The number of occurrences of patterns and constrained patterns in a random permutation

Svante Janson

BIRS, Banff (virtually, alas)
23 September, 2021

Patterns in a permutation

Let $\mathfrak{S}_n$ be the set of permutations of $[n] := \{1, \ldots, n\}$.

If $\sigma = \sigma_1 \cdots \sigma_k \in \mathfrak{S}_k$ and $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$, then an *occurrence* of σ in π is a subsequence $\pi_{i_1} \cdots \pi_{i_k}$, with $1 \leq i_1 < \cdots < i_k \leq n$, that has the same relative order as σ. σ is called a *pattern*.

Example: $\underline{3}14\underline{25}$ is an occurence of 213 in 31425

Let $\mathrm{occ}_\sigma(\pi)$ be the number of occurrences of σ in π.
For example, $\mathrm{occ}_{21}(\pi)$ is the number of inversions in π.

A permutation π *avoids* a pattern σ if there is no occurence of σ in π, i.e., if $\mathrm{occ}_\sigma(\pi) = 0$.

General problem

Let $\boldsymbol{\pi}$ be a random permutation, drawn uniformly from all permutations with a given (large) length n in some given class of permutations.

Let σ be a fixed (small) permutation.

Problem

Study the random variable $\mathrm{occ}_\sigma(\boldsymbol{\pi})$.
In particular, find its asymptotic distribution as $n \to \infty$.

Example

Take $\sigma = 21$.
What is the asymptotic distribution of the number of inversions in a random $\boldsymbol{\pi}$?

Remark

First order properties of $\mathrm{occ}_\sigma(\boldsymbol{\pi})$ are closely connected with permuton limits. (In particular, when there is convergence in distrstribution to a random permuton limit.)

Today's talk will not discuss this. We will look at some cases where $\mathrm{occ}_\sigma(\boldsymbol{\pi})$ is concentrated and study second order properties, more precisely we show asymptotically normal fluctuations around the mean for these cases. (Permuton limits are trivial in the examples discussed today.)

The class of permutations considered here is often (but not always) a *pattern class*, i.e., the class of all permutations that avoid one or several given patterns.

Examples

312 Knuth, *The Art of Computer Programming, vol. 1*
321 Tarjan (1972)
$\{2431, 4231\}$ West (1995)
321; 312; 231; 132; $\{2413, 3142\}$; $\{1342, 1324\}$; $\{4231, 3412\}$;
1342 Stanley, *Enumerative Combinatorics*, Exercises 6.19 x, y, ee, ff, ii, oo, xx; 6.25 g; 6.39 k, l, m; 6.47 a; 6.48.
$\{2413, 3142\}$ (separable permutations) Bassino, Bouvel, Féray, Gerin, Pierrot (2018).

Bóna (2004)

Remark

Many other properties of random permutations from a pattern class have been studied by a number of authors. For example:

consecutive patterns, descents, major index, number of fixed points, position of fixed points, exceedances, longest increasing subsequence, shape and distribution of individual values π_i.

For $\mathrm{occ}_\sigma(\boldsymbol{\pi})$, many permutation classes have been treated, by myself and others.

(I was convinced by Igor Pak that no general results for all pattern classes are possible.)

Today only a few that can be treated by a simple method: *U-statistics* (with some twists).

Unrestricted permutations

As a background, consider random permutations without restrictions.

Theorem (Bóna (2007, 2010), Janson, Nakamura and Zeilberger (2015))

Consider a random permutation $\boldsymbol{\pi}_n \in \mathfrak{S}_n$. Then $\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)$ is asymptotically normally distributed, for any σ: if $k := |\sigma|$ then

$$\frac{\operatorname{occ}_\sigma(\boldsymbol{\pi}_n) - n^k/k!^2}{n^{k-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma_\sigma^2)$$

for some constant $\gamma_\sigma > 0$.

Proof.

A random permutation $\boldsymbol{\pi}_n$ can be obtained by taking i.i.d. random variables $X_1, \ldots, X_n \sim U(0,1)$ and considering their ranks. Then

$$\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) = \sum_{i_1 < \cdots < i_m} f\left(X_{i_1}, \ldots, X_{i_m}\right)$$

for a suitable (indicator) function f.

