The Profile of a Random Permutation

Chaim Even-Zohar, UC Davis
chaim@math.ucdavis.edu

Introduction

Let $\pi \in S_n$ and let $k \leq n$. Consider all $\binom{n}{k}$ restrictions of π to k entries, $\pi(a_1), \ldots, \pi(a_k)$ where $a_1 < a_2 < \cdots < a_k$. The relative ordering of such k values induces a *pattern* $\sigma \in S_k$ as follows. The k-pattern σ is the unique permutation in S_k such that $\sigma(i) < \sigma(j)$ iff $\pi(a_i) < \pi(a_j)$ for every i and j. For example, the restriction of $\pi = \underline{4}\,\underline{1}\,2\,\underline{5}\,3$ to the marked entries induces the 3-pattern $\sigma = 2\,1\,3$.

For each $\sigma \in S_k$ we denote by $N_\sigma(\pi)$ the *number* of times it occurs as a k-pattern in π. The *density* of σ in π is defined as the proportion $P_\sigma(\pi) = N_\sigma(\pi)/\binom{n}{k}$, so that $P_\sigma(\pi) \in [0,1]$. The *k-profile* of π is the $k!$-dimensional vector of all k-pattern densities,

$$\mathbf{P}_k(\pi) \;=\; (\, P_\sigma(\pi)\,)_{\sigma \in S_k} \;\in\; \mathbb{R}^{k!}$$

When $\pi \in S_n$ is sampled uniformly at random, we denote its densities and k-profile by $P_{\sigma n}$ and $\mathbf{P}_{kn}$ respectively.

Pattern densities in permutations give rise to extremal questions [31, 1, 15, 30, 8, 2, 33], and play a role in the construction of limiting objects for permutations [27, 18, 14, 23], and in permutation property testing [19, 26]. The case where some pattern densities vanish [7, 25, 29] is studied extensively. This extended abstract focuses on the study of pattern densities in random permutations [9, 12, 5, 6, 8, 21, 17], which has applications to non-parametric statistics [34, 22, 16, 4, 11, 3].

A first observation is that each entry in the expected profile, $E[\mathbf{P}_{kn}]$, is $1/k!$ regardless of n. One can also establish a law of large numbers, that is, in probability

$$\mathbf{P}_{kn} \;\xrightarrow{n\to\infty}\; \mathbf{U}_k := \left(\frac{1}{k!}, \ldots, \frac{1}{k!}\right)$$

It is hence interesting to understand how the k-profile deviates from this limit. What is the order of magnitude of $(\mathbf{P}_{kn} - \mathbf{U}_k)$ as n grows? What directions in $\mathbb{R}^{k!}$ are typical of this vector? Does it have a natural decomposition into lower-dimensional components? What is the shape of the distribution when properly normalized?

We study these questions using group representations of S_k. Although developed independently, our work extends the discussion by Janson, Nakamura and Zeilberger in Section 4 of [21]. In particular, we address the question in its closing paragraph, on the emerging general structure.

Normalization

We first recall some notions concerning representations of finite groups and the classification of the simple representations of the symmetric group S_k [24, 13, 32].

A *d-dimensional real representation* of a finite group G is a map ρ from G to the linear group of $\mathbb{R}^d$, such that $\rho(g) \circ \rho(h) = \rho(gh)$ for every $g, h \in G$. A represnetation ρ is *simple* if there is no proper subspace $V \subset \mathbb{R}^d$ such that $\rho(g)V = V$ for every $g \in G$. Two representations ρ, ρ' of G are *similar* is there exists a linear map τ such that $\rho'(g) = \tau^{-1} \circ \rho(g) \circ \tau$ for every $g \in G$.

The simple representations of the symmetric group S_k up to similarity are in one-to-one correspondence with integer partitions of k. A *partition* $\lambda \vdash k$ is given by integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell \geq 1$

such that $\lambda_1 + \cdots + \lambda_\ell = k$, and the corresponding simple representation is denoted ρ^λ, of dimension d_λ. We note that $\sum_{\lambda \vdash k} d_\lambda^2 = k!$.

