Check for updates

J. Appl. Probab. 1–20 (2026)
*doi:*10.1017/jpr.2026.10125

THE *k*-WINDOW STATISTIC IN A RANDOM PERMUTATION

HOSAM MAHMOUD,* *The George Washington University*

Abstract

We introduce a new permutation statistic, which we call the k-window statistic, for permutations of $\{1, 2, \ldots, n\}$. The aim is to add to the repertoire of potentially useful tools in the design of efficient algorithms. The k-window statistic counts the number of elements that are at a displacement at most k from their natural positions in the ascendingly sorted permutation. We find the exact mean and exact variance of the k-window statistic, revealing phases in the mean and variance in the different regimes of k. We find a weak (or even strong) law of large numbers in the regime where $k = k_n$ is growing but remains $o(n)$. Via a representation of the statistic into dependent heterogeneous weakly correlated indicators, we establish the complete spectrum of distributions as k_n spans its entire range using an appropriate method for each regime.

Keywords: Permutation; displacement; Poisson limit; Gaussian law; method of factorial moments; combinatorial central limit theorem; Chen–Stein method; dependency neighborhood; phase transition

2020 Mathematics Subject Classification: Primary 05A05
Secondary 60C05; 60F05

1. Introduction

A focal point in the study of permutations is the characterization of the average or distribution of certain statistics. The goal is to discover better data processing algorithms. The hope is that such statistics may recommend or disqualify certain sorting algorithms.

Several statistics on permutations are classic, e.g. the number of cycles, the number of inversions, or the number of ascents, among a few others. They have underlying classical combinatorial numbers, such as Stirling numbers and Eulerian numbers [8, 9].

Our intention is to look at yet another permutation statistic to add to the repertoire of potentially useful tools in the design of efficient algorithms. We consider a statistic that counts the number of elements that are 'close' to their natural position in the sorted permutation in ascending order. The word 'close' here is captured in a prespecified parameter called k.

More precisely, we define the *k-window statistic* in the following way. Suppose $\Pi_n = (\pi_1, \ldots, \pi_n)$ is a given permutation of $[n] := \{1, \ldots, n\}$, with π_i being the position of $i \in [n]$ in the permutation. For $0 \le k \le n-1$, let $C_k(\Pi_n)$ be the number of elements in Π_n that are at a distance no greater than k from their natural position (the *natural position* π_i for i is i). In other

Received 17 January 2026 UTC; accepted 6 July 2026 UTC.
* Postal address: Department of Statistics, The George Washington University, Washington, DC 20052, USA. Email address: hosam@gwu.edu

© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

words, we set

$$C_k(\Pi_n) = |\{i\colon\ |\pi_i - i| \le k\}|;$$

applied to a set $\mathcal{A}$, the notation $|\mathcal{A}|$ is the cardinality or size of the set. For brevity, we drop Π_n from the notation and simply write $C_{n,k}$.

Another way to think of $C_{n,k}$ is to put a viewing window of width $2k+1$ centered at every position $i \in [n]$, and add 1 to the count $C_{n,k}$, for every i seen in its associated window. This suggests a name for the statistic. Let us call it the k-window statistic (a displacement of no more than k positions to the left or k positions to the right).

The 0-window statistic counts the number of fixed points in a permutation. This has been studied; see [4, Example 3.4.2], for instance. To the best of our knowledge, the k-window statistic has not been considered much for other values of k, particularly for k growing with n. We look at the entire spectrum of k; the forthcoming analysis shows phase transitions in the asymptotic average, variance, and distribution of $C_{n,k}$.

We focus on *random permutations*, where the randomness means that all permutations of $[n]$ are equally likely. Indeed, a large class of data has ranks that are random permutations. For instance, the ranks of the observations in a sample of size n taken from *any* continuous distribution are almost surely in a random permutation of $[n]$.

In what follows, we use the notation $\mathbf{1}_{\mathcal{E}}$ for the indicator of the event $\mathcal{E}$. The k-window statistic can be represented as a sum of indicators:

$$C_{n,k} = \sum_{i=1}^{n} \mathbf{1}_{\{|\pi_i - i| \le k\}}. \tag{1.1}$$

The indicators involved are *dependent*. They are heterogeneous in the sense they are not identically distributed and the dependency between different pairs may not be the same. For instance, for $n=3$, and $k=1$, the two indicators $\mathbf{1}_{\{|\pi_1-1|\le 1\}}$ and $\mathbf{1}_{\{|\pi_2-2|\le 1\}}$ are uncorrelated, whereas the covariance between $\mathbf{1}_{\{|\pi_1-1|\le 1\}}$ and $\mathbf{1}_{\{|\pi_3-3|\le 1\}}$ is $\frac{1}{18}$.

Consider an $n \times n$ matrix (which we call $\mathbf{A}_n$) with rows indexed by i and columns indexed by r, and filled with 1s and 0s. The 1s appear in a band around the main diagonal, where the absolute difference between i and r is at most k, and the rest of the entries in the matrix are 0s. Figure 1 shows the matrix $\mathbf{A}_8$, with $k=2$ (windows of width 5 in (a)) and with $k=5$ (windows of width 11 in (b)). The positions containing 1 are colored pink.

Suppose the permutation in the positional notation is

$$3 \quad 8 \quad 1 \quad 6 \quad 2 \quad 7 \quad 5 \quad 4.$$

The bullets in both grids show where the elements of this permutation are. For example, the element 1 appears in the third position in the permutation: cell (1, 3) receives a bullet and is seen in a window of width 5 (in the pink region in (a)), contributing 1 to $C_{8,2}$. It is also seen in a window of length 11 (in the pink region in (b)), contributing 1 to $C_{8,5}$. The element 5 appears at position 2 in the permutation, which is not in the pink area in (a), contributing 0 to $C_{8,2}$, but is in the pink region in (b), contributing 1 to $C_{8,5}$. For this permutation, we compute

$$C_{8,2} = 1+0+1+1+0+1+1+0 = 5, \quad C_{8,5} = 1+0+1+1+1+1+1+1 = 7.$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

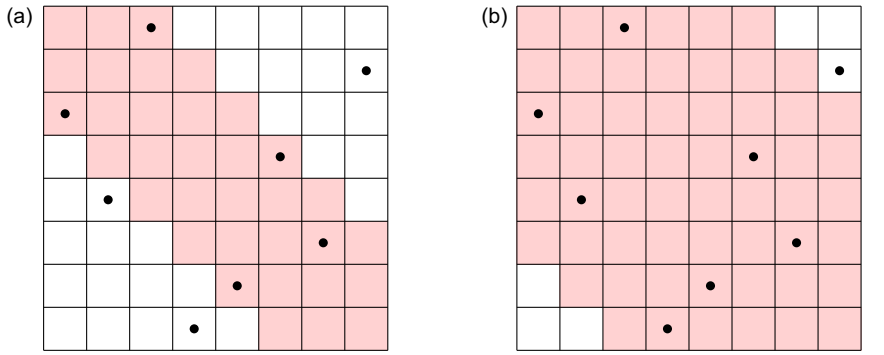


FIGURE 1. The permutation $\mathbf{\Pi}_8 = (3, 8, 1, 6, 2, 7, 5, 4)$ represented as bullets in an 8×8 grid appearing at the positions (i, π_i). The pink shaded cells satisfy $|i - j| \le k$; (a) $k = 2$, (b) $k = 5$.

2. The mean of the *k*-window statistic

We pin down the exact mean of $C_{n,k}$, from which we can extract asymptotics and discern phases within.

Proposition 2.1. *Let k be an integer between 0 and $n-1$. In a random permutation of $[n]$, the average of the k-window statistic is*

$$\mathbb{E}[C_{n,k}] = 2k + 1 - \frac{k}{n}(k+1).$$

Proof. Take expectations of (1.1) to get

$$\mathbb{E}[C_{n,k}] = \sum_{i=1}^{n}\sum_{r=1}^{n} \mathbf{1}_{\{|r-i|\le k\}}\, \mathbb{P}(\pi_i = r) = \frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{n} \mathbf{1}_{\{|r-i|\le k\}}.$$

Here and throughout, we call the sum $\sum_{r=1}^{n} \mathbf{1}_{\{|r-i|\le k\}}$ by the name $x_{n,i,k}$. In view of edge effects, we evaluate $x_{n,i,k}$ piecewise and deal with cases. Thus $x_{n,i,k}$ counts the 1s across row i in this matrix.

