ASYMPTOTIC NORMALITY AND COMBINATORIAL STATISTICS

JOSH SWANSON

Contents

(These are lecture notes for a two-part lecture series given in the informal combinatorics seminar at the University of Washington on April 13th and 20th, 2017. The exposition in part follows some of [BÍ5, Ch. 3], [Pet75], [HZ15], [CJZ12a].)

1. Probability Summary

Definition 1.1. Let X be a real-valued random variable with density function $f(x)$. Recall that

$$\mathbb{E}[g(X)] = \begin{cases} \int_{\mathbb{R}} g(x)f(x)\,dx & \text{if } X \text{ is continuous} \\ \sum_{x=-\infty}^{\infty} g(x)f(x) & \text{if } X \text{ is discrete.} \end{cases}$$

Definition 1.2. The *mean* and *variance* of X are

$$\mu := \mathbb{E}[X]$$
$$\sigma^2 := \mathbb{E}[(X-\mu)^2].$$

More generally, the *dth moment* and *dth central moment* of X are, respectively,

$$\mathbb{E}[X^d] \qquad \text{and} \qquad \mathbb{E}[(X-\mu)^d],$$

where for our purposes each moment will be finite. The *moment-generating function* of X is

$$M(t) := \sum_{d=0}^{\infty} \mathbb{E}[X^d] \frac{t^d}{d!}.$$

Remark 1.3. Note that $M(t)$ involves non-central moments. We can "re-center" moments as follows:

$$\mathbb{E}[(X-\mu)^d] = \sum_{k=0}^{d} \binom{d}{k} \mathbb{E}[X^k](-\mu)^{d-k}.$$

Typically one is actually interested in central moments.

Date: April 20, 2017.

Remark 1.4. Note also that, when X is continuous,

$$\begin{aligned}\sum_{d=0}^{\infty} \mathbb{E}[X^d] \frac{t^d}{d!} &= \sum_{d=0}^{\infty} \left(\int_{\mathbb{R}} x^d f(x)\, dx \right) \frac{t^d}{d!} \\ &= \int_{\mathbb{R}} \left(\sum_{d=0}^{\infty} \frac{(xt)^d}{d!} \right) f(x)\, dx \\ &= \int_{\mathbb{R}} e^{xt} f(x)\, dx.\end{aligned}$$

In general (continuous, discrete, whatever), we have

$$M(t) = \mathbb{E}[e^{Xt}],$$

assuming X has moments of all orders.

Remark 1.5. The *characteristic function* of X is

$$v(t) := \mathbb{E}[e^{iXt}].$$

This always exists for all $t \in \mathbb{R}$ without any convergence issues. Sometimes when working with $M(t)$ it's helpful to really have $v(t)$ in mind, i.e. one imagines evaluating $M(t)$ on the imaginary axis. The characteristic function of a continuous random variable is the usual Fourier transform of its density function. At a high level, one often tries to gain insight into a random variable by manipulating its characteristic function, often tortuously, often using contour integrals and many estimates.

Example 1.6. What is the moment generating function of $X := \mathcal{N}(\mu, \sigma^2)$? Recall that

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right).$$

A standard computation gives

$$M(t) = \mathbb{E}[e^{Xt}] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2 \right).$$

The dth central moment of $\mathcal{N}(\mu, \sigma^2)$ is the dth moment of $\mathcal{N}(0, \sigma^2)$, so

$$\exp\left(\frac{1}{2}\sigma^2 t^2 \right) = \sum_{n=0}^{\infty} \frac{1}{2^{2n} n!} \sigma^{2n} t^{2n} = \sum_{d=0}^{\infty} \mu_d \frac{t^d}{d!}.$$

It follows that

$$\mu_d = \begin{cases} 0 & \text{if } d \text{ is odd} \\ \sigma^d (d-1)!! & \text{if } d \text{ is even.} \end{cases}$$

Definition 1.7. The *cumulant generating function* of X is

$$K(t) := \log M(t) = \log \mathbb{E}[e^{Xt}].$$

The *cumulants* $\kappa_1, \kappa_2, \ldots$ of X are defined by

$$K(t) := \sum_{d=1}^{\infty} \kappa_d \frac{t^d}{d!}.$$

Example 1.8. The cumulant generating function of $\mathcal{N}(\mu, \sigma^2)$ is

$$\log \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2 \right) = \mu t + \frac{1}{2}\sigma^2 t^2,$$

so the cumulants are $\kappa_1 = \mu, \kappa_2 = \sigma^2, \kappa_3 = \kappa_4 = \cdots = 0$.

Remark 1.9. In general we have $\kappa_1 = \mu$, $\kappa_2 = \sigma^2$, and $\kappa_3 = \mu_3$, though the higher cumulants and the higher central moments in general differ. While cumulants may initially be less intuitive than central moments, they're often "better." Some examples:

- The second and higher cumulants of X agree with those for $X - \mu$.

- The cumulants of the sum of independent random variables is the sum of the cumulants.
- The cumulants of a normal distribution are much simpler than its central moments.

