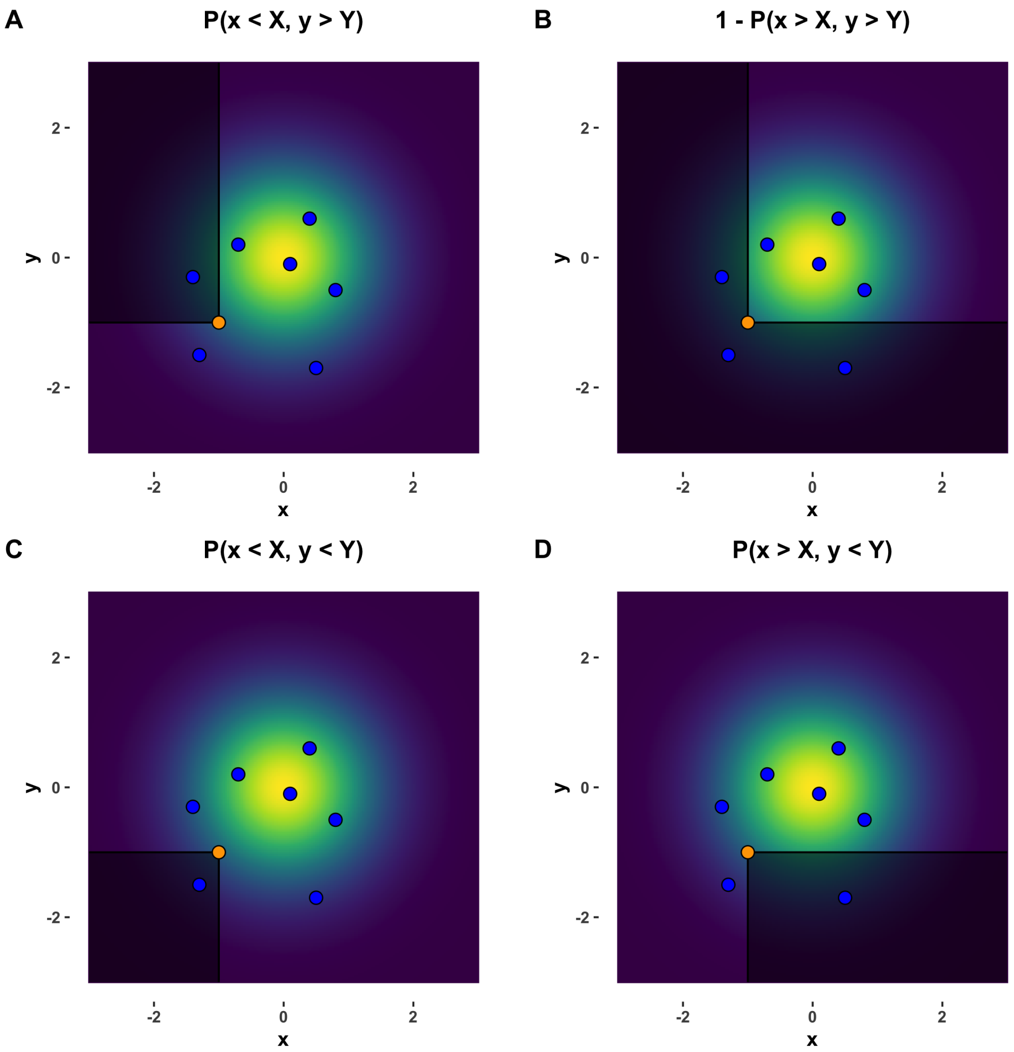

Figure 2: Four ways (3 independent) of defining the cumulative distribution for a given point in 2-D. Here, the orange point (X, Y) is chosen as the origin; the density of observations may be integrated as $\mathbb{P}(x < X, y > Y)$ (A); $\mathbb{P}(x < X \cup y < Y)$ (B); $\mathbb{P}(x < X, y < Y)$ (C); $\mathbb{P}(x > X, y < Y)$ (D).

and quadrants, the 2-dimensional D statistic is defined by the maximal difference of the integrated probabilities between samples.