We use the simple representations of S_k to decompose $\mathbb{R}^{k!}$. Suppose that each representation ρ^λ is given as d_λ-by-d_λ matrices $(\, R^\lambda(\sigma)\,)_{\sigma \in S_k}$ in its similarity class. For $r < k$ consider the subspace

$$V_r \ := \ \text{span} \left\{ \mathbf{R}^\lambda_{ij} \mid {}^{\lambda \vdash k, \ \lambda_1 = k-r}_{1 \le i,j \le d_\lambda} \right\} \qquad \text{where} \qquad \mathbf{R}^\lambda_{ij} \ = \ \left(\, R^\lambda_{ij}(\sigma)\,\right)_{\sigma \in S_k} \in \ \mathbb{R}^{k!}$$

Then $\mathbb{R}^{k!} = V_0 \oplus \cdots \oplus V_{k-1}$ where the direct summands are well-defined, independent of the choice of R^λ in their similarity classes. They are also mutually orthogonal with respect to the the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{\sigma \in S_k} u_\sigma v_\sigma$.

We apply this decomposition to the random k-profile. Denote by $\Pi_r : \mathbb{R}^{k!} \to V_r$ the orthogonal projections on those subspaces.

$$\mathbf{P}_{kn} \ = \ \Pi_0 \mathbf{P}_{kn} + \Pi_1 \mathbf{P}_{kn} + \cdots + \Pi_{k-1} \mathbf{P}_{kn}$$

The first supspace V_0 is spanned by the 1-dimensional *trivial* repsentation $\rho^{(k)}(\sigma) = [1]$. Since $\sum_{\sigma \in S_k} P_{\sigma n} = 1$ deterministically, $\Pi_0 \mathbf{P}_{kn} = \mathbf{U}_k = E[\mathbf{P}_{kn}]$ as noted above. Consequently, the other terms satisfy $E[\Pi_r \mathbf{P}_{kn}] = 0$, for every $n \ge k > r > 0$, so that $\Pi_r \mathbf{P}_{kn} \to \mathbf{0}$ in probability in $\mathbb{R}^{k!}$. The following theorem determines their asymptotic order of magnitude.

Theorem 1. *For every $r < k$*

$$n^r \ E \left[\ \| \Pi_r \mathbf{P}_{kn} \|^2 \ \right] \ \xrightarrow{n \to \infty} \ \sigma_{kr}$$

for some $0 < \sigma_{kr} < \infty$.

We call $\Pi_r \mathbf{P}_{kn}$ the *order r component* of the k-profile, and deonte the *normalized k-profile* by

$$\tilde{\mathbf{P}}_{kn} \ = \ \sum_{r=0}^{k-1} n^{r/2} \, \Pi_r \mathbf{P}_{kn}$$

Another feature of this decomposition is that the normalized components are *asymptotically uncorrelated*. This is stated in the following theorem in terms of the cross-covariance matrix of such two vectors.

Theorem 2. *For every $r < s < k$*

$$E \left[\left(n^{r/2} \, \Pi_r \mathbf{P}_{kn} \right) \left(n^{s/2} \, \Pi_s \mathbf{P}_{kn} \right)^T \right] \ \xrightarrow{n \to \infty} \ 0$$

in the normed space $\mathbb{R}^{k! \times k!}$.

Diagonalization

Theorem 2 already suggests a partial diagonalization of the normalized k-profile distribution. One may proceed by decomposing every component $n^{r/2} \Pi_r \mathbf{P}_{kn}$ according to some orthogonal basis of V_r, so that the covariance matrix of all the $k!$ resulting components is asymptotically diagonal. Such a procedure is sometimes called *decorrelation* or *principal component analysis* (PCA).

It turns out that such a secondary decomposition is again provided by simple representations of S_k. Let $R = \{R^\lambda\}_{\lambda \vdash k}$ be matrix representations in the similarity classes of all ρ^λ as before. We consider all their *matrix elements*,

$$\mathbf{r}^\lambda_{ij} \ = \ \sqrt{\tfrac{d_\lambda}{k!}} \, \mathbf{R}^\lambda_{ij} \qquad \text{where} \qquad \mathbf{R}^\lambda_{ij} \ = \ \left(\, R^\lambda_{ij}(\sigma)\,\right)_{\sigma \in S_k} \in \ \mathbb{R}^{k!}$$

If the representations $R = \{R^\lambda\}_{\lambda \vdash k}$ are all unitary with respect to the usual inner products in $\mathbb{R}^{d_\lambda}$ then $\{\mathbf{r}^\lambda_{ij}\}$ form an orthonormal basis of $\mathbb{R}^{k!}$. This means that any vector $\mathbf{v} \in \mathbb{R}^{k!}$ is written in this basis with the following coefficients.

$$U_R\,\mathbf{v} \;=\; \left(\, \langle \mathbf{r}^\lambda_{ij}, \mathbf{v} \rangle \,\right)_{\lambda \vdash k,\; 1 \le i,j \le d_\lambda}$$

The following theorem indicates that such matrix elements should be used to diagonalize the asymptotic distribution of the k-profile.