Remark 1.10. The (non-central) moments and cumulants of X are related by

$$\mathbb{E}[X^d] = \sum_{\lambda \vdash d} \frac{d!}{z_\lambda} \frac{\kappa_{\lambda_1} \cdots \kappa_{\lambda_k}}{(\lambda_1 - 1)! \cdots (\lambda_k - 1)!}.$$

We can replace the left-hand side with μ_d if we replace κ_1 with 0, leaving the other terms the same. For instance, at $d = 3$ we have

$$\mathbb{E}[X^3] = \frac{3!}{3 \cdot 1!} \frac{\kappa_3}{2!} + \frac{3!}{2 \cdot 1 \cdot 1! \cdot 1!} \frac{\kappa_2 \kappa_1}{1! \cdot 0!} + \frac{3!}{1^3 \cdot 3!} \frac{\kappa_1^3}{0!^3} = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3,$$

hence (again)

$$\mu_3 = \kappa_3.$$

Example 1.11. Fix n and let X measure the length of elements in S_n taken uniformly at random. That is, the density function for X is

$$f(k) = \#\{w \in S_n : \ell(w) = k\}/n!.$$

Recall that the length generating function for S_n is

$$S_n^\ell(q) := \sum_{\sigma \in S_n} q^{\ell(\sigma)} = [n]_q! = \prod_{i=1}^{n} (1 + q + \cdots + q^{i-1}).$$

Note that

$$\begin{aligned} M(t) = \mathbb{E}[e^{Xt}] &= \sum_{k=0}^{\infty} e^{kt} f(k) \\ &= \sum_{k=0}^{\infty} f(k)(e^t)^k = \frac{1}{n!} S_n^\ell(e^t). \end{aligned}$$

That is, $S_n^\ell(q)$ can be thought of as essentially equivalent to the moment generating function. Hence the moment-generating function for the length statistic on S_n is simply

$$M(t) = \frac{[n]_{e^t}!}{n!}.$$

One approach for computing the central moments of X is to compute the cumulant generating function $K(t) = \log M(t)$.

Remark 1.12. To reiterate, we have the following correspondence:

combinatorial generating function $\leftrightarrow$ moment-generating function.

Identities involving one can be interpreted as identities for the other.

Example 1.13. Consider the uniform discrete distribution on $[0, n-1]$, i.e.

$$f(k) := \begin{cases} \frac{1}{n} & \text{if } k = 0, 1, \ldots, n-1 \\ 0 & \text{otherwise.} \end{cases}$$

The moment-generating function is

$$M(t) = \frac{1}{n} \left. \frac{q^n - 1}{q - 1} \right|_{q=e^t},$$

so the cumulant-generating function is

$$K(t) = \log\left(\frac{e^{nt} - 1}{e^t - 1}\right) - \log n.$$

Differentiating and simplifying gives

$$K'(t) = \frac{n}{e^{nt} - 1} - \frac{1}{e^t - 1} + n - 1.$$

Recall that

$$\frac{t}{e^t - 1} = \sum_{m=0}^{\infty} B_m \frac{t^m}{m!},$$

where B'_m is the mth Bernoulli number (where $B'_0 = 1, B'_1 = -\frac{1}{2}, B'_2 = \frac{1}{6}$). Hence

$$K'(t) = (n-1) + \sum_{m=0}^{\infty} \frac{B_m}{m!}(n^m - 1)t^{m-1},$$

so that

$$K(t) = \sum_{m=1}^{\infty} \frac{B_m}{m}(n^m - 1)\frac{t^m}{m!} + (n-1)t.$$

The t coefficient is $\kappa_1 = \mu = \frac{n-1}{2}$. An alternate convention for the Bernoulli numbers uses $B_1 = \frac{1}{2}$ with the rest unchanged. Hence

$$\kappa_d = \frac{B_d}{d}(n^d - 1).$$

Example 1.14. What are the cumulants of the length statistic on S_n? Since

$$M(t) = \frac{[n]_{e^t}!}{n!} = \prod_{i=1}^{n} \frac{[i]_{e^t}}{i},$$

we have

$$K(t) = \log M(t) = \sum_{i=1}^{n} \log \frac{[i]_{e^t}}{i} = \sum_{d=1}^{\infty} \frac{B_d}{d} \sum_{i=1}^{n} (i^d - 1)\frac{t^d}{d!}.$$

Hence the cumulants are given by

$$\kappa_d = \frac{B_d}{d} \sum_{i=1}^{n} (i^d - 1).$$

This yields an explicit formula for the central moments μ_d as a sum over partitions of d of terms involving products of Bernoulli numbers.

