Explicit computation of the variance of the number of maxima in hypercubes

Christian COSTERMANS – HOANG NGOC MINH

Université Lille 2

Let $\Lambda = \{x_1, \dots, x_n\}$ be a set of independent and uniformly distributed random vectors in $[0,1]^d$. A point $x_i = (x_{i_1}, \dots, x_{i_d})$ is said to be dominated by $x_j = (x_{j_1}, \dots, x_{j_d})$ if $x_{i_k} < x_{j_k}$ for all $k \in \{1, \dots, d\}$ and a point x_i is called a maximum of Λ if none of the other points dominates it. The number of maxima of Λ is denoted by $K_{n,d}$.

Combinatoric tools

Previous results : formulas from [1, 2, 3]

$$\mathbb{E}(K_{n,d}^2) = \mu_{n,d} + \sum_{1 \le t \le d-1} \binom{d}{t} \sum_{l=1}^{n-1} \frac{1}{l} \sum^{(*)} \frac{1}{i_1 \dots i_{d-2} j_1 \dots j_{d-1}},$$

where the sum $(*)$ is taken over indices verifying $1 \le i_1 \dots \le i_{t-1} \le l, 1 \le i_t \le \dots \le i_{d-2} \le l$ **and** $l+1 \le j_1 \le \dots \le j_{d-1} \le n$,

$$\mu_{n,d} = \mathbb{E}(K_{n,d}) = \sum_{1 \le i_1 \le \dots \le i_{d-1} \le n} \frac{1}{i_1 \dots i_{d-1}}.$$

Multi-index $(s_1, \dots, s_r) \leftrightarrow$ word $w = y_{s_1} \dots y_{s_r} \in Y^*$, where $Y = \{y_i, i > 0\}$.
Some notations : $w = y_{s_1} \dots y_{s_r}$ has for length r and for weight $s_1 + \dots + s_r$.

$$\mathrm{A}_w(N) := \sum_{N \ge n_1 \ge \dots \ge n_r \ge 1} \frac{1}{n_1^{s_1} \dots n_r^{s_r}}.$$

For $s_1 > 1$, $\mathrm{A}_w(N) \underset{N \to +\infty}{\longrightarrow} \Theta(w) := \sum_{n_1 \ge \dots \ge n_r \ge 1} \frac{1}{n_1^{s_1} \dots n_r^{s_r}}.$

Definition : The polyzêta $\zeta(w)$ is defined for $w = y_{s_1} \dots y_{s_r}$ by

$$\zeta(w) := \sum_{n_1 > \dots > n_r \ge 1} \frac{1}{n_1^{s_1} \dots n_r^{s_r}}.$$

Proposition : $\Theta(w)$ is a linear combination of polyzêtas.
Example :

$$\Theta(y_2 y_1 y_1) = \zeta(y_2 y_1 y_1) + \zeta(y_2 y_2) + \zeta(y_3 y_1) + \zeta(y_4),$$

Or, equivalently

$$\Theta(2,1,1) = \zeta(2,1,1) + \zeta(2,2) + \zeta(3,1) + \zeta(4).$$

Results

$N_i(w)$ stands for the number of occurences of the letter y_i in the word w and $\{y_1, y_2\}_\rho$ for the set of words built over letters $\{y_1, y_2\}$, and of weight ρ.

• Asymptotic equivalent

$$\mathbb{V}ar(K_{n,d}) \sim \left(\frac{1}{(d-1)!} + \kappa_d\right) \ln^{d-1}(n), \text{ proved by Bai et al. [1]},$$

$$\kappa_d = \frac{1}{(d-1)!} \sum_{w \in \{y_1, y_2\}_{d-3}} (-1)^{N_2(w)} \binom{2\,(d-2-N_2(w))}{d-2-N_2(w)} \Theta(y_2 w)$$

Example :

$$\begin{aligned} 5!\kappa_6 &= \binom{8}{4}\Theta(2,1,1,1) - \binom{6}{3}\left(\Theta(2,1,2) + \Theta(2,2,1)\right) \\ &= \binom{8}{4}\left(\zeta(2,1,1,1) + \zeta(2,1,2) + \zeta(2,2,1) + \zeta(2,3) + \dots\right). \end{aligned}$$