The variation defined by Fasano and Franceschini (1987) was to only consider quadrants centered on each observed (x, y) pair to compute the cumulative distribution functions. That is, rather than looking over all n^2 points $(X_i, Y_j) : i, j \in [1, n]$, Fasano and Franceschini only use the observed n points $(X_i, Y_i) : i \in [1, n]$. Thus for any given n points in a two-dimensional space, those n points define $4n$ (rather than $4n^2$) quadrants. The procedure is illustrated in **Figure** 3. The algorithm loops through each point in one sample in turn to define the origin of 4 quadrants (grey dotted lines in **Figure** 3). The fraction of points in each sample is computed in each quadrant, and the quadrant with the maximal difference is designated with the current maximum for the specified origin. By iterating over all data points and quadrants, the test statistic $D_{FF,1}$ is defined by the maximal difference of the integrated probabilities between samples in any quadrant for any origin from the first sample. In **Figure** 3, using the orange point as the origin, the maximal difference is $D_{FF,1} = 0.52$.

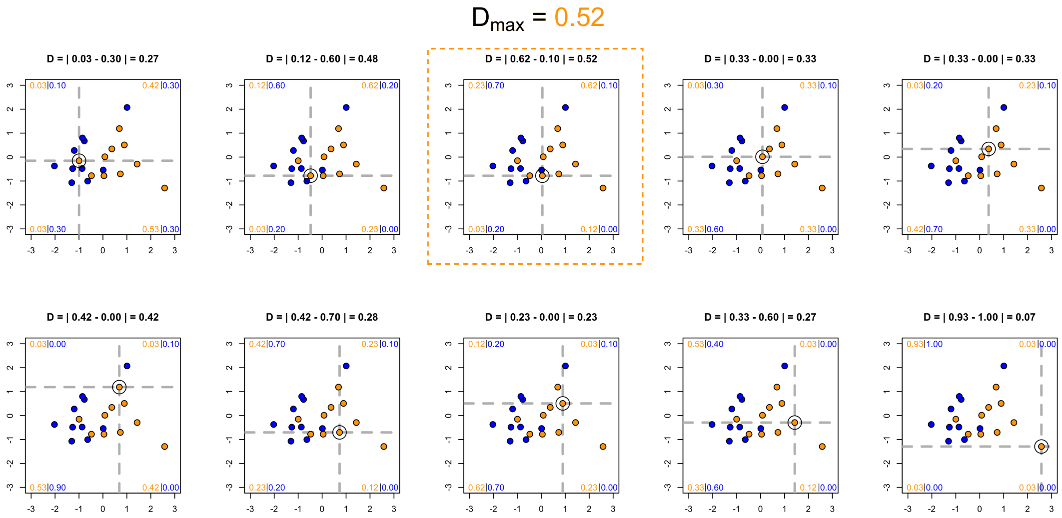

Figure 3: Illustration of the Fasano–Franceschini algorithmic search for the maximal difference ($D_{FF,1}$) between sample 2-D eCDFs. Looping through each point in the sampled data to define a unique origin (grey dotted line), the fraction of orange and blue points in each quadrants are computed (plot corners). For each origin, the quadrant which maximizes the absolute difference in the integrated probabilities is indicated. The origin which maximizes the overall absolute difference in the integrated probabilities between samples is highlighted by the orange box.

This process is repeated using the points from *other* sample as the origins to compute the maximal $D_{FF,2}$ with origins from the second sample. $D_{FF,1}$ and $D_{FF,2}$ are then averaged to compute the overall D_{FF} for hypothesis testing, $D_{FF} = (D_{FF,1} + D_{FF,2})/2$.

