The Expected Number of Blocks in an Ordered Set Partition of n objects is $n/log(4) + O(1)$, its Variance is $(n/log(4))(1/log(4) - 1/2) + O(1)$, and It is Asymptotically Normal! (An Experimental-Mathematical Proof)

Shalosh B. EKHAD[1]

The number of ordered set partitions of n objects into k blocks is $k!S(n,k)$, (where $S(n,k)$ are the Stirling Numbers of the second kind), so the generating polynomial for *all* ordered set partitions, according to the number of blocks is:

$$P_n(t) := \sum_{k=0}^{n} k!S(n,k)t^k \quad .$$

It is well-known (and easy to see) that

$$\sum_{n=0}^{\infty} P_n(t)\frac{x^n}{n!} = \frac{1}{1-t(e^x-1)} \quad .$$

This enabled me, using Maple, to easily crank out the first 280 terms of $\{P_n(t)\}$, then compute the first 280 terms of the sequence of averages (alias expectations) $\{P_n'(1)/P_n(1)\}$, and variance, then *guess*, using Maple's `LeastSquares`, estimates for them (as linear expressions in n), then let Maple `identify` the coefficients of n in these estimates as those in the title, ($n/log(4) + O(1) = (0.72134752044448\ldots) \cdot n + O(1)$ for the expectation and $(n/log(4))(1/log(4) - 1/2) + O(1) = (0.159668485029161\ldots) \cdot n + O(1)$ for the variance) then, by taking higher ordered moments (by further differentiations and plugging-in $t = 1$) estimate expressions for the so-called *alpha*-coefficients, and deduce that the distribution is *indeed* (experimentally!) asymptotically normal.

See `http://www.math.rutgers.edu/~zeilberg/tokhniot/oOSP240` for the Maple output, that was generated by running the input file `http://www.math.rutgers.edu/~zeilberg/tokhniot/inOSP240` using the Maple program `http://www.math.rutgers.edu/~zeilberg/tokhniot/OSP`.

Rigorous mathematicians may consider the assertions in this article as *mere conjectures*, and they are more than welcome to "prove" them "rigorously", but I have better things to do!

Postscript by D. Zeilberger

Shalosh may have better things to do, but the conjectured asymptotic expressions for the expectation and the variances are easy exercises. Using similar methods, one can easily (and Shalosh can be taught to do it automatically!) get *explicit* asymptotic expressions for each moment, from

---

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `c/o zeilberg at math dot rutgers dot edu` , `http://www.math.rutgers.edu/~zeilberg/ekhad.html` . June 28, 2011. Exclusively published in The Personal Journal of S.B. Ekhad and D. Zeilberger (`http://www.math.rutgers.edu/~zeilberg/pj.html`) . Supported in part by the NSF.

which one should deduce asymptotic expressions for the moments-about-the-mean, and should be able to prove that these (normalized) moments converge to those of the normal distribution, up to any prescribed moment. What I don't know how to do yet is how to do it for *every* moment (i.e. derive a formula for α_r for symbolic r), but I believe that even this can be taught to Shalosh (by using induction on r).

Anyway,

$$\sum_{n=0}^{\infty} \frac{P_n(1)}{n!} z^n = \frac{1}{2-e^z} \quad .$$

Since the smallest root (in absolute value) of $e^z - 2 = 0$ is $z = \log 2$, by "partial fraction", $P_n(1)/n!$ is asymptotic to $(1/2) \cdot (1/\log(2))^{n+1}$.

Since

$$\sum_{n=0}^{\infty} \frac{P_n'(1)}{n!} z^n = \frac{1}{(2-e^z)^2} - \frac{1}{2-e^z} \quad ,$$

we get that $P_n'(1)/n!$ is asymptotic to $(n/4) \cdot (1/\log(2))^{n+2}$, and hence $P_n'(1)/P_n(1)$ is asymptotic to $n/(2(\log 2)) = n/\log(4)$.

Similarly for the variance and beyond.