The case $0 \le k < n/2$. We find

$$x_{n,i,k} = \begin{cases} k+i, & \text{for } i = 1, \dots, k, \\ 2k+1, & \text{for } i = k+1, \dots, n-k, \\ k+1+n-i, & \text{for } i = n-k+1, \dots, n. \end{cases}$$

It follows that

$$\begin{aligned} \sum_{i=1}^{n} x_{n,i,k} &= \sum_{i=1}^{k}(k+i) + \sum_{i=k+1}^{n-k}(2k+1) + \sum_{i=n-k+1}^{n}(k+1+n-i) \\ &= (2k+1)n - k(k+1). \end{aligned} \tag{2.1}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

The case $k \geq n/2$. We find

$$x_{n,i,k} = \begin{cases} k+i, & \text{for } i = 1, \ldots, n-k-1, \\ n, & \text{for } i = n-k, \ldots, k+1, \\ k+1+n-i, & \text{for } i = k+2, \ldots, n. \end{cases}$$

It follows that

$$\begin{aligned} \sum_{i=1}^{n} x_{n,i,k} &= \sum_{i=1}^{n-k-1} (k+i) + \sum_{i=n-k}^{k+1} n + \sum_{i=k+2}^{n} (k+1+n-i) \\ &= (2k+1)n - k(k+1), \end{aligned}$$

which is the same expression as the counterpart in the case $k < n/2$. In either case, we get

$$\mathbb{E}[C_{n,k}] = \frac{1}{n} \sum_{i=1}^{n} x_{n,i,k} = (2k+1) - \frac{k}{n}(k+1). \qquad \square$$

2.1. Phases in the average of $C_{n,k}$

We keep in mind that k may depend on n. We write

$$k = k_n = \alpha n + g_n,$$

where g_n is $o(n)$. Depending on context, we simply write k when k is fixed, or appears in huge formulas, or when suggested by aesthetics. In other places, where convergence is discussed, it is preferable to write k_n.

The case $\alpha = 0$ covers slowly growing (sublinear) k, such as when k is a fixed number (in which case $g_n = k$), or a case such as $k_n = 0 + g_n = \lfloor 6n^{3/4} + \pi\sqrt{n} + e \ln n - 5 \rfloor$.

To determine the asymptotics of $\mathbb{E}[C_{n,k}]$, we use the Θ notation, which specifies the exact growth rate of a function. The notation Θ is popular in the literature of computer science. For a function g_n, if there exist two real positive constants c_1 and c_2, and a natural number n_0, such that

$$c_1 |h_n| \leq |g_n| \leq c_2 |h_n| \quad \text{for all } n \geq n_0,$$

we say that g_n is of the same *exact* order of growth as the function h_n and write this as $g_n = \Theta(h_n)$. Consider first the regime in which $\alpha \in [0, 1)$. Plugging the representation of k_n into the exact expression of Proposition 2.1, we get

$$\begin{aligned} \mathbb{E}[C_{n,k}] &= 2(\alpha n + g_n) + 1 - \frac{\alpha n + g_n}{n}(\alpha n + g_n + 1) \\ &= \alpha(2-\alpha)n + 2(1-\alpha)g_n + 1 - \alpha - \frac{g_n^2 + g_n}{n}. \end{aligned}$$

The order of g_n^2/n is always $o(g_n)$. Hence, for k in the sublinear regime (where $\alpha = 0$), we have

$$\mathbb{E}[C_{n,k}] = 2g_n + 1 + \Theta\left(\frac{g_n^2}{n}\right);$$

the average of the k-window statistic is also in a sublinear phase.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

TABLE 1. Asymptotics for some selected k_n functions.

α	k_n	g_n	Asym. $\mathbb{E}[C_{n,k_n}]$	Asym. Var $[C_{n,k_n}]$
0	k	k	$2k+1+\Theta\left(\frac{1}{n}\right)$	$2k+1$
0	$\lfloor \ln n \rfloor$	$\lfloor \ln n \rfloor$	$2\ln n + \mathrm{O}(1)$	$2\ln n$
0	$\left\lceil 6n^{1/4} - \frac{1}{n} \right\rceil$	$\left\lceil 6n^{1/4} - \frac{1}{n} \right\rceil$	$12n^{1/4} + \mathrm{O}(1)$	$12n^{1/4}$
0	$\left\lceil 8n^{3/4} + 1.9n^{3/5} \right\rceil$	$\left\lceil 8n^{3/4} + 1.9n^{3/5} \right\rceil$	$16n^{3/4} + \Theta(n^{3/5})$	$16n^{3/4}$
0	$\left\lceil \frac{n}{\ln n} \right\rceil$	$\left\lceil \frac{n}{\ln n} \right\rceil$	$\frac{2n}{\ln n} + \Theta\left(\frac{n}{\ln^2 n}\right)$	$\frac{2n}{\ln n}$
$\frac{1}{3}$	$\left\lfloor \frac{1}{3}n + en^{1/6} \right\rfloor$	$en^{1/6} - r_n$	$\frac{5}{9}n + \Theta(n^{1/6})$	$\frac{20}{81}n$
$\frac{9}{10}$	$\left\lfloor \frac{9}{10}n + \pi^2 \right\rfloor$	$\pi^2 - \tilde{r}_n$	$\frac{99}{100}n + \Theta(1)$	$\frac{263}{30000}n$
1	$n - \left\lceil \frac{n}{\ln n} \right\rceil$	$-\left\lceil \frac{n}{\ln n} \right\rceil$	$n - \Theta\left(\frac{n}{\ln^2 n}\right)$	$\Theta\left(\frac{n}{\ln^2 n}\right)$
1	$n - \lfloor 3.7\sqrt{n} \rfloor + 9$	$-\lfloor 3.7\sqrt{n} \rfloor + 9$	$n - \Theta(1)$	$\Theta(1)$
1	$n - 2\lceil \ln n \rceil$	$-2\lceil \ln n \rceil$	$n - \Theta\left(\frac{\ln^2 n}{n}\right)$	$\Theta\left(\frac{\ln^2 n}{n}\right)$
1	$n-4$	-4	$n - \Theta\left(\frac{1}{n}\right)$	$\Theta\left(\frac{1}{n}\right)$

In the linear regime $k_n = \alpha n + g_n$ the function g_n is $\mathrm{o}(n)$, and if $\alpha = 1$ this function is negative. With $\alpha \in (0, 1]$, the average comes up into a linear phase too:

$$\mathbb{E}[C_{n,k}] = \alpha(2-\alpha)n + 2(1-\alpha)g_n + 1 - \alpha - \Theta\left(\frac{g_n^2}{n}\right). \tag{2.2}$$

It is evident from the asymptotic approximation that we get a sublinear (in n) average for the k-window statistic, for $\alpha = 0$, then the average transitions into a linear phase, when α is in (0, 1].

In the linear phase, the coefficient of linearity in the average increases with rising α. Table 1 shows the asymptotics of the average of C_{n,k_n} for some selected k_n functions. In the table, the functions r_n and $\tilde{r}_n$ are corrections with values in [0, 1) that appear upon removing the floors.

Remark 2.1. There is no limit to how bizarre the function k_n can be. For example, one could think of a function alternating between high and low values, such as

$$k_n = \begin{cases} 1, & \text{for odd } n, \\ \lfloor \sqrt{n} \rfloor, & \text{for even } n. \end{cases}$$

Such a concocted sequence is not likely to appear in practice, and therefore is of no interest. We confine our attention to functions k_n that are constants or driven by a dominant naturally growing function.

3. The variance of the *k*-window statistic

Toward the exact variance of $C_{n,k}$, we begin with its second moment. Squaring the stochastic equation (1.1), we get

$$C_{n,k}^2 = \sum_{1 \le i,j \le n} \mathbf{1}_{\{|\pi_i - i| \le k\}} \mathbf{1}_{\{|\pi_j - j| \le k\}}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

$$= \sum_{i=1}^{n} \mathbf{1}_{\{|\pi_i - i| \le k\}} + \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \mathbf{1}_{\{|\pi_i - i| \le k\}} \mathbf{1}_{\{|\pi_j - j| \le k\}}$$

$$= C_{n,k} + \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \mathbf{1}_{\{|\pi_i - i| \le k\}} \mathbf{1}_{\{|\pi_j - j| \le k\}}. \tag{3.1}$$

Proposition 3.1. *Let k be an integer between 0 and $n-1$. In a random permutation of $[n]$, the variance of the k-window statistic is*

$$\operatorname{Var}[C_{n,k}] = \begin{cases} (3(2k+1)n^3 - 3(5k^2+5k+1)n^2 \\ \quad + 4k(k+1)(2k+1)n + 3k^2(k+1)^2)/(3n^2(n-1)), & \text{for } 0 \le k < n/2, \\ ((n-k-1)(n-k) \\ \quad \times (2n^2 - (2k+1)n + 3k(k+1))/(3n^2(n-1)), & \text{for } n/2 \le k < n. \end{cases}$$

