Statistics and Probability Letters 123 (2017) 177–182

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Fast calculation of boundary crossing probabilities for Poisson processes

Amit Moscovich *, Boaz Nadler

Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel

ARTICLE INFO

Article history:
Received 3 April 2016
Received in revised form 28 November 2016
Accepted 30 November 2016
Available online 20 December 2016

Keywords:
Boundary crossing
Poisson process
Empirical process
Goodness of fit
Brownian motion
First passage

ABSTRACT

The boundary crossing probability of a Poisson process with n jumps is a fundamental quantity with numerous applications. We present a fast $O(n^2 \log n)$ algorithm to calculate this probability for arbitrary upper and lower boundaries.

1. Introduction

Let $X_1, \ldots, X_n$ be n i.i.d. random variables drawn from a uniform distribution over $[0, 1]$ (henceforth abbreviated $U[0, 1]$) and let $\hat{F}_n$ be their empirical cumulative distribution function,

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \leq t).$$

Given two arbitrary functions $g, h : [0, 1] \to \mathbb{R}$, we define the corresponding *two-sided non-crossing probability* as

$$\Pr\left[\forall t \in [0, 1] : g(t) < \hat{F}_n(t) < h(t)\right]. \tag{1}$$

This probability plays a fundamental role in a wide range of applications, including the computation of p-values and power of sup-type continuous goodness-of-fit statistics (Kolmogorov, 1933; Steck, 1971; Noé and Vandewiele, 1968; Noé, 1972; Durbin, 1973; Kotel'Nikova and Khmaladze, 1983; Friedrich and Schellhaas, 1998; Khmaladze and Shinjikashvili, 2001); construction of confidence bands for empirical distribution functions (Owen, 1995; Frey, 2008; Matthews, 2013); change-point detection (Worsley, 1986); and sequential testing (Dongchu, 1998). Many of these applications consider a more general case, where $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$ for some known continuous distribution F. However, this setting is easily reducible

* Corresponding author.
E-mail addresses: moscovich@gmail.com (A. Moscovich), boaz.nadler@weizmann.ac.il (B. Nadler).

http://dx.doi.org/10.1016/j.spl.2016.11.027
0167-7152/

to the particular case $F = U[0,1]$ by transforming the random variables $X_i \mapsto F(X_i)$ and the boundary functions as $g^*(t) = g(F^{-1}(t))$ and $h^*(t) = h(F^{-1}(t))$.

One popular approach is to estimate Eq. (1) using Monte-Carlo methods. In the simplest of these methods one repeatedly generates $X_1, \ldots, X_n \sim U[0,1]$ and counts the number of times that the inequalities $g(t) < \hat{F}_n(t) < h(t)$ are satisfied for all t. This approach, however, can be extremely slow when the probability of interest is small and the sample size n is large.

The focus of this paper is on the fast computation of the *exact* two-sided crossing probability in Eq. (1) given arbitrary boundary functions. In the *one-sided* case (where either $g(t) < 0$ or $h(t) > 1$ for all $0 \le t \le 1$), Eq. (1) can be computed in $O(n^2)$ operations (Noé and Vandewiele, 1968; Kotel'Nikova and Khmaladze, 1983; Moscovich et al., 2016). Even faster solutions exist for some specific cases, such as a single linear boundary (Durbin, 1973). For general boundaries, however, essentially all existing methods require $O(n^3)$ operations (Steck, 1971; Durbin, 1971; Noé, 1972; Friedrich and Schellhaas, 1998; Khmaladze and Shinjikashvili, 2001).[1]

The main contribution of this paper is the introduction of a fast $O(n^2 \log n)$ algorithm to compute the two-sided crossing probability for general boundary functions. This is done by investigating a closely related problem involving a Poisson process. Specifically, let $\xi_n(t) : [0,1] \to \{0, 1, 2 \ldots\}$ be a homogeneous Poisson process of intensity n and let $g, h : [0,1] \to \mathbb{R}$ be two arbitrary boundaries. As noted in Section 3, there is a well known reduction from the probability of interest in Eq. (1) to the following two-sided non-crossing probability,

$$\Pr\left[\forall t \in [0,1] : g(t) < \xi_n(t) < h(t) \mid \xi_n(1) = k\right]. \tag{2}$$