It may be that some points are tied with the X and/or Y coordinates of the origin, creating an ambiguity when computing the fraction of points in each quadrant. Since the test attempts to define the maximal difference of the cumulative probabilities, a natural solution would be to treat a point that is tied with the current X and/or Y coordinates of the origin as equally likely to have been drawn from any of the tied quadrants. Hence, any data point sharing the same X or Y coordinate as the origin is evenly distributed across the tied quadrants, with each of the two quadrants receiving half a count. Any data point sharing the both the same X and Y coordinates as the current origin (including the origin itself) is evenly distributed across all quadrants, with all four quadrants receiving a quarter count.

Null distribution of D_{FF}

Using Monte Carlo simulation, Fasano and Franceschini created a look-up table of critical values of D_{FF} as a function of D_{FF}, the sample size, and the coefficient of correlation r. Press et al. (2007) later defined an approximate fit to the lookup table as follows. For a single sample of size n,

$$\mathbb{P}(d_{FF} > D_{FF}) = \Phi\left(\frac{D_{FF}\sqrt{n}}{1 + \sqrt{1 - r^2}(0.25 - 0.75/\sqrt{n})}\right). \quad (4)$$

where $\Phi(\cdot)$ is as defined in Eq 1. The two sample case uses the same formula as above, but with the slight variation where

$$n = \frac{n_1 n_2}{n_1 + n_2}. \tag{5}$$

In both cases, r is defined in the usual way as

$$r = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2}\sqrt{\sum_i (Y_i - \bar{Y})^2}}. \tag{6}$$

Power of the Peacock Test (1983) and Fasano–Franceschini Test (1987)

A complete treatmentent of the power of both the Peacock and Fasano–Franceschini tests can be found in the primary literature (Peacock, 1983; Fasano and Franceschini, 1987) and subsequent benchmarking paper (Lopes et al., 2007). In short, results for uncorrelated distributions demonstrated that there is no difference between the power of the two tests. When the correlation coefficient of the model distributions approach unity, the power of the Fasano-Francheschini test is slightly higher (Fasano and Franceschini, 1987). These findings were corroborated by Lopes et al. (2007), who benchmarked the Peacock and Fasano-Francheschini tests. Benchmarking data was comprised of samples containing varying number of points drawn from the same or different underlying distributions. Both methods demonstrated comparable acceptance and rejection metrics, as well high stability (low standard error in the significance calculation) across multiple runs (Lopes et al., 2007).

Illustrations

Fasano–Franceschini test usage

In their paper, Fasano and Franceschini use Monte Carlo simulation to approximate the distribution of D_{FF} as a function of the sample size n and the coefficient of correlation r. Notably, unlike the 1-D KS test, the distribution of D_{FF} is *not* completely independent of the shape of the 2-D distribution of the underlying data, but depends on the correlations between the variables. In the case where the variables X and Y are perfectly correlated ($r = 1$), the 2-D distribution lies along a single line and thus the 1-D KS test could be used; at the other extreme were X and Y are perfectly uncorrelated ($r = 0$), the 2-D distribution is independent in the X and Y directions and one could apply the 1-D KS test on the marginal distributions. Results from Monte Carlo simulation support these expectations, showing that the distribution of D is nearly identical for varying distributions with the same correlation coefficient (Fasano and Franceschini, 1987). The approximation by Press et al. (2007) (Eq 4–5) can be used to test the significance levels for the 2-D K-S test using the following code:

```
> #set seed for reproducible example
> set.seed(123)
>
> #create 2-D samples with the same underlying distributions
> sample1Data <- data.frame(
>   x = rnorm(n = 100, mean = 0, sd = 1),
>   y = rnorm(n = 100, mean = 0, sd = 1)
> )
> sample2Data <- data.frame(
>   x = rnorm(n = 100, mean = 0, sd = 1),
>   y = rnorm(n = 100, mean = 0, sd = 1)
> )
> fasano.franceschini.test(S1 = sample1Data,
>                          S2 = sample2Data)

        Fasano-Francheschini Test

data:  sample1Data and sample2Data
D-stat = 0.105, p-value = 0.7924
sample estimates:
 dff,1  dff,2
0.1075 0.1025
```