This sum is a U-statistic, and the result follows by general results Hoeffding (1948). $\square$

Example

The number of inversions is

$$\mathrm{occ}_{21}(\boldsymbol{\pi}_n) = \sum_{i_1 < i_2} \mathbf{1}\{X_{i_1} > X_{i_2}\}.$$

U-statistics

A *U-statistic* is a sum

$$S_n = S_n(f) = \sum_{i_1 < \cdots < i_m} f(X_{i_1}, \ldots, X_{i_m})$$

where $X_1, \ldots, X_n$ is an i.i.d. sequence of random variables. and f is a measurable function.

X_i may take values in any measurable space. For example, X_i may be real-valued, vectors, or random permutations.

Traditionally (Hoeffding, 1948), f is supposed to be symmetric (equivalently, the sum is taken over all distinct $i_1, \ldots, i_m$). In combinatorics, I usually need the asymmetric version above.

The asymmetric case can be reduced to the symmetric as follows: Let $Y_1, \dots, Y_n$ be uniform random variables on $[0,1]$, independent of (X_i) and each other, and define $Z_i := (X_i, Y_i)$. Let

$$F(Z_1, ..., Z_m) := \sum_{\pi \in \mathfrak{S}_m} f(X_{\pi(1)}, \dots, X_{\pi(m)}) \mathbf{1}\{Y_{\pi(1)} < \cdots < Y_{\pi(m)}\}$$

Then $S_n(F)$ is a symmetric U-statistic, and

$$S_n(F) \stackrel{\mathrm{d}}{=} S_n(f).$$

Theorem (Hoeffding, 1948)

Let $\mathbb{E}\,|f(X_1,\ldots,X_m)|^2 < \infty$. *Then*

$$\frac{S_n - \binom{n}{m}\mu}{n^{m-1/2}} \xrightarrow{\mathrm{d}} N(0,\gamma^2),$$

where

$$\mu = \mathbb{E}\,f(X_1,\ldots,X_m)$$

and

$$\gamma^2 \geq 0.$$

(Explicit formula, but omitted today.)

Degenerate cases

If $\gamma^2 = 0$, then we get non-normal limits with another normalization. Typically an infinite sum of squares of normal variables. (Higher degeneracies lead to higher-degree polynomials.)

Such cases typically do not occur, but they are easily constructed by taking linear combinations.

Example

$$\mathrm{occ}_{123}(\boldsymbol{\pi})+\mathrm{occ}_{231}(\boldsymbol{\pi})+\mathrm{occ}_{312}(\boldsymbol{\pi})-\mathrm{occ}_{132}(\boldsymbol{\pi})-\mathrm{occ}_{213}(\boldsymbol{\pi})-\mathrm{occ}_{321}(\boldsymbol{\pi}).$$

In fact, the space of non-trivial linear combinations of $\mathrm{occ}_\sigma(\boldsymbol{\pi})$, $\sigma \in \mathfrak{S}_k$, has dimension $k! - 1$. The space of normal limits has dimension $(k-1)^2$, so the space of degenerate linear combinations has dimension $k! - 1 - (k-1)^2$. (See further Even-Zohar, 2020.)

Hoeffding's proof

Hoeffding's proof is based on a projection method:
Assume $\mathbb{E}\, F(X_1, \ldots, X_m) = 0$. Define

$$f_i(X_i) = \mathbb{E}\big[f(X_1, \ldots, X_m) \mid X_i\big].$$

Approximate $f(X_1, \ldots, X_m)$ by $f_1(X_1) + \cdots + f_m(X_m)$. The resulting sum is, at least in the symmetric case, asymptotically normal by the standard central limit theorem.
The error has small variance and can be ignored.
QED

$$\gamma^2 = 0 \iff f_i(X_i) = 0 \quad \text{a.s.} \quad \text{for every } i = 1, \ldots, m.$$

Variations: vincular patterns

A *vincular pattern* is a pattern where some entries are marked, and we only count occurrences where a marked entry is adjacent to the next one.

Example

The vincular pattern 2^*13 counts triples $(i, i+1, j)$ with $i+1 < j$ and $\pi_{i+1} < \pi_i < \pi_j$.

In particular, marking every element means that we count only substrings (consecutive patterns) $\pi_i \pi_{i+1} \cdots \pi_{i+m-1}$ that have the right order. (Bóna 2010).