The key observation in this paper, described in Section 2, is that the recursive solution to Eq. (2) given by Khmaladze and Shinjikashvili (2001) can be described as a series of at most $2n$ truncated linear convolutions involving vectors of length at most n. Using the Fast Fourier Transform (FFT), each convolution can thus be computed in $O(n \log n)$ operations, yielding a total running time of $O(n^2 \log n)$.

In Section 4 we present an application of the proposed method to the computation of p-values for a continuous goodness-of-fit statistic. Comparing the run-times of our algorithm to those of Khmaladze and Shinjikashvili (2001) shows that our method yields significant speedups for large sample sizes.

Finally, since Brownian motion can be described as a limit of a Poisson process, one may apply our method to approximate the boundary crossing probability and first passage time of a Brownian motion, see for example Khmaladze and Shinjikashvili (2001). The latter quantity has multiple applications in finance and statistics (Siegmund, 1986; Chicheportiche and Bouchaud, 2014). In this case an accurate approximation may require a fine discretization of the continuous boundaries, or equivalently a large value of n. Hence, fast algorithms are needed. Furthermore, our approach can be extended to higher dimensions, where it may be used to quickly approximate various quantities related to Brownian motion in 2 or 3 dimensions.

2. Boundary crossing probability for a Poisson process

In this section we describe our proposed algorithm for the fast computation of the two-sided non-crossing probability of a Poisson process, given in Eq. (2). We assume that $g(t) < h(t)$ for all $t \in [0,1]$ and that $g(0) < 0 < h(0)$, as otherwise the non-crossing probability is simply zero. Also, since the Poisson process is monotone, w.l.o.g. the two functions $g(t)$ and $h(t)$ may be assumed to be monotone non-decreasing. We start by describing the recursion formula of Khmaladze and Shinjikashvili (2001) whose direct application yields an $O(n^3)$ algorithm, and then show how to reduce the computational cost to $O(n^2 \log n)$ operations.

For every integer $i \in [0, g(1)]$, let $t_i^g = \inf\{t \in [0,1] : g(t) \ge i\}$ be the first time the function $g(t)$ passes the integer i. Similarly for every integer $i \in [h(0), h(1)]$, let $t_i^h = \sup\{t \in [0,1] : h(t) \le i\}$ be the last time the function $h(t)$ is bounded by i. Let $T(g) = \{t_i^g\}_{0 \le i \le g(1)}$ and $T(h) = \{t_i^h\}_{h(0) \le i \le h(1)}$ be the set of all integer crossing times for the two functions. A non-decreasing step function $f : [0,1] \to \{0, 1, 2, \ldots\}$ satisfies $g(t) < f(t) < h(t)$ for all $t \in [0,1]$ if and only if it satisfies these conditions at all discrete times $t \in T(g) \cup T(h) \cup \{1\}$. Hence, to compute the probabilities in equations (1) and (2), it suffices to analyze these inequalities only at a finite set of $N = |T(g) \cup T(h) \cup \{1\}|$ times.

Definition 1. Let $\xi_n(t)$ denote a one-dimensional Poisson process with intensity n. For any $s \in [0,1]$ and $m \in \{0, 1, 2, \ldots, n\}$, define $Q(s,m)$ as the probability that $\xi_n(s) = m$ and that ξ_n does not cross the boundaries $g(t), h(t)$ up to time s, i.e.

$$Q(s,m) := \Pr\left[\forall t \in [0,s] : g(t) < \xi_n(t) < h(t) \text{ and } \xi_n(s) = m\right].$$

[1] The procedure of Steck (1971) is based on the computation of an $n \times n$ matrix determinant and Durbin (1971) is based on solving a system of linear equations. Theoretically, using the Coppersmith–Winograd fast matrix multiplication algorithm, both methods yield an $O(n^{2.373})$ solution. However this method is never used in practice because of the huge constant factors involved.