Proof. Take expectations of (3.1), noting that for distinct i and j, the positions π_i and π_j are distinct, and the joint distribution of (π_i, π_j) is uniform over $[n] \times [n] - \{(i,i) : i = 1, \ldots, n\}$. This yields

$$\mathbb{E}\left[C_{n,k}^2\right] = \mathbb{E}[C_{n,k}] + \frac{1}{n(n-1)} \sum_{\substack{1 \le i,j,r,s \le n \\ i \ne j, r \ne s}} \mathbf{1}_{\{|r-i| \le k\}} \mathbf{1}_{\{|s-j| \le k\}}.$$

For economy of space, we introduce the notation $\mathbf{J}_{p,q,k} = \mathbf{1}_{\{|p-q| \le k\}}$ to write the last relation as

$$\mathbb{E}\left[C_{n,k}^2\right] = \mathbb{E}[C_{n,k}] + \frac{1}{n(n-1)} \sum_{\substack{1 \le i,j,r,s \le n \\ i \ne j, r \ne s}} \mathbf{J}_{r,i,k} \mathbf{J}_{s,j,k}$$

$$= \mathbb{E}[C_{n,k}] + \frac{1}{n(n-1)} \left(\sum_{\substack{1 \le i,j,r,s \le n \\ r \ne s}} \mathbf{J}_{r,i,k} \mathbf{J}_{s,j,k} - \sum_{1 \le i,r,s \le n \ r \ne s} \mathbf{J}_{r,i,k} \mathbf{J}_{s,i,k} \right)$$

$$=: \mathbb{E}[C_{n,k}] + \frac{1}{n(n-1)} (L_{n,k} - M_{n,k}),$$

where $L_{n,k}$ and $M_{n,k}$ are the two four-fold sums in the penultimate line, defined in their order of appearance.

We have an exact expression for $\mathbb{E}[C_{n,k}]$ from Proposition 2.1. The evaluation of these four-fold sums requires a very lengthy computation that can be distracting; we only give the result here and relegate the detailed derivation to the appendix. We have

$$L_{n,k} = \begin{cases} (2k+1)((2k+1)n^2 - (2k^2+4k+1)n) \\ \quad + \frac{1}{3}k(k+1)(3k^2+13k+5), & \text{for } 0 \le k < n/2, \\ \frac{1}{3}n^3 + 2k(2k+1)n^2 - (4k^3+6k^2+2k+\frac{1}{3})n \\ \quad + \frac{1}{3}k(k+1)(3k^2+5k+1), & \text{for } n/2 \le k < n, \end{cases}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

and

$$M_{n,k} = \begin{cases} 2k(2k+1)n - \frac{2}{3}k(k+1)(5k+1), & \text{for } 0 \le k < n/2, \\ -\frac{1}{3}n^3 + (2k+1)n^2 - \frac{2}{3}(3k+1)n - \frac{2}{3}k(k^2-1), & \text{for } n/2 \le k < n. \end{cases}$$

Assembling all the elements and simplifying, we get

$$\mathbb{E}\big[C_{n,k}^2\big] = \begin{cases} (6(k+1)(2k+1)n^2 - 3(4k^3+15k^2+11k+2)n \\ \quad + k(k+1)(3k^2+23k+10))/(3n(n-1)), & \text{for } 0 \le k < n/2, \\ (2n^3 + 6k(2k+1)n^2 - (12k^3+21k^2+9k+2)n \\ \quad + k(k+1)(k+2)(3k+1))/(3n(n-1)), & \text{for } n/2 \le k < n. \end{cases}$$

The variance follows from the computation of $\mathbb{E}\big[C_{n,k}^2\big] - \mathbb{E}^2[C_{n,k}]$ and further algebraic simplification yielding the stated expression. □

3.1. Phases in the variance of $C_{n,k}$

We can now identify phases in the variance. Let $k = k_n = \alpha n + g_n$, with $g_n = o(n)$, as $n \to \infty$. For $\alpha = 0$, we have the asymptotic equivalent

$$\mathrm{Var}\,[C_{n,k}] \sim 2g_n + 1.$$

In contrast, for $0 < \alpha < \frac{1}{2}$, the variance is in the linear phase

$$\mathrm{Var}\,[C_{n,k}] = \alpha\left(\alpha^3 + \frac{8}{3}\alpha^2 - 5\alpha + 2\right)n + \Theta(g_n),$$

whereas for $\frac{1}{2} \le \alpha < 1$, the variance enters the linear phase

$$\mathrm{Var}\,[C_{n,k}] = \left(\alpha^4 - \frac{8}{3}\alpha^3 + 3\alpha^2 - 2\alpha + \frac{2}{3}\right)n + \Theta(g_n).$$

We note that at $\alpha = \frac{1}{2}$, the linearity coefficients $\alpha\big(\alpha^3 + \frac{8}{3}\alpha^2 - 5\alpha + 2\big)$ and $\big(\alpha^4 - \frac{8}{3}\alpha^3 + 3\alpha^2 - 2\alpha + \frac{2}{3}\big)$ are both equal to $\frac{7}{48}$, showing that the two linear phases meet smoothly at the seam line. Also, note that the coefficient of n is 0 at $\alpha = 0$ and again at $\alpha = 1$. Figure 2 shows a plot of $\rho(\alpha)$, the coefficient of n in $\mathrm{Var}\,[C_{n,k}]$, for $\alpha \in [0, 1]$.

The regime in which $\alpha = 1$ offers its own twists. Within this regime, k_n assumes the form $n + g_n$, for some negative function $g_n = o(n)$. The exact variance in Proposition 3.1 in this regime degenerates into

$$\begin{aligned} \mathrm{Var}\,[C_{n,k}] &= \frac{\big(n^2 - \frac{4}{3}ng_n + g_n^2 + \frac{2}{3}n - g_n\big)g_n(g_n-1)}{n^2(n-1)} \\ &\sim \frac{g_n^2}{n} \quad \text{as } n \to \infty. \end{aligned}$$

In this regime of extremely large windows, if the function g_n grows at a rate faster than $\sqrt{n}$ (but remains $o(n)$, of course), we get a growing variance. At $g_n \sim c\sqrt{n}$, the variance is asymptotic to c^2. However, if $g_n = o(\sqrt{n})$, the variance diminishes to 0 as $n \to \infty$. Table 1 shows the asymptotics of the variance of C_{n,k_n} for some selected k_n functions.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

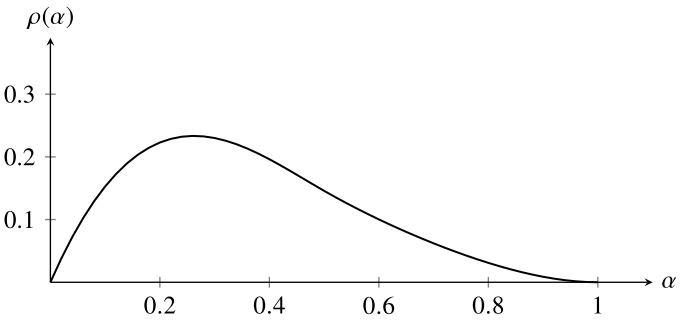


FIGURE 2. The coefficient $\rho(\alpha)$ of n as a function of α.