Variations: vincular patterns

A *vincular pattern* is a pattern where some entries are marked, and we only count occurrences where a marked entry is adjacent to the next one.

Example

The vincular pattern 2^*13 counts triples $(i, i+1, j)$ with $i+1 < j$ and $\pi_{i+1} < \pi_i < \pi_j$.

In particular, marking every element means that we count only substrings (consecutive patterns) $\pi_i \pi_{i+1} \cdots \pi_{i+m-1}$ that have the right order. (Bóna 2010).

More general constraints: gaps at most d, or exactly d.

The marks in a vincular pattern in σ group the entries in *blocks*. (Example: 2^*13 has the two blocks 2^*1 and 3.)

Theorem (Hofer, 2016)

For any vincular pattern σ with b blocks,

$$\frac{S_n - \frac{n^b}{b!}\mu}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2),$$

where $\gamma^2 > 0$ unless σ has length 1.

My proof.

Let again $(X_i)_1^n$ be i.i.d. uniform $U(0,1)$ random variables, and let $\boldsymbol{\pi} \in \mathfrak{S}_n$ be the corresponding permutation.
In the example 2*13 above, define $Y_i := (X_i, X_{i+1}) \in \mathbb{R}^2$. Then

$$\operatorname{occ}_\sigma(\boldsymbol{\pi}) = \sum_{i,j:i+1<j} f(Y_i, Y_j)$$

for a suitable f. This is, up to a negligible error (viz., terms with $j = i + 1$), a U-statistic of order 2 based on (Y_i).

$\square$

My proof.

Let again $(X_i)_1^n$ be i.i.d. uniform $U(0,1)$ random variables, and let $\pi \in \mathfrak{S}_n$ be the corresponding permutation.
In the example 2*13 above, define $Y_i := (X_i, X_{i+1}) \in \mathbb{R}^2$. Then

$$\mathrm{occ}_\sigma(\pi) = \sum_{i,j:i+1<j} f(Y_i, Y_j)$$

for a suitable f. This is, up to a negligible error (viz., terms with $j = i+1$), a U-statistic of order 2 based on (Y_i).
However, the sequence (Y_i) is not i.i.d. !

$\square$

My proof.

Let again $(X_i)_1^n$ be i.i.d. uniform $U(0,1)$ random variables, and let $\pi \in \mathfrak{S}_n$ be the corresponding permutation.

In the example 2*13 above, define $Y_i := (X_i, X_{i+1}) \in \mathbb{R}^2$. Then

$$\operatorname{occ}_\sigma(\pi) = \sum_{i,j:i+1<j} f(Y_i, Y_j)$$

for a suitable f. This is, up to a negligible error (viz., terms with $j = i+1$), a U-statistic of order 2 based on (Y_i).

However, the sequence (Y_i) is not i.i.d. !

No problem!

$\square$

My proof.

Let again $(X_i)_1^n$ be i.i.d. uniform $U(0,1)$ random variables, and let $\pi \in \mathfrak{S}_n$ be the corresponding permutation.
In the example 2*13 above, define $Y_i := (X_i, X_{i+1}) \in \mathbb{R}^2$. Then

$$\mathrm{occ}_\sigma(\pi) = \sum_{i,j: i+1<j} f(Y_i, Y_j)$$

for a suitable f. This is, up to a negligible error (viz., terms with $j = i+1$), a U-statistic of order 2 based on (Y_i).
However, the sequence (Y_i) is not i.i.d. !

No problem!
The sequence is 1-dependent, and this is enough for the central limit theorem (Orey, 1958), and Hoeffding's proof can be modified. (Janson, 2021)
(In general m-dependence is enough.) $\square$

Degenerate cases

New possibilities for degeneracy with vincular patterns.

Example

$$\mathrm{occ}_{1*3*2}(\boldsymbol{\pi}) + \mathrm{occ}_{2*3*1}(\boldsymbol{\pi}) - \mathrm{occ}_{2*1*3}(\boldsymbol{\pi}) - \mathrm{occ}_{3*1*2}(\boldsymbol{\pi}) \in \{0, \pm 1\}.$$

Not completely explored!

I have found a condition for non-degeneracy, which for example shows that for a single σ of length ≥ 2, the asymptotic variance $\gamma^2 > 0$ as claimed above.

Other permutation classes

The trick is to find an encoding of the permutations in a given class such that the number of occurrences of some pattern σ can be written as a U-statistic.

