Variance for the Number of Maxima in Hypercubes and Generalized Euler's γ constants

Christian COSTERMANS – HOANG NGOC MINH

Université Lille 2

Maxima in hypercubes and combinatoric tools

Let $\mathcal{Q} = \{q_1, \ldots, q_n\}$ be a set of independent and uniformly distributed random vectors in $[0,1]^d$. $q_i = (q_{i_1}, \ldots, q_{i_d})$ dominated by $q_j = (q_{j_1}, \ldots, q_{j_d})$ if and only if $q_{i_k} < q_{j_k}$, for any $k \in \{1, \ldots, d\}$. q_i is called a maximum of $\mathcal{Q}$ if non-dominated. The number of maxima of $\mathcal{Q}$ is denoted by $K_{n,d}$.

Our goal : getting the full asymptotic expansion of $\mathbb{V}ar(K_{n,d})$ via ***polylogarithm*** and ***multiple harmonic sum***,

$$\mathrm{Li}_{s_1,\ldots,s_r}(z) = \sum_{n_1 > \ldots > n_r \geq 1} \frac{z^{n_1}}{n_1^{s_1} \ldots n_r^{s_r}} \text{ and } \mathrm{H}_{s_1,\ldots,s_r}(N) = \sum_{N \geq n_1 > \ldots > n_r \geq 1} \frac{1}{n_1^{s_1} \ldots n_r^{s_r}},$$

and via the asymptotic expansion of their noncommutative generating series,

$$\Lambda(z) = \sum_{w \in Y^*} \mathrm{Li}_w(z)\, w \text{ and } \mathrm{H}(N) = \sum_{w \in Y^*} \mathrm{H}_w(N)\, w.$$

Here we use the correspondence between $(s_1, \ldots, s_r)$ and $w = y_{s_1} \ldots y_{s_r}$ over $Y = \{y_i, i > 0\}$, and then $v = x_0^{s_1-1} x_1 \ldots x_0^{s_r-1} x_1$ over $X = \{x_0, x_1\}$. For $w \in Y^* \setminus y_1 Y^*$, ***i.e.*** for $v \in x_0 X^* x_1$, the limits $\lim_{z \to 1} \mathrm{Li}_w(z)$ and $\lim_{N \to +\infty} \mathrm{H}_w(N)$ exist and are equal to the ***convergent polyzêta***

$$\zeta(s_1, \ldots, s_r) = \zeta(y_{s_1} \ldots y_{s_r}) = \zeta(x_0^{s_1-1} x_1 \ldots x_0^{s_r-1} x_1), \text{ for } s_1 > 1.$$

Recall that the noncommutative generating series over X of regularized polyzêtas can be expressed as [4, 5]

$$Z = \prod_{l \in \mathcal{L}ynX \setminus \{x_0, x_1\}}^{\searrow} e^{\zeta(S_l)\, \check{S}_l},$$

where $\mathcal{L}ynX$ denotes the set of Lyndon words over X, $\{S_l\}_{l \in \mathcal{L}ynX}$ (resp. $\{\check{S}_l\}_{l \in \mathcal{L}ynX}$) denotes a transcendental basis (resp. the Lyndon basis) of the ***shuffle algebra*** (resp. ***free Lie algebra***) over $\mathbb{C}$ on X.

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

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

where the sum $(*)$ is taken over indices verifying

$$1 \leq i_1 \ldots \leq i_{t-1} \leq l, 1 \leq i_t \leq \ldots \leq i_{d-2} \leq l \quad \text{and} \quad l+1 \leq j_1 \leq \ldots \leq j_{d-1} \leq n.$$

Bai ***et al.*** proved that [1]

$$\mathbb{V}ar(K_{n,d}) \sim \left(\frac{1}{(d-1)!} + \kappa_d \right) \ln^{d-1}(n),$$

κ_d being a real constant ***the nature*** of which will be precised in the next section.