Permutation version of the Fasano–Franceschini test

It has been noted that the approximation from Press et al. (2007) is only accurate when $n \gtrsim 20$ and the p-value is less than (more significant than) ~ 0.2 (Babu and Feigelson, 2006). While this still allows a simple rejection decision to be made at any $\alpha \leq 0.2$, it is sometimes useful to quantify large p more exactly (such as if one was to do a cross-study concordance analysis comparing p values between studies as in Ness-Cohn et al. (2020)), and to apply it to smaller datasets. To address these limitations, one can empirically compute the significance levels for the particular multidimensional statistic directly from the particular data set under study. As Fasano and Franceschini's paper was originally released in 1987, this approach was unfeasible at scale. Today, modern computers can rapidly compute a permuted null distribution of D_{FF} from the data to test significance.

In the permutation test implementation of **fasano.franceschini.test**, the sample labels are randomly permuted to generate two 2-dimensional data sets with new sample labels `nPermute` times. The frequency count by quadrant is performed for each permuted resampling as described above to compute the D_{FF}. The observed D_{FF} is then compared to the distribution of permuted D_{FF} to compute a p value. The permutation version of the Fasano–Franceschini test can be run as follows (see `fasano.franceschini.test` for further source code details and implementation).

```
> #set seed for reproducible example
> set.seed(123)
>
> #create 2-D samples with the same underlying distributions
> sample1Data <- data.frame(
>  x = rnorm(n = 100, mean = 0, sd = 1),
>  y = rnorm(n = 100, mean = 0, sd = 1)
> )
> sample2Data <- data.frame(
>  x = rnorm(n = 100, mean = 0, sd = 1),
>  y = rnorm(n = 100, mean = 0, sd = 1)
> )
>
> fasano.franceschini.test(S1 = sample1Data,
>                          S2 = sample2Data,
>                          nPermute = 1000,
>                          cores = 1)

           Fasano-Francheschini Test

data:  sample1Data and sample2Data
D-stat = 0.14, p-value = 0.6733
sample estimates:
 dff,1  dff,2
0.1325 0.1475
```

To improve run time, one may adjust the `cores` parameter; see the R **parallel** package and the `mclapply` function for further details. [Note that, due to limitations of **parallel**, the parallelized permutation procedure only works on *nix operating systems (including MacOS, Linux, Unix, and BSD), and not Windows. Parallelization is generally only necessary when working with large values of N and/or `nPermute`; in this case, we recommend using parallelization on a linux-based HPC cluster.]

Real world applications

Much like the 1-D KS test, the 2-D Fasano-Franceschini test is widely applicable across the scientific disciplines. **Figure 4** shows use cases in the three distinct fields of social science, ecology, and cell biology. In **Figure 4A**, the Fasano-Franceschini test was used to detect differences in the distribution of social services within the municipality of Rennes, France (Floch et al., 2018); the Fasano-Franceschini test detected a significant difference in the distribution of clothing stores in relation to doctors offices, schools, and pharmacies ($Bonf = 6.9\text{e}{-18}; Bonf = 2.7\text{e}{-10}; Bonf = 1.6\text{e}{-7}$, respectively). No statistical difference was seen between the distribution of the latter three social services in comparison to each other (doctors offices vs. pharmacies, $Bonf = 1$; schools vs. pharmacies, $Bonf = 1$; doctors offices vs. schools, $Bonf = 1$). Pragmatically, such analysis can help to identify geographic service disparities and inform future city planning. In **Figure 4B**, the Fasano-Franceschini test was used to gain insight on the differences in the distribution of species in an ecosystem. Comparing the distribution of tree species with the Paracou Forest in French Guiana (Marcon et al., 2015), a statistically significant difference is seen between *V. Americana* and *Q. Rosea* ($p = 3.6\text{e}{-5}$). While this was a trivial analysis of two species, multiple pairwise comparisons of species in a geographic area can identify ecological variation, providing useful insights for conservation and restoration efforts. In **Figure 4C**, the Fasano-Franceschini test was used to detect difference in the localization of distinct cell types with a tissue. Immunohistochemistry was used to stain tumor-infiltrated lymph node cross section for T-cells and tumor cell (Set, 2010). A statistically significant difference is seen between the two cell types ($p < 2.2\text{e}{-16}$). In the cell biology context, the Fasano-Franceschini test lends itself for uses in high-throughput histological diagnostic assays; with initial findings support this proof of concept analysis, as deferentially distributed cell types within a tumor sample are shown to be linked to health outcomes (Set, 2010).