Other permutation classes

The trick is to find an encoding of the permutations in a given class such that the number of occurrences of some pattern σ can be written as a U-statistic.

Possible only for some permutation classes!

Other permutation classes

The trick is to find an encoding of the permutations in a given class such that the number of occurrences of some pattern σ can be written as a U-statistic.

Possible only for some permutation classes!

Let $\mathfrak{S}_n(\tau_1, \dots, \tau_k)$ denote the set of permutations in $\mathfrak{S}_n$ that avoid $\tau_1, \dots, \tau_k$.

Avoiding $\{132, 312\}$

Theorem

Let $m \geq 2$ and $\sigma \in \mathfrak{S}_m(132, 312)$. If $\boldsymbol{\pi}_n$ is random in $\mathfrak{S}_n(132, 312)$. then as $n \to \infty$,

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) - 2^{1-m} n^m / m!}{n^{m-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2).$$

Proof.

A permutation π belongs to the class $\mathfrak{S}_*(132, 312)$ if and only if every entry π_i is either a maximum or a minimum.
[Simion and Schmidt, 1995].
Encode $\pi \in \mathfrak{S}_n(132, 312)$ by a sequence $\xi_2, \ldots, \xi_n \in \{\pm 1\}^{n-1}$, where $\xi_j = 1$ if π_j is a maximum in π, and $\xi_j = -1$ if π_j is a minimum. This is a bijection. Hence the code for a uniformly random $\boldsymbol{\pi}_n$ has $\xi_2, \ldots, \xi_n$ i.i.d. with $\mathbb{P}(\xi_j = 1) = \mathbb{P}(\xi_j = -1) = \frac{1}{2}$.

Let $\sigma \in \mathfrak{S}_m(132, 312)$ have the code $\eta_2, \ldots, \eta_m$. Then $\pi_{i_1} \cdots \pi_{i_m}$ is an occurrence of σ in π if and only if $\xi_{i_j} = \eta_j$ for $2 \leq j \leq m$. Consequently, $\mathrm{occ}_\sigma(\boldsymbol{\pi}_n)$ is a U-statistic

$$\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) = \sum_{i_1 < \cdots < i_m} f\left(\xi_{i_1}, \ldots, \xi_{i_m}\right),$$

where

$$f\left(\xi_1, \ldots, \xi_m\right) := \prod_{j=2}^{m} \mathbf{1}[\xi_j = \eta_j].$$

The result follows from Hoeffding's theorem. □

Example

For the number of inversions, we have $\sigma = 21$ and $m = 2$, $\eta_2 = -1$. A calculation yields $\mu = \frac{1}{2}$ and $\gamma^2 = \frac{1}{12}$, and thus

$$\frac{\mathrm{occ}_{21}(\pi_n) - n^2/4}{n^{3/2}} \xrightarrow{\mathrm{d}} N\left(0, \tfrac{1}{12}\right),$$

Block decompositions of permutations

If $\sigma \in \mathfrak{S}_m$ and $\tau \in \mathfrak{S}_n$, their (direct) *sum* $\sigma \oplus \tau \in \mathfrak{S}_{m+n}$ is defined by letting τ act on $[m+1, m+n]$ in the natural way; more formally, $\sigma \oplus \tau = \pi \in \mathfrak{S}_{m+n}$ where $\pi_i = \sigma_i$ for $1 \leq i \leq m$, and $\pi_{j+m} = \tau_j + m$ for $1 \leq j \leq n$.

A permutation $\pi \in \mathfrak{S}_*$ is *decomposable* if $\pi = \sigma \oplus \tau$ for some $\sigma, \tau \in \mathfrak{S}_*$, and *indecomposable* otherwise; we also call an indecomposable permutation a *block*.

It is easy to see that any permutation $\pi \in \mathfrak{S}_*$ has a unique decomposition $\pi = \pi_1 \oplus \cdots \oplus \pi_\ell$ into indecomposable permutations (blocks) $\pi_1, \ldots, \pi_\ell$; we call these the *blocks of* π.

$\{231, 312\}$-avoiding permutations

Theorem

Let $\sigma \in \mathfrak{S}_*(231, 312)$ have b blocks. Then, for a random $\boldsymbol{\pi}_n \in \mathfrak{S}_n(231, 312)$,

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) - n^b/b!}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2)$$

for some constant γ^2.