• Next terms

Proposition : Let $\mathcal{Z}$ be the $\mathbb{Q}$-algebra generated by polyzêtas, i.e. $\{\zeta(w), w \in Y^* \setminus y_1 Y^*\}$ and let $\mathcal{Z}'$ be the $\mathbb{Q}[\gamma]$-algebra generated by $\mathcal{Z}$. Then there exist algorithmically computable coefficients $b_i \in \mathcal{Z}'$, $\kappa_i \in \mathbb{N}$ and $\eta_i \in \mathbb{Z}$ such that, for any $w \in Y^*$,

$$\mathrm{A}_w(N) \sim \sum_{i=0}^{+\infty} b_i N^{\eta_i} \log^{\kappa_i}(N), \text{ for } N \to +\infty.$$

Theorem : There exist algorithmically computable coefficients $\alpha_i, \beta_{j,k} \in \mathcal{Z}'$ such that, for any dimension d and any order M,

$$\mathbb{V}ar(K_{n,d}) = \sum_{i=0}^{d-1} \alpha_i \ln^i(n) + \sum_{j=1}^{M} \frac{1}{n^j} \sum_{k=0}^{2d-2} \beta_{j,k} \ln^k(n) + \mathrm{o}\left(\frac{1}{n^M}\right).$$

Computations

The following results can be reached thanks to a final step, which is reducing into irreducible polyzêtas [4].

$$\begin{aligned} \kappa_2 &= 0 \\ \kappa_3 &= \zeta(2) \\ \kappa_4 &= 2\,\zeta(3) \\ \kappa_5 &= \frac{33}{40}\zeta(2)^2 \end{aligned}$$

$$\begin{aligned} \kappa_6 &= \frac{5}{4}\zeta(5) + \frac{1}{6}\zeta(2)\,\zeta(3) \\ \kappa_7 &= \frac{1451}{7560}\zeta(2)^3 + \frac{7}{72}\zeta(3)^2 \\ \kappa_8 &= \frac{1729}{5760}\zeta(7) + \frac{181}{3600}\zeta(3)\,\zeta(2)^2 + \frac{13}{360}\zeta(2)\,\zeta(5) \end{aligned}$$

$$\begin{aligned} \kappa_9 &= -\frac{17}{1920}\zeta(6,2) + \frac{11}{160}\zeta(3)\,\zeta(5) \\ &+ \frac{1}{320}\zeta(2)\,\zeta(3)^2 + \frac{1891}{89600}\zeta(2)^4 \\ \kappa_{10} &= \frac{529}{75600}\zeta(2)^2\zeta(5) + \frac{33941}{6350400}\zeta(2)^3\zeta(3) \\ &+ \frac{17}{3360}\zeta(2)\,\zeta(7) + \frac{199271}{4354560}\zeta(9) + \frac{11}{12960}\zeta(3)^3 \end{aligned}$$

$$\begin{aligned} \mathbb{V}ar(K_{n,3}) &= \left(\frac{1}{2} + \kappa_3\right)\ln^2(n) + \left(-10\zeta(3) + 2\zeta(2)\gamma + \gamma\right)\ln(n) \\ &+ \frac{1}{2}\gamma^2 - 10\zeta(3)\gamma + \frac{83}{10}\zeta(2)^2 + \zeta(2)\gamma^2 + \frac{1}{2}\zeta(2) + \mathrm{o}(1) \\ \mathbb{V}ar(K_{n,4}) &= \left(\frac{1}{3!} + \kappa_4\right)\ln^3(n) + \left(-\frac{53}{5}\zeta(2)^2 + 6\,\zeta(3)\,\gamma + \frac{1}{2}\gamma\right)\ln^2(n) \\ &+ \left(97\,\zeta(5) - \frac{106}{5}\zeta(2)^2\gamma + 16\,\zeta(2)\,\zeta(3) + 6\,\zeta(3)\,\gamma^2 + \frac{1}{2}\zeta(2) + \frac{1}{2}\gamma^2\right)\ln(n) \\ &+ \frac{1}{3}\zeta(3) - \frac{53}{5}\zeta(2)^2\gamma^2 - \frac{3719}{70}\zeta(2)^3 + \frac{1}{6}\gamma^3 + \frac{1}{2}\zeta(2)\,\gamma \\ &+ 16\,\zeta(2)\,\zeta(3)\,\gamma - 3\,\zeta(3)^2 + 2\,\zeta(3)\,\gamma^3 + 97\,\zeta(5)\,\gamma + \mathrm{o}(1) \end{aligned}$$