Beyond these three practical examples, the Fasano-Franchescini test has already been used in a myriad of other contexts, including the analysis of star clusters in the field of astrophysics (Met, 2002) and seizure progression in the field of neuroscience (Chen et al., 2012). Ultimately, the nonparametric multivariate Fasano-Francheschini test proves to be a versatile analysis strategy that transcends domains and can provide practical insights.

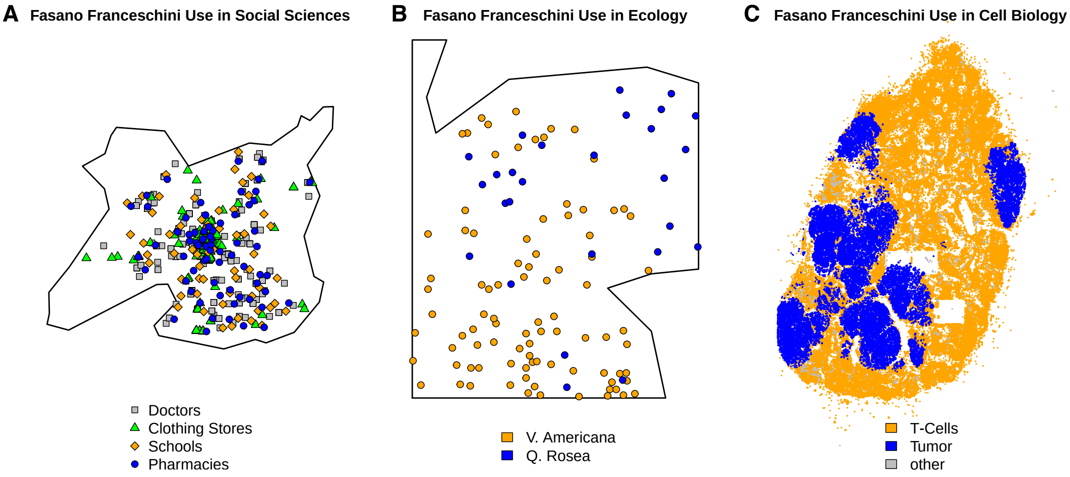

Figure 4: Example application settings for the Fasano-Francheschini test. **A:** Geographic locations of fours distinct facilities in the municipality of Rennes, France (Floch et al., 2018). **B:** Geographic location of tree species in the plot 16 field station of the Paracou Forest, French Guiana (Marcon et al., 2015). **C:** Immunohistochemistry stained tumor-infiltrated lymph node cross section, with tumor and T-cells colored in blue and orange respectively (Set, 2010).

Computational efficiency

To assess the computational efficiency, we benchmarked the package as follows. Using the **rbenchmark** package to evaluate runtime, the Fasano–Franceschini test and Peacock test were run under four different samples sizes ($n = 10, 100, 1000, 5000$), with 10 replicates for each run. The Fasano–Franceschini test permutation procedure was further evaluated under four different permuted iterations (no permutation, 10, 100, 1000), again using 10 replicates for each run. Reported results represent the average run time of the 10 replicate benchmarks. All benchmark tests were run on a 2018 macBook Pro Mac (macOS Catalina) with a 2.7-GHz Quad-Core Intel Core i7 processor and 16 GB of 2133 MHz LPDDR3 memory.