Of particular interest are the values $Q(1, m)$ which correspond to Poisson processes that never cross the boundaries. Clearly $Q(0, 0) = 1$ and $\forall m > 0 : Q(0, m) = 0$. Let $t_1 < \cdots < t_N = 1$ denote the sorted set of times from $T(g) \cup T(h) \cup \{1\}$. For any $i \in \{0, \ldots, N-1\}$ and any $m \in \{0, 1, 2, \ldots\}$ the Chapman–Kolmogorov equations give

$$Q(t_{i+1}, m) = \begin{cases} \sum_{\ell=g(t_i)}^{m} Q(t_i, \ell) \cdot \Pr[Z_i = m - \ell] & \text{if } g(t_{i+1}) < m < h(t_{i+1}) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where Z_i is a Poisson random variable with intensity $n(t_{i+1} - t_i)$. This formula was proposed by Khmaladze and Shinjikashvili (2001) in order to compute $Q(1, n)$. All quantities up to the final time $t_N = 1$ can be computed recursively as follows: first calculate explicitly the probabilities $Q(t_1, 0), \ldots, Q(t_1, n)$ at time t_1. Next, calculate all probabilities at time t_{i+1} using the quantities from time t_i, and so on. Since each $Q(t_i, m)$ is a sum of up to $m + 1 \le n + 1$ terms and since $N \le 2n + 1$. the total run-time is at most $O(n^3)$, but may be smaller if the boundary functions $g(t)$, $h(t)$ are close to each other.

Next, we describe a faster procedure. Let $Q_{t_i} = (Q(t_i, 0), Q(t_i, 1), \ldots, Q(t_i, n))$ and let $\pi_\lambda = (\Pr[Z = 0], \Pr[Z = 1], \ldots, \Pr[Z = n])$, where $Z \sim \text{Poisson}(\lambda)$. The key observation is that the vector $Q_{t_{i+1}}$ in Eq. (3) is nothing but a *truncated linear convolution* of the vectors Q_{t_i} and $\pi_{n(t_{i+1}-t_i)}$. Hence we may apply the circular convolution theorem to compute it in the following fashion:

1. Append n zeros to the end of the two vectors Q_{t_i} and $\pi_{n(t_{i+1}-t_i)}$, denoting the resulting vectors Q^{2n} and π^{2n} respectively.
2. Compute the Fourier transform of the zero-extended vectors $\mathcal{F}\{Q^{2n}\}$ and $\mathcal{F}\{\pi^{2n}\}$.
3. Use the convolution theorem to obtain the Fourier transform of the convolution,
$$C^{2n} = \mathcal{F}\{Q^{2n} \star \pi^{2n}\} = \mathcal{F}\{Q^{2n}\} \cdot \mathcal{F}\{\pi^{2n}\},$$
where $\star$ denotes cyclic convolution and $\cdot$ denotes pointwise multiplication.
4. Compute the inverse Fourier transform of C^{2n} to yield the vector $Q_{t_{i+1}}$
$$Q_{t_{i+1}}(m) = \begin{cases} \mathcal{F}^{-1}\{C^{2n}\}(m) & \text{if } g(t_{i+1}) < m < h(t_{i+1}) \\ 0 & \text{otherwise.} \end{cases}$$

Using the FFT algorithm, each Fourier Transform takes $O(n \log n)$ time. Repeating these four steps for all times $t \in T(g) \cup T(h) \cup \{1\}$ yields a worst-case total run-time of $O(n^2 \log n)$. However, it may be much lower if the functions $g(t)$ and $h(t)$ are close to each other. For more details on the FFT and the computation of discrete convolutions, we refer the reader to Press et al. (1992, Chapters 12, 13).