$$\begin{aligned} \mathbb{V}ar(K_{n,5}) &= \left(\frac{1}{4!} + \kappa_5\right)\ln^4(n) + \left(\frac{1}{6}\gamma - \frac{98}{3}\zeta(5) + \frac{33}{10}\zeta(2)^2\gamma - \frac{13}{3}\zeta(2)\,\zeta(3)\right)\ln^3(n) \\ &+ \left(\frac{10123}{140}\zeta(2)^3 + \frac{47}{2}\zeta(3)^2 + \frac{99}{20}\zeta(2)^2\gamma^2 + \frac{1}{4}\gamma^2 + \frac{1}{4}\zeta(2) - 13\,\zeta(2)\,\zeta(3)\,\gamma\right. \\ &- \left.98\,\zeta(5)\,\gamma\right)\ln^2(n) + \left(\frac{1}{6}\gamma^3 + \frac{33}{10}\zeta(2)^2\gamma^3 + \frac{1}{2}\zeta(2)\,\gamma - 950\,\zeta(7)\right. \\ &- 13\,\zeta(2)\,\zeta(3)\,\gamma^2 + 47\,\zeta(3)^2\gamma + \frac{1}{3}\zeta(3) - \frac{317}{5}\zeta(3)\,\zeta(2)^2 + \frac{10123}{70}\zeta(2)^3\gamma \\ &- \left.98\,\zeta(5)\,\gamma^2 - 222\,\zeta(2)\,\zeta(5)\right)\ln(n) - \frac{13}{3}\zeta(2)\,\zeta(3)\,\gamma^3 + \frac{47}{2}\zeta(3)^2\gamma^2 \\ &- \frac{317}{5}\zeta(3)\,\zeta(2)^2\gamma - \frac{98}{3}\zeta(5)\,\gamma^3 + \frac{33}{40}\zeta(2)^2\gamma^4 + \frac{32}{3}\zeta(3)\,\zeta(5) + \frac{10123}{140}\zeta(2)^3\gamma^2 \\ &- 222\,\zeta(2)\,\zeta(5)\,\gamma + \frac{1}{24}\gamma^4 - 950\,\zeta(7)\,\gamma + 50\,\zeta(6,2) + \frac{1}{4}\zeta(2)\,\gamma^2 + \frac{1}{3}\zeta(3)\,\gamma \\ &+ \frac{9}{40}\zeta(2)^2 + \frac{95}{6}\zeta(2)\,\zeta(3)^2 + \frac{134739}{350}\zeta(2)^4 + \mathrm{o}(1) \end{aligned}$$