4. Concentration

We use the notation $\xrightarrow{\mathbb{P}}$ for convergence in probability, and $\xrightarrow{\text{a.s.}}$ for convergence almost surely. In the regime where $\alpha = 0$, but g_n is growing and the regime $0 < \alpha < 1$, the variance is small enough relative to the square of the mean, and this gives us a weak law of large numbers for the k-window statistic in the following way. According to the Chebyshev inequality, in the sublinear regime k_n, with g_n growing, for all $\varepsilon > 0$, we have

$$\mathbb{P}(|C_{n,k} - \mathbb{E}[C_{n,k}]| > \varepsilon\, \mathbb{E}[C_{n,k}]) \le \frac{\text{Var}\,[C_{n,k}]}{\varepsilon^2\, \mathbb{E}^2[C_{n,k}]} = \frac{\Theta(g_n)}{\Theta(g_n^2)} \to 0 \quad \text{as } n \to \infty.$$

Hence, for $\alpha = 0$ and growing g_n, we have the convergence $C_{n,k}/\mathbb{E}[C_{n,k}] \xrightarrow{\mathbb{P}} 1$. Yet, we have $\mathbb{E}[C_{n,k}]/g_n \to 2$. Combining these convergence relations, we get

$$\frac{C_{n,k}}{g_n} \xrightarrow{\mathbb{P}} 2 \quad \text{as } n \to \infty.$$

In the linear regime $k_n \sim \alpha n$ with $\alpha \in (0, 1)$, we have $\mathbb{E}[C_{n,k}] \sim \alpha(2-\alpha)n$, and $\text{Var}\,[C_{n,k}] \sim \rho(\alpha)n$. For all $\varepsilon > 0$, we have

$$\mathbb{P}(|C_{n,k} - \mathbb{E}[C_{n,k}]| > \varepsilon\, \mathbb{E}[C_{n,k}]) = \frac{\Theta(n)}{\Theta(n^2)} \to 0 \quad \text{as } n \to \infty.$$

Hence we have the convergence $C_{n,k}/\mathbb{E}[C_{n,k}] \xrightarrow{\mathbb{P}} 1$. Yet, we have $\mathbb{E}[C_{n,k}]/n \to \alpha(2-\alpha)$. Combining these convergence relations, we get

$$\frac{C_{n,k}}{n} \xrightarrow{\mathbb{P}} \alpha(2-\alpha) \quad \text{as } n \to \infty.$$

In the regime $\alpha = 1$, the Chebyshev bound $\text{Var}\,[C_{n,k}]/(\varepsilon^2\mathbb{E}^2[C_{n,k}])$ is $\Theta(g_n^2/n^3)$. Hence we have the convergence $C_{n,k}/\mathbb{E}[C_{n,k}] \xrightarrow{\mathbb{P}} 1$. Yet, we have $\mathbb{E}[C_{n,k}]/n \to 1$. Combining these convergence relations, we get

$$\frac{C_{n,k}}{n} \xrightarrow{\mathbb{P}} 1 \quad \text{as } n \to \infty.$$

If the Chebyshev bound is $\text{O}(n^{-(1+\delta)})$, for some $\delta > 0$, the probability on the left-hand side of Chebyshev's inequality becomes summable. Via the Borel–Cantelli lemma, the convergence

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

is upgraded to be almost sure. That is, in the regime $k_n = \alpha n - \mathrm{O}(n^{1-\delta/2})$, for some $\delta > 0$, we have $C_{n,k}/n \xrightarrow{\text{a.s.}} 1$. For instance, for $k_n = n - 5\lfloor n^{3/4} \rfloor$, we get the bound $\Theta(n^{-3/2})$ in the Chebyshev inequality, with the conclusion that $C_{n,k}/n \xrightarrow{\text{a.s.}} 1$.

Remark 4.1. The regime of fixed k is excluded from this approach to the law of large numbers.

Remark 4.2. As in (2.2) and the concluding discussion before Remark 4.1, there is something special about the regime $\alpha = 1$. The presence of $1 - \alpha = 0$ in the formula (2.2) for $\mathbb{E}[C_{n,k}]$ eliminates the terms $2(1-\alpha)g_n$ and $1 - \alpha$ before it gets down to the diminishing order g_n^2/n.

5. The full spectrum of distributions

We denote the Poisson random variable with parameter $\lambda > 0$ by $\mathrm{Poi}\,(\lambda)$, and denote the centered normal random variable with variance σ^2 by $\mathcal{N}(0, \sigma^2)$, with cumulative distribution function $\mathbb{P}(\mathcal{N}(0, 1) \le x) = \Phi(x)$. We denote convergence in distribution by the symbol $\xrightarrow{\mathcal{D}}$. For the r-times falling factorial of a number y, we use Pochhammer's symbol

$$(y)_r = y(y-1)\cdots(y-r+1).$$

The method of convergence of all falling factorial moments is used to establish convergence in distribution in the case of fixed k. In certain regimes, Bolthausen's theorem [6] is instrumental for establishing the rate of convergence to Gaussianity. It is one of several variations and improvements over the classic combinatorial central limit theorem by Hoeffding [10]. For other improvements, see [3] and [4]. For the reader's convenience, we give a statement of this theorem.

Theorem 5.1. (Bolthausen [6].) *Let $\mathbf{A}_n = (a_{i,j})$ be an $n \times n$ matrix of real numbers. Let*

$$\mu_{\mathbf{A}_n} = n a_{..}, \quad \sigma^2_{\mathbf{A}_n} = \frac{1}{n-1} \sum_{1 \le i,j \le n} (a_{ij} - a_{i.} - a_{.j} + a_{..})^2,$$

where

$$a_{i.} = \frac{1}{n}\sum_{j=1}^{n} a_{ij}, \quad a_{.j} = \frac{1}{n}\sum_{i=1}^{n} a_{ij}, \quad a_{..} = \frac{1}{n^2} \sum_{1 \le i,j \le n} a_{ij}.$$

For a uniform random permutation $(\pi_1, \ldots, \pi_n)$, we have

$$\sup_x \left(\left| \mathbb{P}\left(\frac{a_{i,\pi_n} - \mu_{\mathbf{A}_n}}{\sigma_{\mathbf{A}_n}} \le x \right) - \Phi(x) \right| \right) \le \frac{K}{n} \sum_{1 \le i,j \le n} \left| \frac{a_{ij} - a_{i.} - a_{.j} + a_{..}}{\sigma_{\mathbf{A}_n}} \right|^3, \tag{5.1}$$

for some positive constant K.

We apply Theorem 5.1 to the specialized k-window matrix, $\mathbf{A}_n$, described at the end of Section 1. Note that in the notation of the present paper, $a_{i.}$ is $x_{n,i,k}/n$. For this specific matrix, its symmetry renders $a_{.j} = x_{n,j,k}/n$ too.

In certain regimes, the Chen–Stein method is the right tool. We use the version stated in [2] (proved in [1]). For a wider scope see [4]. We state the theorem here for the reader's convenience.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

The theorem is based on bounding the total variation distance. For non-negative integer-valued random variables X and Y, let $\mathbb{P}_X$ and $\mathbb{Q}_Y$ be the measures specifying the distributions of X and Y over $\mathbb{N} \cup \{0\}$. The total variation distance is defined as

$$d_{TV}(\mathbb{P}_X, \mathbb{Q}_Y) = \sup_{A \subseteq \mathbb{N} \cup \{0\}} |\mathbb{P}_X(A) - \mathbb{Q}_Y(A)| = \frac{1}{2} \sum_{j=0}^{\infty} |\mathbb{P}_X(X = j) - \mathbb{Q}_Y(Y = j)|.$$

Many authors directly write $d_{TV}(X, Y)$ for $d_{TV}(\mathbb{P}_X, \mathbb{Q}_Y)$. We follow this simplifying convention, but it should be understood that the total variation distance is between probability laws, not between random variables. It is well known that convergence in total variation distance implies convergence in distribution.

Theorem 5.2. (Arratia, Goldstein, and Gordon [2].) *Let $\mathcal{I}$ be an arbitrary index set, and for $\alpha \in \mathcal{I}$, let X_α be a Bernoulli random variable with $\mathbb{P}(X_\alpha = 1) > 0$. Let*

$$W = \sum_{\alpha \in \mathcal{I}} X_\alpha \quad \textit{and} \quad \lambda = \mathbb{E}[W] = \sum_{\alpha \in \mathcal{I}} \mathbb{E}[X_\alpha].$$

We assume that $\lambda \in (0, \infty)$. For each $\alpha \in \mathcal{I}$, suppose we have chosen $B_\alpha \subseteq \mathcal{I}$ with $\alpha \in B_\alpha$. Define

$$b_1 = \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in B_\alpha} \mathbb{E}[X_\alpha]\, \mathbb{E}[X_\beta],$$

$$b_2 = \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in B_\alpha - \{\alpha\}} \mathbb{E}[X_\alpha X_\beta],$$

$$b_3 = \sum_{\alpha \in \mathcal{I}} \mathbb{E}|\mathbb{E}[X_\alpha - \mathbb{E}[X_\alpha] \mid \{X_\beta : \beta \notin B_\alpha\}]|;$$

for each α in the indexing set, the conditioning is on the sigma field generated by the variables not in the neighborhood of α. The total variation distance between W and Poi (λ) *is bounded in the following way:*

$$d_{TV}(W, \text{Poi}\,(\lambda)) \leq 2(b_1 + b_2 + b_3).$$

Theorem 5.3. *Let $C_{n,k}$ be the k-window statistic in a permutation of $[n]$.*

(a) *For k fixed, we have*

$$C_{n,k} \xrightarrow{\mathcal{D}} \text{Poi}\,(2k+1) \quad \textit{as } n \to \infty.$$

(b) *For $k = k_n \to \infty$, in such a way that it remains* $o(n)$*, we have*

$$\frac{C_{n,k_n} - 2k_n}{\sqrt{k_n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2) \quad \textit{as } n \to \infty.$$