3. Boundary crossing probability for the empirical CDF

We now return to the problem of calculating the probability in Eq. (1), that an empirical CDF will cross prescribed upper and lower boundaries. To simplify notation, we look at the scaled function $n\hat{F}_n(t)$ instead of $\hat{F}_n(t)$, and similarly to the previous section, consider the probabilities

$$R(s, m) = \Pr\left[\forall t \in [0, s] : g(t) < n\hat{F}_n(t) < h(t) \text{ and } n\hat{F}_n(t) = m\right].$$

Let $0 = t_0 < t_1 < \cdots < t_N = 1$ be as before, and let

$$Z_{\ell,i} \sim \text{Binomial}\left(n - \ell, \tfrac{t_{i+1}-t_i}{1-t_i}\right).$$

The Chapman–Kolmogorov equations give the recursive relations of Friedrich and Schellhaas (1998)

$$R(t_{i+1}, m) = \begin{cases} \sum_{\ell=g(t_i)}^{m} R(t_i, \ell) \cdot \Pr\left[Z_{\ell,i} = m - \ell\right] & \text{if } g(t_{i+1}) < m < h(t_{i+1}) \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In contrast to Eq. (3), the expression for $R_{t_{i+1}}$, the vector of probabilities at time t_{i+1}, is *not* in the form of a straightforward convolution, and hence cannot be directly computed using the FFT. While not the focus of our work, we note that by some algebraic manipulations, it is possible to compute Eq. (4) using a convolution and an additional $O(n)$ operations. Instead, we present a simpler construction that builds upon the results of the previous section. To this end we recall a well-known reduction from the empirical CDF to the Poisson process (Shorack and Wellner, 2009, Chapter 8, Proposition 2.2):

Lemma 1. *The distribution of the process $n\hat{F}_n(t)$ is identical to that of a Poisson process $\xi_n(t)$ with intensity n, conditioned on $\xi_n(1) = n$.*

According to this lemma, the non-crossing probability of an empirical CDF can be efficiently computed by a reduction to the Poisson case, since

$$\begin{aligned}\Pr\left[\forall t : g(t) < n\hat{F}_n(t) < h(t)\right] &= \Pr\left[\forall t : g(t) < \xi_n(t) < h(t) | \xi_n(1) = n\right] \\ &= \frac{\Pr\left[\forall t : g(t) < \xi_n(t) < h(t) \text{ and } \xi_n(1) = n\right]}{\Pr\left[\text{Poisson}(n) = n\right]} = \frac{Q(n, n)}{n^n e^{-n}/n!}\end{aligned} \tag{5}$$

and $Q(n, n)$ can be computed efficiently, as described in Section 2.

4. Computing *p*-values for goodness-of-fit statistics

The results of the previous sections can be used to compute the p-value of several two-sided continuous goodness-of-fit statistics such as Kolmogorov–Smirnov, and their power against specific alternatives. Our algorithm may also be applied to one-sided statistics such as the Higher-Criticism statistic of Donoho and Jin (2004).

To this end, we first recall the setup in the classical continuous goodness-of-fit testing problem (Lehmann and Romano, 2005). Let $x_1, \ldots, x_n \in \mathbb{R}$ be a sample of n observations. We wish to assess the validity of a null hypothesis that $x_1, \ldots, x_n$ are sampled i.i.d. from a known (and fully specified) continuous distribution function F against the alternative hypothesis that they are sampled from an unknown and arbitrary alternative G. Let $u_i = F(x_i)$ be the *probability integral transform* of the ith sample, and $u_{(1)} \leq u_{(2)} \leq \ldots \leq u_{(n)}$ be the sorted sequence of transformed samples. Under the null hypothesis, each u_i is uniformly distributed in $[0, 1]$ and therefore $u_{(i)}$ is the ith order statistic of a uniform distribution.