Theorem 3. *Let $k \le 6$. There exist unitary simple representations $R = \{R^\lambda\}_{\lambda \vdash k}$, such that*

$$E\left[\left(U_R\,\tilde{\mathbf{P}}_{kn}\right)\left(U_R\,\tilde{\mathbf{P}}_{kn}\right)^T\right] \xrightarrow{n\to\infty} \Sigma$$

where $\Sigma \in \mathbb{R}^{k!\times k!}$ is diagonal with positive entries.

Extending Theorem 3 to every $k \in \mathbb{N}$ is work in progress. The discovery of these bases and the verification of the theorem were done by computer-aided exploration. The matrix representations R^λ that come up are not those of any of the classical constructions: Young's semi-normal form, Young's orthogonal form, or Young's natural form. It would be interesting to understand how they are related. Thus we also hope to resolve some degrees of freedom in our choice of matrices for certain representations of S_6.

Our treatment parallels the framework of spectral analysis of statistical data defined on non-abelian groups, as introduced by Diaconis [10, Section 8B]. An important practical issue that rises there is the arbitrary choice of Fourier bases, which might depend on matters of interpretation and convenience. The bases described here may provide an answer in the case of S_k, at least when the samples relate to occurrences of ordering types, and are associated with the ordering of a larger sequence.

Ditribution

The next challenge would be to understand the asymptotic shape of the profile's distribution. By the results of Janson, Nakamura and Zeilberger [21], the first order component $\sqrt{n}\,\Pi_1\mathbf{P}_{kn}$ is asymptotically multi-normal in the $(k-1)^2$-dimensional space V_1.

In this case, that corresponds to the partition $\lambda = (k-1, 1)$, we can characterize the unitary representation that appears. It is given by $R^\lambda(\sigma) = U^T A_\sigma U$, where $A_\sigma \in \mathbb{R}^{k\times k}$ permutes the standard basis of $\mathbb{R}^k$ according to σ. The matrix $U \in \mathbb{R}^{k\times(k-1)}$ is semi-orthogonal, where U_{ij} is a polynomial of degree j evaluated at i. This yields an orthonormal basis of V_1 in which the asymptotic distribution is given by $(k-1)^2$ independent normal variables.

As for higher order components, some scalar projections $\langle \mathbf{v}, \mathbf{P}_{kn} \rangle$ were studied in the statistical literature [16, 4, 11, 3]. In general, if $\mathbf{v} \in V_r$ then $\langle \mathbf{v}, \mathbf{P}_{kn} \rangle$ is a U-statistic of degree k, degenerate to order r [28]. We omit the details of this description. It means that after scaling by $n^{r/2}$ this scalar projection has the asymptotic distribution of a sum of degree r polynomials in independent Gaussians [20].

It is also interesting to understand the joint distribution of several profile projections of order greater than one. It seems that the decomposition according to the matrix elements $\{\mathbf{r}^\lambda_{ij}\}$ may be useful in studying those.

Applications

We briefly mention a few interesting special cases, as consequences of the structure of the profile's distribution described above. Some of them are related to well-studied permutation statistics.

1. $\lambda = (1,1)$: This 1-dimensional representation, $\mathrm{sign} : S_2 \to [\pm 1]$, corresponds to the inversion number of $\pi \in S_n$, or to Kendall's τ [22]. It is asymptotically normal of order $1/\sqrt{n}$.

$$\langle \mathbf{R}_{11}^{1+1}, \mathbf{P}_k(\pi) \rangle \;=\; \tau(\pi) \;=\; P_{12}(\pi) - P_{21}(\pi)$$

2. $\lambda = (2,1)$: In this case $d_{2+1} = 2$ and a suitable two-dimensional representation is generated by

$$R^{2+1}(321) \;=\; \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad R^{2+1}(132) \;=\; \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$$

The four limit distributions are independently normal, of order $1/\sqrt{n}$, and the proportion between their variances is as in $\left[\begin{smallmatrix} 25 & 5 \\ 5 & 1 \end{smallmatrix}\right]$.