(c) *For $k = k_n \sim \alpha n$, with $\alpha \in (0, \frac{1}{2})$,we have*

$$\frac{C_{n,k_n} - \alpha(2-\alpha)n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \alpha^3 + \frac{8}{3}\alpha^2 - 5\alpha + 2\right) \quad \textit{as } n \to \infty.$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

(d) *For $k = k_n \sim \alpha n$, with $\alpha \in [\frac{1}{2}, 1)$, we have*

$$\frac{C_{n,k_n} - \alpha(2-\alpha)n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \alpha^4 - \frac{8}{3}\alpha^3 + 3\alpha^2 - 2\alpha + \frac{2}{3}\right) \quad \text{as } n \to \infty.$$

(e) *For $k = n - g_n$, with $g_n = \mathrm{o}(n)$ growing faster than $\sqrt{n}$, we have*

$$\frac{C_{n,k_n} - n}{\sqrt{g_n^2/n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \quad \text{as } n \to \infty.$$

(f) *For $k = n - g_n$, with $g_n \sim c\sqrt{n}$, for $c > 0$, we have*

$$C_{n,k_n} - n \xrightarrow{\mathcal{D}} -\operatorname{Poi}(c^2) \quad \text{as } n \to \infty.$$

(g) *For $k = n - g_n$, with $g_n = \mathrm{o}(\sqrt{n})$, we have*

$$C_{n,k_n} - n \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty.$$

Proof. The parts of the theorem are treated by different methods. In certain parts, we use Bolthausen's theorem (Theorem 5.1) or the Arratia–Goldstein–Gordon theorem (Theorem 5.2) specialized to the k-window matrix $\mathbf{A}_n$ described in the paragraph before Section 2.

(a) Let k be a fixed natural number. We start from the representation (1.1) of $C_{n,k}$ as a sum of indicators. The factorial moments of a sum of indicators have a special form [11, Chapter 2]:

$$\mathbb{E}[(C_{n,k})_r] = \sum_{i_1,\dots,i_r \in Q_{n,r}} \mathbb{E}\left[\mathbf{1}_{\{|\pi_{i_1} - i_1| \le k\}} \cdots \mathbf{1}_{\{|\pi_{i_r} - i_r| \le k\}}\right], \tag{5.2}$$

where $Q_{n,r}$ is the set of tuples of r distinct integers from $[n]$. We first establish an upper bound. For any fixed r, let n be large enough to pack r non-intersecting windows centered at $i_1, \dots, i_r$. There are at most $(2k+1)$ placements for each of the chosen indices within its window that give rise to success (value 1) of the corresponding indicator. (We say 'at most' to take into account possible edge effects for the first and the last indices.)

The rest of the elements of $[n]$ can be distributed freely over the remaining positions in $(n-r)!$ ways. That is, all the indicators in the summand will be 1 and their product is 1 with probability at most $(n-r)!\,(2k+1)^r/n!$. If some of the windows intersect, $(2k+1)$ remains as an upper bound on the number of positions for the success of an indicator in the summand. This counting leads to the bound

$$\begin{aligned}
\mathbb{E}[(C_{n,k})_r] &\le \sum_{i_1,\dots,i_r \in Q_{n,r}} \frac{(n-r)!\,(2k+1)^r}{n!} \\
&= \frac{(n-r)!\,(2k+1)^r}{n!}\, n(n-1)\cdots(n-r+1) \\
&= (2k+1)^r.
\end{aligned}$$

Toward a lower bound, consider the set $Q'_{n,r}$ of r-tuples of *well-separated* indices, in which the difference between any two indices in the tuple is at least $2k+1$, and the smallest and largest of these indices are at distance at least k from the edges.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

We break up the sum in (5.2) into two sums, one over $Q'_{n,r}$ and another over $Q_{n,r} - Q'_{n,r}$. We keep only the sum over $Q'_{n,r}$. On the set $Q'_{n,r}$ there are exactly $2k+1$ positions for i_j to fall within its window, for $j = 1, \ldots, r$, and obtain the lower bound

$$\mathbb{E}[(C_{n,k})_r] \geq \sum_{i_1,\ldots,i_r \in Q'_{n,r}} \frac{(n-r)!\,(2k+1)^r}{n!} = \frac{(n-r)!\,(2k+1)^r}{n!}\,|Q'_{n,r}|.$$

The size of $Q'_{n,r}$ can be computed by figuring out the number of ways to place elements in the gaps between the well-separated blocks as follows. Let the number of elements preceding the first window be δ_0, let the distance (number of keys) in the gap between the ith and $(i+1)$th windows be δ_i, and let the number of elements following the rth window be δ_r. We need to solve

$$\delta_0 + \delta_1 + \cdots + \delta_r = n - (2k+1)r$$

in non-negative integers. By a standard stars-and-bars count, we find

$$|Q'_{n,r}| = r!\binom{n-2kr}{r}.$$

We go further with the lower bound to get

$$\mathbb{E}[(C_{n,k})_r] \geq \frac{(n-r)!\,(2k+1)^r}{n!}\,(n-2kr)_r.$$

From the lower and upper bounds, we tightly sandwich the rth factorial moment of the k-window statistic:

$$(2k+1)^r\left(1 - \Theta\left(\frac{1}{n}\right)\right) \leq \mathbb{E}[(C_{n,k})_r] \leq (2k+1)^r. \tag{5.3}$$

Take the limit (as $n \to \infty$), for any fixed k, to find that $\mathbb{E}[(C_{n,k})_r] \to (2k+1)^r$. We know that $(2k+1)^r$ are the factorial moments of Poi $(2k+1)$, and convergence of all factorial moments implies convergence in distribution, while uniquely characterizing the limit. Since the factorial moments converge to those of a Poi(λ) law, convergence in distribution to Poi $(2k+1)$ is implied in the standard moment theorem [5, Theorem 30.2].

(b) We next consider $k = k_n$ growing, but remaining $o(n)$, as $n \to \infty$. The characteristic function of a random variable, V, has a standard expansion in terms of factorial moments (see e.g. [11, Chapter 2]):

$$\phi_V(t) = \sum_{r=0}^{\infty} \mathbb{E}[(V)_r]\,\frac{(e^{it}-1)^r}{r!}.$$

By (5.3), the characteristic function of $C_{n,k}$ is

$$\begin{aligned}\phi_{n,k_n}(t) &\sim \sum_{r=0}^{\infty} (2k_n)^r\,\frac{(e^{it}-1)^r}{r!}\\ &= e^{2k_n(e^{it}-1)}\\ &= e^{2k_n(it - t^2/2 + O(t^3))}.\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

Hence, as $n \to \infty$, we have

$$\phi_{n,k_n}\left(\frac{t}{\sqrt{k_n}}\right)\exp\left(-\frac{2\mathrm{i}k_n t}{\sqrt{k_n}}\right) \to \exp\left(2k_n\left(-\frac{t^2}{2k_n}+\mathrm{O}\left(\frac{t^3}{k_n^{3/2}}\right)\right)\right),$$

from which we conclude that

$$\frac{C_{n,k}-2k_n}{\sqrt{k_n}} \to \mathrm{e}^{-t^2}.$$

The right-hand side in the last convergence relation is the characteristic function of $\mathcal{N}(0, 2)$. By Lévy's continuity theorem [5, Theorem 26.3], part (b) is proved.

(c)–(d) In the linear regime of these two parts, we have $k_n \sim \alpha n$, with $\alpha \in (0, 1)$. Let $\sigma^2_{\mathbf{A}_n} = \mathrm{Var}\,[C_{n,k}]$. By the symmetry of $\mathbf{A}_n$, we have $x_{n,j,k} = x_{n,n-j+1,k}$. With every $a_{i,j}$ being 0 or 1, and $a_{i.} = x_{n,i,k}/n$, $a_{.j} = x_{n,n-j+1,k}/n$, and the proportion of 1s in the matrix all being fractions, each of the four factors in Bolthausen's bound is absolutely bounded by 1. We have the bound

$$\sup_x\left(\left|\mathbb{P}\left(\frac{C_{n,k_n}-\mathbb{E}[C_{n,k}]}{\sigma_{\mathbf{A}_n}}\le x\right)-\Phi(x)\right|\right) \le \frac{K}{n}\sum_{1\le i,j\le n}\left(\frac{4}{\sigma_{\mathbf{A}_n}}\right)^3 = \frac{64Kn}{\sigma^3_{\mathbf{A}_n}}.$$

The variance of C_{n,k_n} in this regime is asymptotic to $\rho(\alpha)n$ (see Proposition 3.1). Asymptotic normality follows. After cleaning up via Slutsky's theorem, we get

$$\frac{C_{n,k_n}-\alpha(2-\alpha)n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,\,\rho(\alpha)) \quad \text{as } n\to\infty;$$

the rate of convergence is $\mathrm{O}(1/\sqrt{n}\,)$. Recall that the linearity coefficient $\rho(\alpha)$ in the variance in the regime $(0, \frac{1}{2})$ is different from the linearity coefficient in the regime $[\frac{1}{2}, 1)$ (see the discussion in Section 3.1). For each part, we use the corresponding $\rho(\alpha)$. Parts (c) and (d) are proved.