A common approach to goodness-of-fit testing is to measure the distance of the different order statistics from their expectation under the null. A classical example is the Kolmogorov–Smirnov statistic $K_n := \max\{K_n^-, K_n^+\}$, where K_n^- and K_n^+ are the one-sided KS statistics, defined as

$$K_n^- = \sqrt{n} \max_{i=1,\ldots,n} \left(u_{(i)} - \frac{i-1}{n}\right), \qquad K_n^+ = \sqrt{n} \max_{i=1,\ldots,n} \left(\frac{i}{n} - u_{(i)}\right).$$

More generally, given a sequence of monotone-increasing functions $r_1, \ldots, r_n : \mathbb{R} \to \mathbb{R}$ and a sequence of decreasing functions $s_1, \ldots, s_n : \mathbb{R} \to \mathbb{R}$, one may define one-sided goodness-of-fit statistics by

$$R := \max_{i=1,\ldots,n} r_i(u_{(i)}) \quad \text{and} \quad S := \max_{i=1,\ldots,n} s_i(u_{(i)}) \tag{6}$$

and a two-sided statistic by

$$T := \max\{R, S\}. \tag{7}$$

Statistics of this form include the supremum Anderson–Darling statistic and other weighted Kolmogorov–Smirnov statistics (Kolmogorov, 1933; Anderson and Darling, 1952), the R_n statistic of Berk and Jones (1979) and Phi-divergence supremum statistics (Jager and Wellner, 2007). Similarly, the one-sided Higher Criticism statistic of Donoho and Jin (2004) and its variants follow the form of the one-sided statistic S in Eq. (6).

It is easy to verify that $T < t$ if and only if $s_i^{-1}(t) < u_{(i)} < r_i^{-1}(t)$ holds for all i. Therefore, the p-value of $T = t$ is equal to

$$\Pr\left[T > t | \mathcal{H}_0\right] = 1 - \Pr\left[\forall 1 \leq i \leq n : s_i^{-1}(t) < U_{(i)} < r_i^{-1}(t)\right], \tag{8}$$

where $U_{(1)}, \ldots, U_{(n)}$ are the order statistics of n draws from $U[0, 1]$. Define two functions $g_t(x)$ and $h_t(x)$ as follows,

$$g_t(x) = \sum_{i=1}^{n} \mathbf{1}(r_i^{-1}(t) \leq x), \qquad h_t(x) = \sum_{i=1}^{n} \mathbf{1}(s_i^{-1}(t) \leq x),$$

then the probability of Eq. (8) is equal to that of Eq. (5) which we can compute in time $O(n^2 \log n)$.

4.1. Simulation results

We evaluate the empirical run-time of our procedure for computing p-values of the two-sided M_n and one-sided M_n^+ goodness-of-fit statistics of Berk and Jones (1979). These statistics have the form of Eqs. (6) and (7) but with a minimum instead of a maximum (see Moscovich et al., 2016, Section 3).

To this end we wrote an efficient implementation of the proposed procedure using the FFTW3 library by Frigo and Johnson (2005) and compared it to a direct implementation of the Khmaladze and Shinjikashvili (2001) recursion relations (denoted "KS 2001"). In addition, we implemented the $O(n^2)$ *one-sided* algorithm of Moscovich et al. (2016) (denoted "MNS 2016"). We find that both two-sided procedures are numerically stable using standard double-precision (64-bit) floating point numbers, even for sample sizes as large as $n = 250{,}000$. In contrast, the one-sided procedure (Moscovich et al., 2016) requires a careful numerical implementation using extended-precision (80-bit) floating point numbers and breaks down completely