Example The number of inversions.

$$\frac{\mathrm{occ}_{21}(\boldsymbol{\pi}_n) - n}{n^{1/2}} \xrightarrow{\mathrm{d}} N(0, 6).$$

Proof.

- $\pi \in \mathfrak{S}_*(231, 312) \iff$ each block is decreasing: $\ell(\ell-1)\cdots 21$ [Simion and Schmidt, 1995].
- If the block lengths of $\boldsymbol{\pi}_n$ are $\ell_1, \dots, \ell_m$, and the block lengths of σ are $s_1, \dots, s_b$, then

$$\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) = \sum_{i_1 < \cdots < i_b} \prod_{j=1}^{b} \binom{\ell_{i_j}}{s_j}.$$

- If the block lengths of $\boldsymbol{\pi}_n$ are $\ell_1, \dots, \ell_m$, then $\sum_i \ell_i = n$, and $(\ell_1, \dots, \ell_m)$ is a uniformly random composition of n. Thus, the block lengths $\ell_1, \dots, \ell_m$ can be realized as the first elements, up to sum n, of an i.i.d. sequence $L_1, L_2, \dots$ of random variables with a Geometric $\mathrm{Ge}(1/2)$ distribution. I.e., define $N(n) := \max\{k : \sum_1^k L_i \geq n\}$. Then the block lengths can be taken as $(L_1, \dots, L_{N(n)})$ (with the last term truncated if necessary).

- Hence, up to a negligble error (from the last block),

$$\mathrm{occ}_\sigma(\pi_n) = \sum_{1 \le i_1 < \cdots < i_b \le N(n)} \prod_{j=1}^{b} \binom{L_{i_j}}{s_j}.$$

This is a U-statistic, based on the i.i.d. sequence (L_i).

But the sum is up to the random $N(n)$ and not to a fixed n.

- Hence, up to a negligble error (from the last block),

$$\mathrm{occ}_\sigma(\pi_n) = \sum_{1 \le i_1 < \cdots < i_b \le N(n)} \prod_{j=1}^{b} \binom{L_{i_j}}{s_j}.$$

  This is a U-statistic, based on the i.i.d. sequence (L_i).

  But the sum is up to the random $N(n)$ and not to a fixed n.
- No problem!

- Hence, up to a negligble error (from the last block),

$$\mathrm{occ}_\sigma(\pi_n) = \sum_{1 \le i_1 < \cdots < i_b \le N(n)} \prod_{j=1}^{b} \binom{L_{i_j}}{s_j}.$$

  This is a U-statistic, based on the i.i.d. sequence (L_i).

  But the sum is up to the random $N(n)$ and not to a fixed n.
- No problem!
  Renewal theory shows that Hoeffding's proof can be adapted.
  (Janson, 2018)

$\{231, 312, 321\}$-avoiding permutations

Theorem

Let $\sigma \in \mathfrak{S}_*(231, 312, 321)$ *have* b *blocks. Then, for a random* $\boldsymbol{\pi}_n \in \mathfrak{S}_n(231, 312, 321)$*,*

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) - \mu n^b}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2)$$

for some constants μ, γ*.*

Example The number of inversions. $\sigma = 21$. $b = 1$. A calculation yields $\mu = (3 - \sqrt{5})/2$ and $\gamma^2 = 5^{-3/2}$.

$$\frac{\mathrm{occ}_{21}(\boldsymbol{\pi}_n) - \frac{3-\sqrt{5}}{2} n}{n^{1/2}} \xrightarrow{\mathrm{d}} N\big(0, 5^{-3/2}\big).$$

Proof.

- $\pi \in \mathfrak{S}_*(231, 312, 321) \iff$ each block is of the type 1 or 21. [Simion and Schmidt, 1995].
- Thus π is determined by its sequence of block lengths $\ell_1, \dots, \ell_m$ with $\ell_i \in \{1, 2\}$ and $\sum_i \ell_i = n$.
- Let $p := (\sqrt{5} - 1)/2$, the golden ratio, so that $p + p^2 = 1$. Let $X_1, X_2, \dots$ be an i.i.d. sequence of random variables with

$$\mathbb{P}(X_i = 1) = p, \qquad \mathbb{P}(X_i = 2) = p^2.$$

Let $S_k := \sum_{i=1}^{k} X_i$ and $N(n) := \min\{k : S_k \geq n\}$. Then, the sequence $\ell_1, \dots, \ell_B$ of block lengths of a uniformly random permutation $\boldsymbol{\pi}_n \in \mathfrak{S}_*(231, 312, 321)$ has the same distribution as $(X_1, \dots, X_{N(n)})$ conditioned on $S_{N(n)} = n$.
Consequently, $\mathrm{occ}_\sigma(\boldsymbol{\pi}_n)$ can be expressed as a U-statistic based on $X_1, \dots, X_{N(n)}$, conditioned as above.