(e) Consider next the regime $k_n = n - g_n$, for $g_n = \mathrm{o}(n)$ but growing faster than $\sqrt{n}$. Again, we use Bolthausen's theorem (Theorem 5.1). We examine the geometry of $\mathbf{A}_n$. It has two corners filled with zeros. These are positions excluded from viewing. The top row has $g_n - 1$ excluded cells flushed to its right end, the second row has $g_n - 2$ excluded cells flushed to its right end, and so on until we reach row $g_n - 1$, which has one excluded cell at its rightmost position. There are $(g_n - 1) + (g_n - 2) + \cdots + 1 = g_n(g_n - 1)/2$ excluded cells in this top right corner. The lower left corner is a mirror image of the top right one. So we have exactly $(g_n - 1)g_n$ excluded cells.

To the double sum (5.1) in Bolthausen's theorem, a cell containing $a_{i,j} = 1$ contributes

$$\begin{aligned}
&\left|1-\frac{x_{n,i,k}}{n}-\frac{x_{n,n-j+1,k}}{n}+\frac{n^2-g_n(g_n-1)}{n^2}\right|^3\\
&=\left|1-\frac{n-\mathrm{O}(g_n)}{n}-\frac{n-\mathrm{O}(g_n)}{n}+\frac{n^2-g_n(g_n-1)}{n^2}\right|^3\\
&=\mathrm{O}\left(\frac{g_n^3}{n^3}\right),
\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

whereas a cell containing $a_{i,j} = 0$ contributes

$$\left|0 - \left(1 + \mathrm{O}\left(\frac{g_n}{n}\right)\right) - \left(1 + \mathrm{O}\left(\frac{g_n}{n}\right)\right) + \left(1 - \mathrm{o}\left(\frac{g_n}{n}\right)\right)\right|^3 = 1 + \mathrm{O}\left(\frac{g_n}{n}\right).$$

The double sum (5.1) in Bolthausen's theorem amasses $(n^2 - g_n(g_n - 1))\mathrm{O}(g_n^3/n^3)$ from the cells containing 1 and $g_n(g_{n-1})(1 + \mathrm{O}(g_n/n))$ from the excluded zeros at the corners, a total double sum of order g_n^2. The variance is asymptotic to g_n^2/n and the overall bound is

$$\frac{\mathrm{O}\left(g_n^2\right)}{n\Theta\left(g_n^2/n\right)^{3/2}} = \Theta\left(\frac{\sqrt{n}}{g_n}\right).$$

The bound tends to 0 for the regime $k_n = n - g_n$, with $n = \mathrm{o}(g_n^2)$. Limiting Gaussianity follows. After cleaning up via Slutsky's theorem, part (e) is proved.

(f) In the critical regime $k_n = n - g_n$, with $g_n \sim c\sqrt{n}$, and $c \in \mathbb{R}^+$, we prove a Poisson limit. Let us first focus on $C'_{n,k_n} = n - C_{n,k_n}$, which counts the number of permutation elements lying *outside* their viewing window.

The natural indexing by the set $\{1, \ldots, n\}$ gives bounds on the total variation distance that are too high to be of use. It is then helpful to work with refined indexing $\mathcal{I}$, derived from the representation of $\{|\pi_i - i| > k_n\}$ as a union of individual atoms specified by locations on row i.

On row $i \in \{1, \ldots, g_n - 1\}$, the set $\mathcal{I}$ collects

$$\{(i, i + k + 1)\} \cup \{(i, i + k + 2)\} \cup \cdots \cup \{(i, n)\}.$$

On row $i \in \{n - g_n + 2, \ldots, n\}$, the set $\mathcal{I}$ collects

$$\{(i, 1)\} \cup \{(i, 2)\} \cup \cdots \cup \{(i, i + g_n - n - 1)\}.$$

On rows $i = g_n, \ldots, n - g_n + 1$, the sets $\{|\pi_i - i| > k_n\}$ are empty and add nothing to $\mathcal{I}$. So the indexing set can be written as

$$\mathcal{I} = \left(\bigcup_{i=1}^{g_n - 1} \bigcup_{j=i+k+1}^{n} \{(i, j)\}\right) \bigcup \left(\bigcup_{i=n-g_n+2}^{n} \bigcup_{j=1}^{i+g_n-n-1} \{(i, j)\}\right).$$

The size of the indexing set is $2((g_n - 1) + (g_n - 2) + \cdots + 1) = g_n(g_n - 1)$.

For $(i, j) \in [n] \times [n]$, define $X_{(i,j)} = \mathbf{1}_{\{\pi_i = j\}}$. Now we have

$$C'_{n,k_n} = \sum_{i=1}^{n} \mathbf{1}_{\{|\pi_i - i| > k_n\}} = \sum_{i=1}^{n} \sum_{\alpha=(i,j)\in\mathcal{I}} X_{(i,j)}.$$

After all, with this representation, C'_{n,k_n} is the sum of a large number of Bernoulli random variables, making the case a candidate for Poisson approximation via the Chen–Stein method.

In the Arratia–Goldstein–Gordon version of the Chen–Stein method (Theorem 5.2), take the dependency neighborhood corresponding to $\alpha = (i, j) \in \mathcal{I}$ to be

$$B_\alpha = \{\beta = (i', j') \colon i' = i \text{ or } j' = j\}.$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

In other words, we take X_β to be in the dependency neighborhood of X_α if the cell specified by the pair $\beta = (i', j')$ in the matrix $\mathbf{A}_n$ is in the same row or column as $\alpha = (i, j)$. The size of the dependency neighborhood of $\alpha \in \mathcal{I}$ is at most $g_n + g_n - 1 = 2g_n - 1$.

We next go through the calculation of the three components b_1, b_2, b_3. For b_1, we use that the size of B_α of any $\alpha \in \mathcal{I}$ is at most $2g_n - 1$, and the size of $\mathcal{I}$ is $g_n(g_n - 1)$, and that we are in the regime $g_n \sim c\sqrt{n}$ to get

$$b_1 = \sum_{\alpha \in \mathcal{I}} \frac{|B_\alpha|}{n^2} \leq \frac{|\mathcal{I}|(2g_n - 1)}{n^2} = \Theta\left(\frac{1}{\sqrt{n}}\right).$$

As no two elements can occupy the same location in the permutation, and no element appears in two places in the permutation, if $X_\alpha = 1$, then $X_\beta = 0$ for all $\beta \in B_\alpha - \{\alpha\}$. Otherwise, $X_\alpha = 0$. In both cases $X_\alpha X_\beta = 0$ and $b_2 = 0$.

The computation of b_3 requires conditioning on the variables X_β, with β in the subset $H_\alpha = \{\beta \in \mathcal{I} \colon \beta \notin B_\alpha\}$. By the geometry of the neighborhoods, if $X_\beta = 1$, for $\beta \in H_\alpha$, then this deprives X_α from a position in the permutation, as $\beta \in H_\alpha$ specifies a cell in $\mathbf{A}_n$ in a different row and a different column from α. Therefore, for $\alpha = (i, j) \in \mathcal{I}$, conditionally, there are $n - \sum_{\beta \in H_\alpha} X_\beta$ equally likely positions for i.