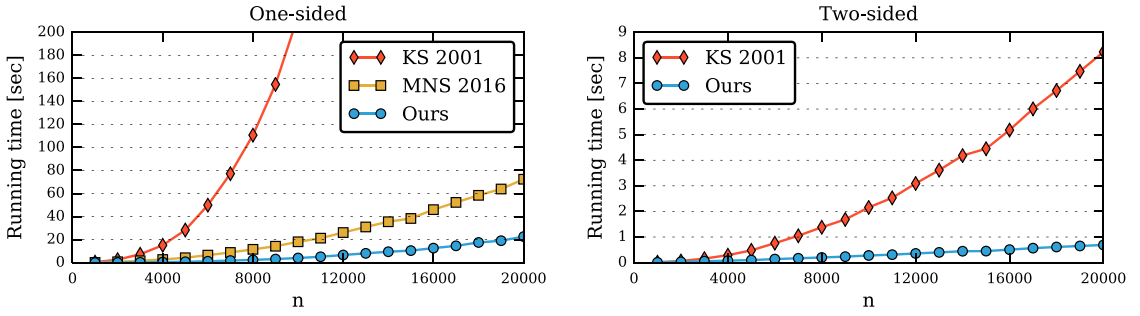

Fig. 1. Runtime comparison of our algorithm compared to that of Khmaladze and Shinjikashvili (2001) (KS 2001) and to the one-sided method described in Moscovich et al. (2016) (MNS 2016). The boundaries were chosen such that the p-value of M_n, equal to its two-sided boundary crossing probability, is 5%. Note that the one-sided case is much slower to compute.

for sample sizes $n > 50{,}000$. Fig. 1 presents a runtime comparison of the three algorithms for computing one-sided and two-sided crossing probabilities.[2]

Somewhat counter-intuitively, the one-sided case is much more expensive than the two-sided case. This is made clear by examining Eq. (3) and noting that in the one-sided case the variable m has a large valid range averaging around $n/2$, whereas in the two-sided case this range is typically much smaller. In all cases, our procedure is the fastest of all 3 methods. Surprisingly, this is true even in the one-sided case where the $O(n^2)$ procedure of Moscovich et al. (2016) is asymptotically superior.

Finally, we note that for large sample sizes, one may be inclined to forgo exact computation of p-values and instead use the asymptotic null distribution of the particular test statistic in use (assuming it is known). However, this does not always provide an adequate approximation, particularly as in several cases the finite sample distribution converges very slowly to its limiting form. Depending on the application, even the currently best known approximations may not be sufficiently accurate. For more on this topic, see Li and Siegmund (2015).

Acknowledgment

The authors would like to thank Jonathan Rosenblatt for computing infrastructure.