Introducing $\quad \mathrm{A}_{y_{s_1} \ldots y_{s_r}}(N) = \sum_{N \geq n_1 \geq \ldots \geq n_r \geq 1} \frac{1}{n_1^{s_1} \ldots n_r^{s_r}},$

we get $\quad \mathbb{E}(K_{n,d}^2) = \mathrm{A}_{y_1^{d-1}}(n) + \sum_{1 \leq t \leq d-1} \binom{d}{t} \sum_{l=1}^{n-1} \frac{1}{l} \mathrm{A}_{y_1^{t-1}}(l) \mathrm{A}_{y_1^{d-t-1}}(l) \mathrm{A}_{y_1^{d-1}}(n; l+1),$

where the notation $\mathrm{A}_w(n; l)$ stands for the same object than $\mathrm{A}_w(n)$, but with indices bounded by $l + 1$. For any $w \in Y^*$, A_w is a linear combination of multiple harmonic sums. For example,

$$\mathrm{A}_{y_2 y_1 y_1} = \mathrm{H}_{y_2 y_1 y_1} + \mathrm{H}_{y_2 y_2} + \mathrm{H}_{y_3 y_1} + \mathrm{H}_{y_4} \quad \Longleftrightarrow \quad \mathrm{A}_{2,1,1} = \mathrm{H}_{2,1,1} + \mathrm{H}_{2,2} + \mathrm{H}_{3,1} + \mathrm{H}_4.$$

For $w \in Y^* \setminus y_1 Y^*$, the limit $\underline{\zeta}(w) = \lim_{N \to +\infty} \mathrm{A}_w(N)$ also exists and is a linear combination of polyzêtas.

Results *à l'Abel* and Generalized Euler's γ constants

The ***quasi-shuffle*** product $\uplus$ of $u = y_i u'$ and $v = y_j v'$ is the polynomial in Y^* defined recursively by $\epsilon \uplus u = u \uplus \epsilon = u$ and $u \uplus v = y_i(u' \uplus v) + y_j(u \uplus v') + y_{i+j}(u' \uplus v')$.
Definition : Let $\zeta_{\uplus} : (\mathbb{C}\langle Y \rangle, \uplus) \to (\mathbb{C}, .)$ the algebra morphism verifying for $w \in Y^* \setminus y_1 Y^*$, $\zeta_{\uplus}(w) = \zeta(w)$ and such that $\zeta_{\uplus}(y_1) = \gamma$.

Theorem : Let π_Y stand for the projector from $\mathbb{C}\langle X \rangle$ over $\mathbb{C}\langle Y \rangle$ erasing monomials ending by x_0. Then

$$\lim_{z \to 1} \exp\left(-y_1 \log \frac{1}{1-z} \right) \Lambda(z) = \lim_{N \to +\infty} \exp\left(\sum_{k \geq 1} \mathrm{H}_{y_k}(N) \frac{(-y_1)^k}{k} \right) \mathrm{H}(N) = \pi_Y Z.$$

Therefore,

$$\mathrm{H}(N) \underset{N \to \infty}{\widetilde{\quad}} \exp\left(-\sum_{k \geq 1} \mathrm{H}_{y_k}(N) \frac{(-y_1)^k}{k} \right) \pi_Y Z.$$

So we get, looking at the constant part in this and introducing the Bell polynomials $b_{n,k}(t_1, \ldots, t_{n-k+1})$ in $\{t_i\}_{i \geq 1}$ with the specialization $t_i = (-1)^{i+1}(i-1)!\zeta_{\uplus}(y_i)$,

Corollary : Let $Z_{\uplus}$ be the noncommutative generating series of the constants $\{\zeta_{\uplus}(w)\}_{w \in Y^*}$. Then

$$Z_{\uplus} = \sum_{w \in Y^*} \zeta_{\uplus}(w)\, w = \left[1 + \sum_{n \geq 1} \left(\sum_{k=1}^{n} b_{n,k}(\gamma, -\zeta(2), 2\zeta(3), \ldots) \right) \frac{y_1^n}{n!} \right] \pi_Y Z.$$

Identifying coefficients of $y_1^k w$ leads to :

Corollary : For any $k \geq 0$ and for any $w \in x_0 X^* x_1$, ***i.e.*** $w = x_0 u$ and $\pi_Y w \in Y^* \setminus y_1 Y^*$, we have

$$\zeta_{\uplus}(x_1^k w) = \sum_{i=0}^{k} \frac{\zeta(x_0[(-x_1)^{k-i} \sqcup\!\sqcup\, u])}{i!} \left[\sum_{j=1}^{i} b_{i,j}(\gamma, -\zeta(2), 2\zeta(3), \ldots) \right],$$

where $\sqcup\!\sqcup$ denotes the well known ***shuffle*** over X^*.