Observe that at most one among the Bernoulli random variables in a row within $\mathcal{I}$ can equal 1 (an element of $[n]$ cannot be at two positions in the permutation). In other words, the sum of all the Bernoulli random variables on the same row of $\mathcal{I}$ is one combined Bernoulli random variable. Conditionally, there are

$$n - \sum_{\substack{i'=1 \\ i' \neq i}}^{g_n - 1} \tilde{X}_{i'} - \sum_{\substack{i''=n-g_n+2 \\ i'' \neq i}}^{n} \tilde{X}_{i''},$$

for Bernoulli random variables

$$\tilde{X}_{i'} = \sum_{j=n-g_n+i'+1}^{n} X_{(i',j)} \quad \text{for } i' \in \{1, \ldots, g_n - 1\},$$

and

$$\tilde{X}_{i''} = \sum_{j=1}^{i''-n+g_n-1} X_{(i'',j)} \quad \text{for } i'' \in \{n - g_n + 2, \ldots, n\}.$$

So we have

$$b_3 = \sum_{\alpha=(i,j)\in\mathcal{I}} \mathbb{E}\left[\frac{1}{n - \sum_{\substack{i'=1 \\ i' \neq i}}^{g_n-1} \tilde{X}_{i'} - \sum_{\substack{i''=n-g_n+2 \\ i'' \neq i}}^{n} \tilde{X}_{i''}} - \frac{1}{n}\right]$$

$$= \sum_{\alpha \in \mathcal{I}} \mathbb{E}\left[\frac{\sum_{\substack{i'=1 \\ i' \neq i}}^{g_n-1} \tilde{X}_{i'} + \sum_{\substack{i''=n-g_n+2 \\ i'' \neq i}}^{n} \tilde{X}_{i''}}{n\left(n - \sum_{\substack{i'=1 \\ i' \neq i}}^{g_n-1} \tilde{X}_{i'} - \sum_{\substack{i''=n-g_n+2 \\ i'' \neq i}}^{n} \tilde{X}_{i''}\right)}\right]$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

$$\leq \sum_{\alpha \in \mathcal{I}} \frac{2(g_n - 1)}{n(n - 2(g_n - 1))}$$
$$= \frac{2|\mathcal{I}|\,(g_n - 1)}{n(n - 2(g_n - 1))}$$
$$= \frac{\mathrm{O}(g^3(n))}{\Theta(n^2)}$$
$$= \mathrm{O}\left(\frac{1}{\sqrt{n}}\right).$$

With $b_2 = 0$ and both b_1 and b_3 being of order $1/\sqrt{n}$, in Theorem 5.2, we get to the conclusion that

$$d_{TV}\big(C'_{n,k_n}, \mathrm{Poi}\big(\mathbb{E}\big[C'_{n,k_n}\big]\big)\big) = \mathrm{O}\left(\frac{1}{\sqrt{n}}\right).$$

By the triangle inequality, we have

$$d_{TV}\big(C'_{n,k_n}, \mathrm{Poi}\,(c^2)\big) \leq d_{TV}\big(C'_{n,k_n}, \mathrm{Poi}\big(\mathbb{E}\big[C'_{n,k_n}\big]\big)\big) + d_{TV}\big(\mathrm{Poi}\big(\mathbb{E}\big[C'_{n,k_n}\big], \mathrm{Poi}\,(c^2)\big)\big)$$
$$= \mathrm{O}\left(\frac{1}{\sqrt{n}}\right) + d_{TV}\big(\mathrm{Poi}\big(\mathbb{E}\big[C'_{n,k_n}\big], \mathrm{Poi}\,(c^2)\big).$$

As shown in [12, Lemma 4.2], the total variation distance between two Poisson distributions is at most the absolute difference between their parameters. So we have

$$d_{TV}\big(\mathrm{Poi}\big(\mathbb{E}\big[C'_{n,k_n}\big], \mathrm{Poi}\,(c^2)\big)\big) = \big|n - \mathbb{E}\big[C_{n,k_n}\big] - c^2\big|.$$

The expectation of C_{n,k_n} is given in Proposition 2.1. After a calculation, we find that the latter total variation distance is $\mathrm{O}(1/\sqrt{n}\,)$ too. We have arrived at

$$d_{TV}\big(n - C_{n,k_n}, \mathrm{Poi}\,(c^2)\big) \to 0 \quad \text{as } n \to \infty,$$

which implies that $n - C_{n,k_n} \xrightarrow{\mathcal{D}} \mathrm{Poi}\,(c^2)$. Part (f) is proved.

(g) In the extremal regime $n - g_n$, with $g_n = \mathrm{o}(\sqrt{n}\,)$, the variance is $\mathrm{O}(g_n^2/n)$. For every $\varepsilon > 0$, the Chebyshev inequality gives

$$\mathbb{P}\big(\big|C_{n,k_n} - \mathbb{E}\big[C_{n,k_n}\big]\big| > \varepsilon\big) \leq \frac{\mathrm{O}\left(g_n^2/n\right)}{\varepsilon^2} \to 0 \quad \text{as } n \to \infty,$$

which means $C_{n,k_n} - \mathbb{E}[C_{n,k_n}] \xrightarrow{\mathbb{P}} 0$. It is discussed in Section 2.1 that $\mathbb{E}[C_{n,k_n}] - n \to 0$. The addition rule of probabilistic convergence relations proves part (g). □

6. A retrospective look at the phases

For the purpose of convergence, we kept k fixed in Theorem 5.3(a). In the case of k growing, the bounds in (5.3) remain valid, but it is crucial to keep k_n in the low $\mathrm{o}(n)$ regime. The reason is that for the argument of the lower bound, for any fixed r, however large, we need to be able

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

to pack r non-intersecting windows of width $2k_n + 1$ over n positions to construct Q'_{n,k_n}. This requires $r(2k_n + 1) \le n$. For r arbitrarily large, we need $r \le n/(2k_n + 1)$ for $n \to \infty$. Thus n/k_n must diverge to infinity. Without the stipulation that $k_n = o(n)$, the argument breaks down for not keeping the lower bound of the same order as the upper bound.

A case in point is $k_n = n - 1$, for which the product of the indicators is a Bernoulli random variable with success probability 1 (every element is seen within its associated window, $C_{n,n-1} = n$) and the rth factorial moment of $C_{n,n-1}$ is $(n)_r$, for $r \ge 2$, not bounded below by $(2(n-1)+1)^r$.

For $k_n = o(n)$, we can approximate the distribution of the k_n-window statistic by that of Poi $(2k_n + 1)$, which in turn is normal in view of the size of its parameter, as in part (b) of Theorem 5.3, for example.

Chen–Stein Poisson approximation could be used to prove part (a). It is also possible to use the method in the regime $k_n = n - g_n$, with $g_n = o(\sqrt{n})$. However, in this case we get convergence of $C_{n,k_n} - n$ to Poi (0), which would look a bit artificial. In any case, such convergence to Poi (0) is identical to the result in part (g).

In the Arratia–Goldstein–Gordon version [1] of the Chen–Stein method, the authors speak of dependency neighborhood sets B_α of X_α, outside which the variables X_β, for $\beta \notin B_\alpha$, are independent or 'nearly independent' of X_α. In practice, such neighborhoods are not unique. The author ultimately selected B_α as in the proof of Theorem 5.3(f) to keep the argument in the spirit of Example 1.3 in [2] and create a context for comparison.

Remark 6.1. The distribution of the 0-window statistic (the number of fixed points) has long been known to be Poi (1) (see [4, Example 3.4.2]). This agrees with the statement of Theorem 5.3(a), when $k = 0$.

Remark 6.2. The convergence in part (f) to Poi (c^2) is consistent with the computations of the mean and variance in Sections 2.1 and 3.1: the Poison distribution Poi (c^2) has mean c^2 and variance c^2, conforming to $\lim_{n\to\infty} \mathbb{E}[C_{n,k}] = c^2$ (from Proposition 2.1) and $\lim_{n\to\infty}$ Var $[C_{n,k}] = c^2$ (from Proposition 3.1).

7. A note on the use of AI

The author attempted to use AI and it was a very mixed experience, often frustrating, but there were also bright moments with happy endings. AI cannot construct mathematical logic and often makes gross mistakes. Nevertheless, on occasion it suggested references and useful directions, even though it could not execute them well.

The specific AI consulted is ChatGPT-5.2. One example of where AI was misleading is as follows. The author asked for references on the Poisson limit for fixed points. ChatGPT claimed it is in Feller's book [7], specifying Section 6.2. Upon checking Section 6.2, one cannot find a discussion of fixed points in permutations; rather, one finds a discussion of derangements, which count something different. The author went back and pressed ChatGPT for better answers. The chat became defensive and evasive, claiming that there are multiple editions of the book and possibly the discussion is in the Indian edition. This was a waste of time with no good outcome.

As an example of lack of technical ability, the author consulted ChatGPT on the dependency neighborhoods in the Chen–Stein method. The author did not have extensive experience with the Chen–Stein method when applied to dependent random variables. The announcement in [2] that the complement of the dependency neighborhoods includes independent or 'nearly independent' random variables left the author baffled. He then asked ChatGPT for clarification.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

The answer was reasonable: these neighborhoods are not unique. The author then chose the one in the proof of Theorem 5.3(f) which works. He then wondered if a simpler neighborhood like just the set $\{\alpha\}$ might work. ChatGPT concurred that it is a 'risky' dependency neighborhood. The author then asked ChatGPT to go through the calculation with this reduced neighborhood. It came back concluding that it does not work. After hand checking, the author found that indeed b_1, b_2, and b_3 converge to 0. Chen–Stein Poisson approximation goes through with this choice; the chat's answer is false. This gross error was reported to ChatGPT. It acknowledged its mistake and apologized ☺.