The main distinction between the Peacock and Fasano–Franceschini tests is in computational efficiency, with Fasano–Franceschini scaling as $O(n^2)$ relative to Peacock's complexity of $O(n^3)$ (Lopes et al., 2007). Our benchmarks also show this advantage, as shown in **Figure 5A**. While the implementation of the permutation procedure increases runtime in comparison to the approximate fit from Press et al. (2007), parallelization of the Fasano–Franceschini test shows a four-fold reduction in run time when parallelized across 8 cores (**Figure 5B**).

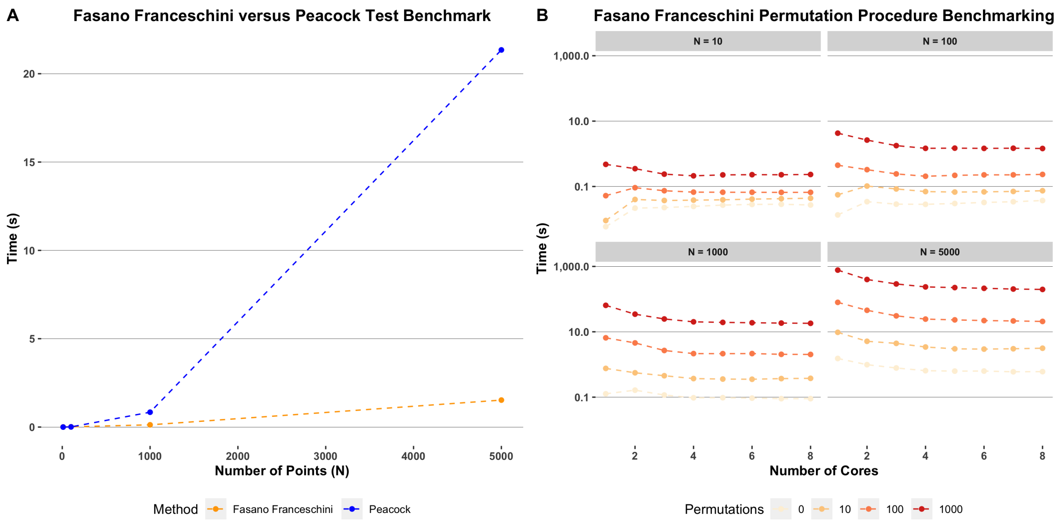

Figure 5: Computational efficiency benchmarks. **A:** Runtime of the Fasano–Franceschini test relative to the Peacock test at four different sample sizes ($n = 10, 100, 1000, 5000$). Points represent the average of 10 benchmark runs. **B:** Runtime of the Fasano–Franceschini permutation procedure for various sample sizes ($n = 10, 100, 1000, 5000$) as a function of the number of cores used. Within each panel, lines are colored by the number of permuted iterations (no permutation, 10, 100, 1000). Points represent the average of 10 benchmark runs. Note the logarithmic y-axis in (B).

Summary and discussion

The **fasano.franceschini.test** package is an R implementation of the 2-D two-sample KS test as defined by Fasano and Franceschini (Fasano and Franceschini, 1987). It improves upon existing packages by implementing a fast algorithm and a parallelized permutation procedure for improved statistical testing. Complete package documentation and source code is available via the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/` and the package website at `https://nesscoder.github.io/fasano.franceschini.test/`.

Computational details

The results in this paper were obtained using R 4.0.3 with the **fasano.franceschini.test** 1.0.0 package. R itself and all package dependencies (**methods** 4.0.3; **parallel** 4.0.3) are available from the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/`.

Acknowledgments

Research reported in this publication was supported by the NSF-Simons Center for Quantitative Biology at Northwestern University, an NSF-Simons MathBioSys Research Center. This work was supported by a grant from the Simons Foundation/SFARI (597491-RWC) and the National Science Foundation (1764421). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation and Simons Foundation.