The upper left element $\langle \mathbf{R}_{11}^{2+1}, \mathbf{P}_k \rangle$ is equivalent to Spearman's ρ [34], up to adding $\frac{1}{4}\mathbf{R}_{11}^{1+1+1}$ which has negligible order $1/n$. The lower right one $\langle \mathbf{R}_{22}^{2+1}, \mathbf{P}_k \rangle$ corresponds, up to constants, to the count of monotone triplets $P_{123} + P_{321}$.

3. $\lambda = (1,1,1)$: The alternating representation, $\mathrm{sign} : S_3 \to [\pm 1]$, yields Fisher and Lee's statistic for circular rank correlation [11], or Gepner's statistic [35]:

$$\langle \mathbf{R}_{11}^{1+1+1}, \mathbf{P}_k \rangle \;=\; P_{123} + P_{231} + P_{312} - P_{132} - P_{213} - P_{321}$$

It has order $1/n$, and its limiting distribution was studied [11].

4. $\lambda = (1,\ldots,1)$: Extending the cases $k = 2, 3$, the alternating representation of S_k yields

$$\langle \mathbf{R}_{11}^{1+\cdots+1}, \mathbf{P}_k \rangle \;=\; \left(\#_{k\text{-patterns}}^{\text{even}} \;-\; \#_{k\text{-patterns}}^{\text{odd}} \right) / \tbinom{n}{k}$$

Its distribution was mentioned as being particularly narrow, of order $n^{-(k-1)/2}$ [21]. Since $\lambda_1 = 1$ we can see it is indeed the single most narrow projection of the k-profile.

5. $\lambda = (3,1)$: This case yields $d_{3+1}^2 = 9$ projections, with a multi-normal axis-aligned distribution, of order $1/\sqrt{n}$. The representation is generated by the matrices

$$R^{3+1}(1243) \;=\; \begin{bmatrix} 0.8 & \sqrt{0.2} & -0.4 \\ \sqrt{0.2} & 0 & \sqrt{0.8} \\ -0.4 & \sqrt{0.8} & 0.2 \end{bmatrix} \qquad R^{3+1}(3142) \;=\; \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

The variances are proportional to the rank-1 matrix $\left[\begin{smallmatrix} 441 & 147 & 21 \\ 147 & 49 & 7 \\ 21 & 7 & 1 \end{smallmatrix}\right]$.

6. $\lambda = (2,2)$: This representation has the same 2-by-2 matrices as $(2,1)$ above, with $R^{2+2}(1243) = R^{2+1}(321)$ and $R^{2+2}(3142) = R^{2+1}(132)$. This yields four uncorrelated statistics of degenrate order $1/n$. The lower right element was proposed by Bergsma and Dassios [3] as $\tau^\star(\pi)$, a consistent test of independence for paired samples.

$$\tau^\star \;=\; \tfrac{2}{3}\,\langle \mathbf{R}_{22}^{2+2}, \mathbf{P}_k \rangle \;=\; P_{1234} + P_{1243} + P_{2134} + P_{2143} + P_{3412} + P_{3421} + P_{4312} + P_{4321} - \tfrac{1}{3}$$

It can be shown by our decomposition that $\tau^\star$ is in fact asymptotically equaivalent to classical independence tests by Hoeffding [16] and Blum–Kiefer–Rosenblatt [4]. They only differ by order $n^{-3/2}$ terms that come from the representations $(2,2,1)$ and $(3,2,1)$ respectively. Král' and Pikhurko [27] show that it follows from several projections of the asymptotic 4-profile.

7. $\lambda = (2,1,1)$: To complete the treatment of $k = 4$, let

$$R^{2+1+1}(1243) \;=\; \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad R^{2+1+1}(3142) \;=\; \begin{bmatrix} 0 & 0 & -1 \\ 0 & +1 & 0 \\ +1 & 0 & 0 \end{bmatrix}$$

This representation has a nice form of signed permutation matrices. Each of the nine order $1/n$ statistics counts the occurrences of some four 4-patterns minus those of another four.

References

[1] Michael H. Albert, Mike D. Atkinson, Chris C Handley, Derek A. Holton, and Walter Stromquist. On packing densities of permutations. *Electron. J. Combin*, 9(1), 2002.