References

Anderson, T., Darling, D., 1952. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Ann. Math. Stat. 23 (2), 193–212. http://doi.org/10.1214/aoms/1177729437.
Berk, R.H., Jones, D.H., 1979. Goodness-of-fit test statistics that dominate the Kolmogorov statistics. Probab. Theory Related Fields 47, 47–59. http://doi.org/10.1007/BF00533250.
Chicheportiche, R., Bouchaud, J.-P., 2014. Some applications of first-passage ideas to finance. In: First-passage Phenom. Their Appl.. http://doi.org/10.1142/9789814590297_0018.
Dongchu, S., 1998. Exact computation for some sequential tests. Sequential Anal. 17 (2), 127–150. http://doi.org/10.1080/07474949808836403.
Donoho, D., Jin, J., 2004. Higher criticism for detecting sparse heterogeneous mixtures. Ann. Statist. 32 (3), 962–994. http://doi.org/10.1214/009053604000000265.
Durbin, J., 1971. Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov–Smirnov test. J. Appl. Probab. 8 (3), 431–453. http://doi.org/10.2307/3212169.
Durbin, J., 1973. Distribution Theory for Tests Based on Sample Distribution Function. SIAM, http://doi.org/10.1137/1.9781611970586.
Frey, J., 2008. Optimal distribution-free confidence bands for a distribution function. J. Statist. Plann. Inference 138, 3086–3098. http://doi.org/10.1016/j.jspi.2007.12.001.
Friedrich, T., Schellhaas, H., 1998. Computation of the percentage points and the power for the two-sided Kolmogorov–Smirnov one sample test. Statist. Papers 39, 361–375. http://doi.org/10.1007/BF02927099.
Frigo, M., Johnson, S.G., 2005. The design and implementation of FFTW3. Proc. IEEE 93 (2), 216–231. http://doi.org/10.1109/JPROC.2004.840301.
Jager, L., Wellner, J.A., 2007. Goodness-of-fit tests via phi-divergences. Ann. Statist. 35 (5), 2018–2053. http://doi.org/10.1214/0009053607000000244.
Khmaladze, E., Shinjikashvili, E., 2001. Calculation of noncrossing probabilities for Poisson processes and its corollaries. Adv. Appl. Probab. 33, 702–716. http://doi.org/10.1239/aap/1005091361.
Kolmogorov, A.N., 1933. Sulla determinazione empirica di una legge di distribuzione. G. dell'instituto Ital. degli attuari 4, 83–91.
Kotel'Nikova, V.F., Khmaladze, E.V., 1983. On computing the probability of an empirical process not crossing a curvilinear boundary. Theory Probab. Appl. 27, 640–648. http://doi.org/10.1137/1127075.
Lehmann, E.L., Romano, J.P., 2005. Testing Statistical Hypotheses, third ed.. In: Springer Texts in Statistics, Springer, New York, ISBN: 0-387-98864-5, http://doi.org/10.1007/0-387-27605-X.
Li, J., Siegmund, D., 2015. Higher criticism: p-values and criticism. Ann. Statist. 43 (3), 1323–1350. http://doi.org/10.1214/15-aos1312.
Matthews, D., 2013. Exact nonparametric confidence bands for the survivor function. Int. J. Biostat. 9 (2), 185–204. http://doi.org/10.1515/ijb-2012-0046.
Moscovich, A., Nadler, B., Spiegelman, C., 2016. On the exact Berk–Jones statistics and their p-value calculation. Electron. J. Stat. 10 (2), 2329–2354. http://doi.org/10.1214/16-EJS1172.

[2] C++ source code for all procedures is freely available at http://www.wisdom.weizmann.ac.il/~amitmo. The code was compiled using GCC 4.8.4 with maximum optimizations. The running times were measured on an Intel® Xeon® E5-4610 v2 2.30 GHz CPU.

Noé, M., 1972. The calculation of distributions of two-sided Kolmogorov–Smirnov type statistics. Ann. Math. Stat. 43 (1), 58–64. http://doi.org/10.1214/aoms/1177692700.
Noé, M., Vandewiele, G., 1968. The calculation of distributions of Kolmogorov–Smirnov type statistics including a table of significance points for a particular case. Ann. Math. Stat. 39 (1), 233–241. http://doi.org/10.1214/aoms/1177698523.
Owen, A.B., 1995. Nonparametric likelihood confidence bands for a distribution function. J. Amer. Statist. Assoc. 90 (430), 516–521. http://doi.org/10.2307/2291062.
Press, W.H., Flannery, B.P., Teukolsky, S.A., W T, V., 1992. Numerical Recipes in C: The Art of Scientific Computing, second ed. Cambridge University Press, http://apps.nrbook.com/c/index.html.
Shorack, G.R., Wellner, J.A., 2009. Empirical Processes with Applications to Statistics. SIAM, http://doi.org/10.1137/1.9780898719017.
Siegmund, D., 1986. Boundary crossing probabilities and statistical applications. Ann. Statist. 14 (2), 361–404. http://doi.org/10.1214/aos/1176349928.
Steck, G.P., 1971. Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. Ann. Math. Stat. 42 (1), 1–11. http://doi.org/10.1214/aoms/1177693490.
Worsley, K.J., 1986. Confidence regions and tests for a change point in a sequence of exponential family random variables. Biometrika 73 (1), 91–104. http://doi.org/10.1093/biomet/73.1.91.