However, ChatGPT was helpful in resolving multiple technical writing issues that would otherwise consume an enormous amount of time to investigate and find the right sources. As an example, the author impressed on the AI that the scientific community prefers for citations to books and long papers to specify a location, such as a range of pages or a section. ChatGPT concurs that this is a better style. It was then asked how to produce such a citation. ChatGPT gave a useful tip on this, providing the LaTeX code

\cite[Example 3.4.2]{Holst}

for the citation [4, Example 3.4.2] used in Section 1.

The author is sharing this information in the hope that it will help other researchers, particularly junior researchers, understand the scope, limitations, pitfalls, and areas of strength of AI. It is also hoped that it will help AI developers improve their designs.

Appendix

We expand the four-fold sums in $L_{n,k}$ to write it in the form

$$\begin{aligned}
L_{n,k} &= \sum_{1\le i,j\le n}\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\sum_{\substack{s=1\\ s\neq r}}^{n}\mathbf{J}_{s,j,k}\\
&= \sum_{1\le i,j\le n}\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\left(\left(\sum_{s=1}^{n}\mathbf{J}_{s,j,k}\right)-\mathbf{J}_{r,j,k}\right)\\
&= \sum_{1\le i,j\le n}\left(\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\right)\left(\sum_{s=1}^{n}\mathbf{J}_{s,j,k}\right)-\sum_{1\le i,j\le n}\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\,\mathbf{J}_{r,j,k}\\
&= \sum_{1\le i,j\le n}\left(\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\right)\left(\sum_{s=1}^{n}\mathbf{J}_{s,j,k}\right)-\sum_{1\le i,j\le n}\sum_{r=1}^{n}\mathbf{J}_{i,r,k}\,\mathbf{J}_{j,r,k} \quad (\mathbf{A}_n \text{ is symmetric})\\
&= \sum_{1\le i,j\le n}\left(\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\right)\left(\sum_{s=1}^{n}\mathbf{J}_{s,j,k}\right)-\sum_{r=1}^{n}\sum_{i=1}^{n}\mathbf{J}_{i,r,k}\sum_{j=1}^{n}\mathbf{J}_{j,r,k}.
\end{aligned}$$

Similarly, we have

$$M_{n,k}=\sum_{i=1}^{n}\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\sum_{\substack{s=1\\ s\neq r}}^{n}\mathbf{J}_{s,i,k}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

$$= \sum_{i=1}^{n}\sum_{r=1}^{n} \mathbf{J}_{r,i,k}\left(\left(\sum_{s=1}^{n}\mathbf{J}_{s,i,k}\right) - \mathbf{J}_{r,i,k}\right)$$
$$= \sum_{i=1}^{n}\left(\sum_{r=1}^{n}\mathbf{J}_{r,i,k}\right)\left(\sum_{s=1}^{n}\mathbf{J}_{s,i,k}\right) - \sum_{i=1}^{n}\sum_{r=1}^{n}\mathbf{J}_{r,i,k}.$$

Again, the sum $x_{n,i,k} = \sum_{r=1}^{n}\mathbf{J}_{r,i,k}$ is key to these calculations. In the proof of Proposition 2.1, we have worked out the structure of $x_{n,i,k}$ in both the lower regime $0 \le k < n/2$ and the upper regime $n/2 \le k$. In these terms, we write

$$L_{n,k} = \left(\sum_{i=1}^{n} x_{n,i,k}\right)^2 - \sum_{i=1}^{n} x_{n,i,k}^2 \tag{A.1}$$

and

$$M_{n,k} = \sum_{i=1}^{n} x_{n,i,k}^2 - \sum_{i=1}^{n} x_{n,i,k}. \tag{A.2}$$

The sum $\sum_{i=1}^{n} x_{n,i,k}$ is evaluated in (2.1) and the formula for the regime $0 \le k < n/2$ is the same as for the regime $n/2 \le k \le n-2$. Unlike this situation, the sum $\sum_{i=1}^{n} x_{n,i,k}^2$ is different in the two regimes, and we have to consider the two cases separately.

The regime $k < n/2$. In this regime we have

$$\sum_{i=1}^{n} x_{n,i,k}^2 = \sum_{i=1}^{k}(k+i)^2 + \sum_{i=k+1}^{n-k}(2k+1)^2 + \sum_{i=n-k+1}^{n}(k+1+n-i)^2$$
$$= (2k+1)\left((2k+1)n - \frac{5}{3}k(k+1)\right). \tag{A.3}$$

We plug (2.1) and (A.3) into (A.1) to find

$$L_{n,k} = (2k+1)((2k+1)n^2 - (2k^2+4k+1)n) + \frac{1}{3}k(k+1)(3k^2+13k+5).$$

We finish the computation by substituting (2.1) and (A.3) in (A.2), which gives

$$M_{n,k} = 2k(2k+1)n - \frac{2}{3}k(k+1)(5k+1).$$

The regime $k \ge n/2$. The sum $\sum_{i=1}^{n} x_{n,i,k}$ is the same expression as the counterpart in the case $k < n/2$.

Again, $\sum_{i=1}^{n} x_{n,i,k}^2$ does not come out the same as its counterpart in the case $k < n/2$. Here we have

$$\sum_{i=1}^{n} x_{n,i,k}^2 = \sum_{i=1}^{n-k-1}(k+i)^2 + \sum_{i=n-k}^{k+1} n^2 + \sum_{i=k+2}^{n}(k+n+1-i)^2$$
$$= \frac{1}{3}(-n^3 + 3(2k+1)n^2 + n - k(k+1)(2k+1)). \tag{A.4}$$

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press

Using (2.1) and (A.4) in (A.1), we get

$$L_{n,k} = \frac{1}{3}n^3 + 2k(2k+1)n^2 - \left(4k^3 + 6k^2 + 2k + \frac{1}{3}\right)n + \frac{1}{3}k(k+1)(3k^2+5k+1).$$

We finish the computation by substituting (2.1) and (A.4) in (A.2), which gives

$$M_{n,k} = -\frac{n^3}{3} + (2k+1)n^2 - \left(2k + \frac{2}{3}\right)n - \frac{2k}{3}(k^2-1).$$

Acknowledgements

The author is indebted to Svante Janson for encouragement and sound advice directing him to lines of attack that proved quite effective in this research. The careful critique by two anonymous referees improved the exposition and rid it of some small errors.

This work was completed while the author was on sabbatical leave at The Catholic University of America in Fall 2025. He thanks the Department of Mathematics and Statistics there for their gracious reception and hospitality, and for an environment conducive to research under the leadership of Kiran Bhutani.

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ARRATIA, R., GOLDSTEIN, L. AND GORDON, L. (1989). Two moments suffice for Poisson approximations: The Chen–Stein method. *Ann. Prob.* **17**, 9–25.
[2] ARRATIA, R., GOLDSTEIN, L. AND GORDON, L. (1990). Poisson approximation and the Chen–Stein method. *Statist. Sci.* **5**, 403–424.
[3] BARBOUR, A. AND JANSON, S. (2009). A functional combinatorial central limit theorem. *Electron. J. Prob.* **14**, 2352–2370.
[4] BARBOUR, A., HOLST, L. AND JANSON, S. (1992). *Poisson Approximation*. Clarendon Press, Oxford.
[5] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd edn. Wiley, New York, NY.
[6] BOLTHAUSEN, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrscheinlichkeitsth.* **66**, 379–386.
[7] FELLER, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd edn. John Wiley, New York.
[8] FLAJOLET, P. AND SEDGEWICK, R. (2009). *Analytic Combinatorics*. Cambridge University Press.
[9] GRAHAM, R. L., KNUTH, D. E. AND PATASHNIK, O. (1994). *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, Reading, MA.
[10] HOEFFDING, W. (1951). A combinatorial central limit theorem. *Ann. Math. Statist.* **22**, 558–566.
[11] JOHNSON, N. L., KEMP, A. W. AND KOTZ, S. (2005). *Univariate Discrete Distributions*. John Wiley, Hoboken, NJ.
[12] LYON, M. R. AND MAHMOUD, H. M. (2020). Trees grown under young-age preferential attachment. *J. Appl. Prob.* **57**, 911–927.

https://doi.org/10.1017/jpr.2026.10125 Published online by Cambridge University Press