In particular, we have

$$\zeta_{\uplus}(x_1^k) = \frac{1}{k!} \sum_{j=1}^{k} b_{k,j}(\gamma, -\zeta(2), 2\zeta(3), \ldots).$$

Example :

$$\zeta_{\uplus}(1,1,2) = 3\zeta(2,1,1) - 2\zeta(2,1)\gamma + \frac{\zeta(2)}{2}\left(-\zeta(2) + \gamma^2 \right),$$

$$\zeta_{\uplus}(1,1) = \frac{\gamma^2 - \zeta(2)}{2}, \quad \zeta_{\uplus}(1,1,1) = \frac{\gamma^3 - 3\zeta(2)\gamma + 2\zeta(3)}{6}.$$

Proposition :

$$\kappa_d = \frac{1}{(d-1)!} \sum_{w \in \{y_1, y_2\}_{d-3}} (-1)^{|w|_2} \binom{2(d-2-|w|_2)}{d-2-|w|_2} \underline{\zeta}(y_2 w),$$

where $|w|_i$ 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 (the ***weight*** of $y_{s_1} \ldots y_{s_r}$ being $\sum_{i=1}^r s_i$).

Example :

$$\begin{aligned} 5!\kappa_6 &= \binom{8}{4} \underline{\zeta}(2,1,1,1) - \binom{6}{3} \left(\underline{\zeta}(2,1,2) + \underline{\zeta}(2,2,1) \right) \\ &= 70\left(\zeta(2,1,1,1) + \zeta(3,1,1) + \zeta(4,1) \right) \\ &+ 50\left(\zeta(2,1,2) + \zeta(2,2,1) \right) + 30\left(\zeta(2,3) + \zeta(3,2) + \zeta(5) \right). \end{aligned}$$

Finally, we get

Theorem : Let $\mathcal{Z}$ be the $\mathbb{Q}$-algebra generated by $\{\zeta(w), w \in Y^* \setminus y_1 Y^*\}$ and let $\mathcal{Z}'$ be the $\mathbb{Q}[\gamma]$-algebra generated by $\mathcal{Z}$. There exist algorithmically computable $\alpha_i, \beta_{j,k} \in \mathcal{Z}'$ such that, for any d and any 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 [6].

$$\kappa_2 = 0, \quad \kappa_3 = \zeta(2), \quad \kappa_4 = 2\zeta(3), \quad \kappa_5 = \frac{33}{40}\zeta(2)^2, \quad \kappa_6 = \frac{5}{4}\zeta(5) + \frac{1}{6}\zeta(2)\zeta(3), \quad \kappa_7 = \frac{1451}{7560}\zeta(2)^3 + \frac{7}{72}\zeta(3)^2, \quad \kappa_8 = \frac{1729}{5760}\zeta(7) + \frac{181}{3600}\zeta(3)\zeta(2)^2 + \frac{13}{360}\zeta(2)\zeta(5),$$

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

$$\mathbb{V}ar(K_{n,3}) = \left(\frac{1}{2} + \kappa_3 \right) \ln^2(n) + \left(-10\zeta(3) + 2\zeta(2)\gamma + \gamma \right) \ln(n) + \frac{1}{2}\gamma^2 - 10\zeta(3)\gamma + \frac{83}{10}\zeta(2)^2 + \zeta(2)\gamma^2 + \frac{1}{2}\zeta(2) + \mathrm{o}(1)$$

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

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

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*, pages 90–100, 1976.

[4] Hoang Ngoc Minh, M. Petitot, and J. Van der Hoeven. Polylogarithms and shuffle algebra. *Proceedings of FPSAC'98*, 1998.

[5] Hoang Ngoc Minh, G. Jacob, N.E. Oussous, and M. Petitot. De l'algèbre des ζ de riemann multivariées à l'algèbre des ζ de hurwitz multivariées. *Journal électronique du Séminaire Lotharingien de Combinatoire B44e*, 2001.

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