[2] József Balogh, Ping Hu, Bernard Lidickỳ, Oleg Pikhurko, Balázs Udvari, and Jan Volec. Minimum number of monotone subsequences of length 4 in permutations. *Combinatorics, Probability and Computing*, 24(04):658–679, 2015.

[3] Wicher Bergsma, Angelos Dassios, et al. A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli*, 20(2):1006–1028, 2014.

[4] Julius R Blum, J Kiefer, and Murray Rosenblatt. Distribution free tests of independence based on the sample distribution function. *The annals of mathematical statistics*, pages 485–498, 1961.

[5] Miklós Bóna. The copies of any permutation pattern are asymptotically normal. *arXiv:0712.2792*, 2007.

[6] Miklós Bóna. On three different notions of monotone subsequences. *Permutation patterns*, 376:89, 2010.

[7] Miklós Bóna. *Combinatorics of permutations*. CRC Press, 2012.

[8] Alexander Burstein and Peter Hästö. Packing sets of patterns. *European Journal of Combinatorics*, 31(1):241–253, 2010.

[9] Joshua N Cooper. Quasirandom permutations. *Journal of Combinatorial Theory, Series A*, 106(1):123–143, 2004.

[10] Persi Diaconis. Group representations in probability and statistics, 1988.

[11] NI Fisher and AJ Lee. Nonparametric measures of angular-angular association. *Biometrika*, pages 315–321, 1982.

[12] Jason Fulman et al. Stein's method and non-reversible Markov chains. In *Stein's Method*, pages 66–74. Institute of Mathematical Statistics, 2004.

[13] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.

[14] Roman Glebov, Andrzej Grzesik, Tereza Klimošová, et al. Finitely forcible graphons and permutons. *Journal of Combinatorial Theory, Series B*, 110:112–135, 2015.

[15] Peter A Hasto. The packing density of other layered permutations. *Journal of Combinatorics*, 9(2):1, 2002.

[16] Wassily Hoeffding. A non-parametric test of independence. *The annals of mathematical statistics*, pages 546–557, 1948.

[17] Lisa Hofer. A central limit theorem for vincular permutation patterns. *arXiv:1704.00650*, 2017.

[18] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, Balázs Ráth, and Rudini Menezes Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, 2013.

[19] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, and Rudini Menezes Sampaio. Testing permutation properties through subpermutations. *Theoretical Computer Science*, 412(29):3555–3567, 2011.

[20] Svante Janson. *Gaussian hilbert spaces*, volume 129. Cambridge university press, 1997.

[21] Svante Janson, Brian Nakamura, and Doron Zeilberger. On the asymptotic statistics of the number of occurrences of multiple permutation patterns. *Journal of Combinatorics*, 6:117–143, 2015.

[22] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

[23] Richard Kenyon, Daniel Kral, Charles Radin, and Peter Winkler. Permutations with fixed pattern densities. *arXiv:1506.02340*, 2015.

[24] Adalbert Kerber and GD James. *The Representation Theory of the Symmetric Group (Encyclopedia of Mathematics and Its Applications)*. Cambridge University Press, 1984.

[25] Sergey Kitaev. *Patterns in permutations and words*. Springer Science & Business Media, 2011.

[26] Tereza Klimošová and Daniel Král. Hereditary properties of permutations are strongly testable. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1164–1173. Society for Industrial and Applied Mathematics, 2014.

[27] Daniel Král' and Oleg Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23(2):570–579, 2013.

[28] Justin Lee. U-statistics: Theory and practice, 1990.

[29] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 107(1):153–160, 2004.

[30] Cathleen Battiste Presutti and Walter Stromquist. Packing rates of measures and a conjecture for the packing density of 2413. *Permutation patterns*, 376:287–316, 2010.

[31] Alkes Long Price. Packing densities of layered* patterns, 1997.

[32] Bruce Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203. Springer Science & Business Media, 2013.

[33] Jakub Sliacan and Walter Stromquist. Improving bounds on packing densities of 4-point permutations. *arXiv:1704.02959*, 2017.

[34] Charles Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101, 1904.

[35] Doron Zeilberger. Doron Gepner's statistics on words in {1, 2, 3} is (most probably) asymptotically logistic. *arXiv:1604.00663*, 2016.