E.N.C developed the **fasano.franceschini.test** package and produced the tutorials/documentation; E.N.C. and R.B. wrote the paper.

Bibliography

A two-dimensional Kolmogorov-Smirnov test for crowded field source detection: ROSAT sources in NGC 6397. *Monthly Notices of the Royal Astronomical Society*, 335(1):73–83, sep 2002. ISSN 00358711. doi: 10.1046/j.1365-8711.2002.05595.x. URL `https://academic-oup-com.turing.library.northwestern.edu/mnras/article/335/1/73/1033717`. [p7]

Quantitative, architectural analysis of immune cell subsets in tumor-draining lymph nodes from breast cancer patients and healthy lymph nodes. *PLoS ONE*, 5(8):e12420, 2010. ISSN 19326203. doi: 10.1371/journal.pone.0012420. URL `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0012420`. [p7]

S. Atasoy, L. Roseman, M. Kaelen, M. L. Kringelbach, G. Deco, and R. L. Carhart-Harris. Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under lsd. *Scientific Reports*, 7(1):1–18, 2017. [p1]

G. Babu and E. Feigelson. Astrostatistics: Goodness-of-fit and all that! In *Astronomical Data Analysis Software and Systems XV*, volume 351, page 127, 2006. URL `https://ui.adsabs.harvard.edu/abs/2006ASPC..351..127B/abstract`. [p1, 6]

W. Chen, D. O. Cahoy, J. G. Tasker, and A. W. Chiu. Kernel duration and modulation gain in a coupled oscillator model and their implications on the progression of seizures. *Network: Computation in Neural Systems*, 23(1-2):59–75, mar 2012. ISSN 0954898X. doi: 10.3109/0954898X.2012.678463. URL `https://www.tandfonline.com/doi/abs/10.3109/0954898X.2012.678463`. [p7]

F. Chiang, O. Mazdiyasni, and A. AghaKouchak. Amplified warming of droughts in southern united states in observations and model simulations. *Science Advances*, 4(8):eaat2380, 2018. [p1]

G. Fasano and A. Franceschini. A multidimensional version of the kolmogorov-smirnov test. *Monthly Notices of the Royal Astronomical Society*, 225(1):155–170, mar 1987. ISSN 0035-8711. doi: 10.1093/mnras/225.1.155. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/225.1.155`. [p1, 4, 5, 8]

J.-M. Floch, E. Marcon, and F. Puech. Spatial distribution of points. In V. Loonis and M.-P. d. Bellefon, editors, *Handbook of Spatial Analysis*, pages 71–111. Insee-Eurostat, 2018. URL `https://www.insee.fr/en/information/3635545`. [p6, 7]

J. M. Hahne, M. A. Schweisfurth, M. Koppe, and D. Farina. Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users. *Science Robotics*, 3(19), 2018. [p1]

S. M. Hargreaves, W. M. C. Araújo, E. Y. Nakano, and R. P. Zandonadi. Brazilian vegetarians diet quality markers and comparison with the general population: A nationwide cross-sectional study. *PloS One*, 15(5):e0232954, 2020. [p1]

S. Kaczanowska, D. W. Beury, V. Gopalan, A. K. Tycko, H. Qin, M. E. Clements, J. Drake, C. Nwanze, M. Murgai, Z. Rae, W. Ju, K. A. Alexander, J. Kline, C. F. Contreras, K. M. Wessel, S. Patel, S. Hannenhalli, M. C. Kelly, and R. N. Kaplan. Genetically engineered myeloid cells rebalance the core immune suppression program in metastasis. *Cell*, 184(8):2033–2052, 2021. [p1]

M. G. Kendall and A. Stuart. *The Advanced Theory of Statistics*, volume 2. Griffin, London, 1946. [p2]

A. N. Kolmogorov. Sulla Determinazione Empirica di Una Legge di Distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, (4):83–91, 1933a. [p1]