- This is almost as in the preceding case.

  But in this case, we also condition on the event $S_{N(n)} = n$, i.e., that some sum S_k exactly equals n.

- This is almost as in the preceding case.

  But in this case, we also condition on the event $S_{N(n)} = n$, i.e., that some sum S_k exactly equals n.
- No problem!

- This is almost as in the preceding case.

  But in this case, we also condition on the event $S_{N(n)} = n$, i.e., that some sum S_k exactly equals n.

- No problem!

  More renewal theory shows that Hoeffding's proof can be adapted to this case too. (Janson, 2018)

Forest permutations $= \{321, 3412\}$-avoiding

If π is a permutation of $[n]$, then its *permutation graph* G_π is the graph with an edge ij for each inversion (i, j) in π.

Acan and Hitczenko (2016) define π to be a *tree permutation* [*forest permutation*] if G_π is a tree [forest].

$$\{\text{forest permutations}\} = \mathfrak{S}(321, 3412).$$

A permutation is a forest permutation $\iff$ every block is a tree permutation.

Define a random tree permutation (of random length) $\boldsymbol{\tau}$ such that, for every tree permutation σ,

$$\mathbb{P}(\boldsymbol{\tau} = \sigma) = p^{|\sigma|},$$

with $p = (3 - \sqrt{5})/2$ chosen such that $\sum_\sigma \mathbb{P}(\boldsymbol{\tau} = \sigma) = 1$.

Let $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \ldots,$ be i.i.d. random tree permutations with this distribution. Let $S_k := \sum_{i=1}^{k} |\boldsymbol{\tau}_k|$, the total length of the k first, and let $N(n) := \min\{k : S_k \geq n\}$. Then, conditioned on $S_{N(n)} = n$, the sum $\boldsymbol{\pi} := \boldsymbol{\tau}_1 \oplus \cdots \oplus \boldsymbol{\tau}_{N(n)}$ is a uniformly distributed forest permutation of length n.

Let $\sigma = \sigma_1 \oplus \ldots, \oplus \sigma_b$ be a forest permutation, decomposed into tree permutations σ_i. Then, up to a small error,

$$\mathrm{occ}_\sigma(\boldsymbol{\pi}) = \sum_{i_1 < \cdots < i_b} \prod_{j=1}^{b} \mathrm{occ}_{\sigma_j}(\boldsymbol{\tau}_{i_j}).$$

This is a U-statistic based on the i.i.d. sequence $(\boldsymbol{\tau}_i)$.

Theorem

For a random forest permutation

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) - \mu n^b}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2)$$

for some constants μ, γ.

Proof.

Hoeffding's theorem, with modifications as above. $\square$

Random tree permutations

"Theorem". For a random tree permutation $\boldsymbol{\pi}_n$ of length n, and a tree permutation σ,

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) - \mu n^b}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2)$$

for some $b \geq 1$ and constants μ, γ.

Proof.

To be completed next week.

$\square$

Random tree permutations

"Theorem". For a random tree permutation $\boldsymbol{\pi}_n$ of length n, and a tree permutation σ,

$$\frac{\operatorname{occ}_\sigma(\boldsymbol{\pi}_n) - \mu n^b}{n^{b-1/2}} \xrightarrow{\mathrm{d}} N(0, \gamma^2)$$

for some $b \geq 1$ and constants μ, γ.

Proof.
To be completed next week.
Uses a coding of tree permutations by a sequence of runs of 0's or 1's, which again permits $\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)$ to be written as a U-statistic.
This time we have to take a vincular U-statistic, and also use renewal theory as above.
Hence the two variations of U-statistics are combined.
Hoeffding's proof can still be adapted. (I believe.) $\square$