$$\begin{aligned} \mathbb{V}ar(K_{n,6}) &= \left(\frac{1}{5!} + \kappa_6\right)\ln^5(n) + \left(\frac{1}{24}\gamma + \frac{25}{4}\zeta(5)\,\gamma + \frac{5}{6}\zeta(2)\,\zeta(3)\,\gamma - \frac{25}{6}\zeta(3)^2\right. \\ &- \left.\frac{22711}{2520}\zeta(2)^3\right)\ln^4(n) + \left(\frac{1}{12}\gamma^2 + \frac{1231}{30}\zeta(3)\,\zeta(2)^2 - \frac{50}{3}\zeta(3)^2\gamma\right. \\ &+ \frac{8729}{24}\zeta(7) + \frac{127}{2}\zeta(2)\,\zeta(5) + \frac{1}{12}\zeta(2) + \frac{5}{3}\zeta(2)\,\zeta(3)\,\gamma^2 - \frac{22711}{630}\zeta(2)^3\gamma \\ &+ \left.\frac{25}{2}\zeta(5)\,\gamma^2\right)\ln^3(n) + \left(-55\,\zeta(6,2) - \frac{241}{6}\zeta(2)\,\zeta(3)^2 + \frac{1231}{10}\zeta(3)\,\zeta(2)^2\gamma\right. \\ &+ \frac{1}{12}\gamma^3 + \frac{1}{4}\zeta(2)\,\gamma + \frac{8729}{8}\zeta(7)\,\gamma - \frac{2331589}{4200}\zeta(2)^4 + \frac{1}{6}\zeta(3) + \frac{25}{2}\zeta(5)\,\gamma^3 \\ &+ \frac{5}{3}\zeta(2)\,\zeta(3)\,\gamma^3 - 342\,\zeta(3)\,\zeta(5) - 25\,\zeta(3)^2\gamma^2 - \frac{22711}{420}\zeta(2)^3\gamma^2 \\ &+ \left.\frac{381}{2}\zeta(2)\,\zeta(5)\,\gamma\right)\ln^2(n) + \left(\frac{8729}{8}\zeta(7)\,\gamma^2 - \frac{2331589}{2100}\zeta(2)^4\gamma + \frac{381}{2}\zeta(2)\,\zeta(5)\,\gamma^2\right. \\ &- \frac{241}{3}\zeta(2)\,\zeta(3)^2\gamma + \frac{1231}{10}\zeta(3)\,\zeta(2)^2\gamma^2 - 19\,\zeta(3)^3 + \frac{1}{4}\zeta(2)\,\gamma^2 + \frac{9}{40}\zeta(2)^2 \\ &- \frac{50}{3}\zeta(3)^2\gamma^3 - 110\,\zeta(6,2)\,\gamma + \frac{25}{4}\zeta(5)\,\gamma^4 + \frac{1}{3}\zeta(3)\,\gamma + \frac{21919}{20}\zeta(2)^2\zeta(5) \\ &- \frac{22711}{630}\zeta(2)^3\gamma^3 + \frac{135593}{315}\zeta(2)^3\zeta(3) + \frac{182179}{18}\zeta(9) + \frac{5}{6}\zeta(2)\,\zeta(3)\,\gamma^4 \\ &- \left.684\,\zeta(3)\,\zeta(5)\,\gamma + \frac{19209}{8}\zeta(2)\,\zeta(7) + \frac{1}{24}\gamma^4\right)\ln(n) + \frac{127}{2}\zeta(2)\,\zeta(5)\,\gamma^3 \\ &+ \frac{1231}{30}\zeta(3)\,\zeta(2)^2\gamma^3 - \frac{241}{6}\zeta(2)\,\zeta(3)^2\gamma^2 - 342\,\zeta(3)\,\zeta(5)\,\gamma^2 + \frac{1}{6}\zeta(2)\,\zeta(3)\,\gamma^5 \\ &+ \frac{1}{6}\zeta(2)\,\zeta(3) + \frac{182179}{18}\zeta(9)\,\gamma + \frac{9}{40}\zeta(2)^2\gamma - 19\,\zeta(3)^3\gamma + \frac{1}{6}\zeta(3)\,\gamma^2 \\ &- \frac{2331589}{4200}\zeta(2)^4\gamma^2 + \frac{1}{5}\zeta(5) - 55\,\zeta(6,2)\,\gamma^2 - 325\,\zeta(8,2) + \frac{135593}{315}\zeta(2)^3\zeta(3)\,\gamma \\ &- 55\,\zeta(2)\,\zeta(6,2) + \frac{1}{12}\zeta(2)\,\gamma^3 + \frac{8729}{24}\zeta(7)\,\gamma^3 - \frac{22711}{2520}\zeta(2)^3\gamma^4 - \frac{25}{6}\zeta(3)^2\gamma^4 + \frac{1}{120}\gamma^5 \\ &- 945\,\zeta(5)^2 - \frac{6237237}{2200}\zeta(2)^5 + \frac{767}{30}\zeta(2)^2\zeta(3)^2 - \frac{9031}{12}\zeta(3)\,\zeta(7) - 392\,\zeta(5)\,\zeta(2)\,\zeta(3) \\ &+ \frac{5}{4}\zeta(5)\,\gamma^5 + \frac{21919}{20}\zeta(2)^2\zeta(5)\,\gamma + \frac{19209}{8}\zeta(2)\,\zeta(7)\,\gamma + \mathrm{o}(1). \end{aligned}$$

References

[1] Z.D. Bai, C.C. Chao, H.K. Hwang, and W.Q. Liang. On the variance of the number of maxima in random vectors and its applications. ***Annals of Applied Probability***, 8:886–895, 1998.

[2] O. Barndorff-Nielsen and M. Sobel. On the distribution of the number of admissible points in a vector random sample. ***Theory Proba. Appl.***, 11:249–269, 1966.

[3] V.M. Ivanin. Calculation of the dispersion of the number of elements of the pareto set for the choice of independent vectors with independent components. ***Theory of Optimal Decisions, Akad. Nauk. Ukrain. SSR Inst. Kibernet., Kiev***, pages 90–100, 1976.

[4] Hoang Ngoc Minh and M. Petitot. Lyndon words, polylogarithms and the riemann ζ function. ***Discrete Mathematics***, 217:273–292, 2000.