A. N. Kolmogorov. Über die Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Bull. Acad. Sci. URSS*, 1933(3):363–372, 1933b. URL `http://mi.mathnet.ru/eng/izv5009`. [p1]

R. H. C. Lopes, I. Reid, and P. R. Hobson. The two-dimensional Kolmogorov-Smirnov test. In *XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, 2007. URL `http://bura.brunel.ac.uk/handle/2438/1166`. [p1, 5, 7]

E. Marcon, S. Traissac, F. Puech, and G. Lang. Tools to characterize point patterns: Dbmss for R. *Journal of Statistical Software*, 67(CODESNIPPET3):1–15, oct 2015. ISSN 15487660. doi: 10.18637/jss.v067.c03. URL `https://www.jstatsoft.org/index.php/jss/article/view/v067c03/v67c03.pdfhttps://www.jstatsoft.org/index.php/jss/article/view/v067c03`. [p6, 7]

E. Ness-Cohn, M. Iwanaszko, W. L. Kath, R. Allada, and R. Braun. TimeTrial: An interactive application for optimizing the design and analysis of transcriptomic time-series data in circadian biology research. *Journal of Biological Rhythms*, 35(5):439–451, oct 2020. ISSN 15524531. doi: 10.1177/0748730420934672. URL `http://www.ncbi.nlm.nih.gov/pubmed/32613882`. [p6]

J. A. Peacock. Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society*, 202(3):615–627, mar 1983. ISSN 0035-8711. doi: 10.1093/mnras/202.3.615. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/202.3.615`. [p1, 2, 5]

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, USA, 3 edition, 2007. ISBN 0521880688. [p1, 4, 5, 6, 7]

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org/`. ISBN 3-900051-07-0. [p1]

N. V. Smirnov. Sur la distribution de ω^2 (criterium de M.R. v. Mises). *Com. Rend. Acad. Sci. (Paris)*, 202: 449–452, 1936. [p1]

N. V. Smirnov. On the distribution of the mises ω^2 criterion [in Russian]. *Rec. Math. N.S. [Mat. Sbornik]*, 2:973–993, 1937. URL `http://mi.mathnet.ru/msb5636`. [p1]

N. V. Smirnov. On the deviations of the empirical distribution curve [in Russian]. *Rec. Math. N.S. [Mat. Sbornik]*, 6(48):3–26, 1939. URL `http://mi.mathnet.ru/msb5810`. [p1]

N. V. Smirnov. Approximate laws of distribution of random variables from empirical data. *Uspehi Matem. Nauk*, 10:179–206, 1944. ISSN 0042-1316. [p1]

N. V. Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 1948. ISSN 0003-4851. doi: 10.1214/aoms/1177730256. [p1]

M. A. Stephens. Introduction to Kolmogorov (1933) On the Empirical Determination of a Distribution. pages 93–105. Springer, New York, NY, 1992. doi: 10.1007/978-1-4612-4380-9_9. URL `https://link.springer.com/chapter/10.1007/978-1-4612-4380-9_9`. [p2]

F. Wong and J. J. Collins. Evidence that coronavirus superspreading is fat-tailed. *Proceedings of the National Academy of Sciences*, 117(47):29416–29418, 2020. [p1]

Elan Ness-Cohn
Department of Molecular Biosciences
and
NSF-Simons Center for Quantitative Biology
Northwestern University
Evanston, IL 60208
ORCiD: 0000-0002-3935-6667
E-mail: `elan.ness-cohn@northwestern.edu`
Website: `https://www.nesscoder.com`

Rosemary Braun
Department of Molecular Biosciences,
Department of Engineering Sciences and Applied Mathematics,
Department of Physics and Astronomy,
Northwestern Institute on Complex Systems,
and
NSF-Simons Center for Quantitative Biology
Northwestern University
Evanston, IL 60208
ORCiD: 0000-0001-9668-9866
E-mail: `rbraun@northwestern.edu`
Website: `https://sites.northwestern.edu/braunlab/`