2. Asymptotic Normality

Definition 2.1. Given a real-valued random variable X, the corresponding *normalized* random variable is

$$X^* := \frac{X - \mu}{\sigma}.$$

Note that X^* has mean 0 and variance 1. The central moments μ_d of X are related to the (central) moments μ_d^* of X^* by

$$\mu_d^* = \frac{\mu_d}{\sigma^d}.$$

Similarly

$$\kappa_d^* = \frac{\kappa_d}{\sigma^d}.$$

Definition 2.2. Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables. Let X_n^* have cumulative distribution function $\Phi_n(x)$. Say $N(0,1)$ has cumulative distribution function

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt.$$

We say that $X_1, X_2, \ldots$ is *asymptotically normal* if, for all $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \Phi_n(x) = \Psi(x).$$

Theorem 2.3. *Suppose $X_1, X_2, \ldots$ is a sequence of real-valued random variables with moments of all orders. Suppose X_n^* has dth central moment $\mu_{d;n}^*$. If for each d,*

$$\lim_{n\to\infty} \mu_{d;n}^* = \nu_d$$

where ν_d is the dth central moment of $N(0,1)$, then $X_1, X_2, \ldots$ is asymptotically normal.

Remark 2.4. One may check cumulants instead of moments. That is, if $\kappa^*_{d;n}$ is the dth cumulant of X^*_n, it suffices to show that for each d

$$\lim_{n\to\infty} \kappa^*_{d;n} = \begin{cases} 0 & \text{if } d = 1 \\ 1 & \text{if } d = 2 \\ 0 & \text{if } d \geq 3. \end{cases}$$

Corollary 2.5. *The length of elements of S_n is asymptotically normal.*

Proof. We have

$$\kappa^*_{d;n} = \frac{1}{\sigma^d}\frac{B_d}{d}\sum_{i=1}^{n}(i^d - 1)$$

where

$$\sigma^2 = \frac{1}{12}\sum_{i=1}^{n}(i^2 - 1).$$

Hence $\kappa^*_{d;n} = O(n^{d+1-3d/2}) \to 0$ for $d \geq 3$, as required. □

Remark 2.6. Asymptotic normality shows up all the time with combinatorial statistics. Some examples:

Statistic	Set	Generating Function	Reference
Length/major index	S_n	$[n]_q!$	[Fel45], [Gon44]; see [B1́5, Thm. 3.3.4]
Length/major index	S_n/S_α, words content α	$\binom{n}{\alpha}_q$	Mann; Whitney; Diaconis; Canfield-Janson-Zeilberger; see [CJZ12a] and corrigendum
Major index	SYT(λ)	$q^{b(\lambda)} \frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q}$	Billey-Konvalinka-Swanson (forthcoming)
# parts	partitions of n with distinct parts	$\prod_{m=1}^{\infty}(1 + xy^m)$	[EL41]; see [B1́5, Ex. 3.5.2]
# elements	subsets of n	$(1+q)^n$	Classical
# left-to-right minima; # cycles	S_n	$\prod_{i=0}^{n-1}(q+i)$	[Fel45], [Gon44]; see [B1́5, Thm. 3.3.2]
# blocks	set partitions of $[n]$	$\sum_k S(n,k)q^k$	[Har67]; see [B1́5, Thm. 3.4.3]

There are many other examples as well. One of the common themes is particularly good control over the roots of the moment generating function.

Remark 2.7. There are several major techniques for proving asymptotic normality results. We summarize several here.

- The method of moments (cumulants) from last time.
- Central limit-type theorems.
  - Averages of iid random variables (e.g. fair coin flips; $\frac{1}{2^n}(1+q)^n$).
  - Sums of independent random variables (e.g. Berry-Esseen theorem, see [B1́5, Thm. 3.2.4]; inversion number on S_n by sum of Lehmer code).
  - A million zillion variations.
- Control of roots of combinatorial generating functions $f(q) = \sum_{i=0}^{n} a_i q^i$ (estimating the a_i).
  - Real, non-negative roots. (Sum of Bernoulli's. See e.g. [Pit97]; works for instance for binomial coefficients, Eulerian numbers.)

  - Complex norm 1 roots. (See [HZ15]; works for instance for Gaussian binomial coefficient polynomials.)
- Direct estimates of the characteristic function under Fourier inversion.

3. Gaussian Multinomials

Next we'll summarize results of Hwang-Zacharovas [HZ15] and Canfield-Janson-Zeilberger [CJZ12a] which shed light on, among other things, the distribution of coefficients of q-Gaussian multinomials.

Theorem 3.1 ([HZ15, Thm. 1.1])**.** *Let $\{X_n\}$ be a sequence of discrete random variables whose probability generating functions $f_n(q) := \mathbb{E}[q^{X_n}]$ are polynomials of degree n with all roots ρ_j lying on the unit circle $|\rho_j| = 1$. Then:*

- $1 \leq \mathbb{E}[(X_n^*)^4] < 3$*;*
- $\{X_n\}$ *is asymptotically normal if and only if*

$$\lim_{n\to\infty} \mathbb{E}[(X_n^*)^4] = 3.$$

- $\{X_n\}$ *is "asymptotically Bernoulli" (p.g.f. $\frac{1}{2}(q^{-1} + q^1)$) if and only if*

$$\lim_{n\to\infty} \mathbb{E}[(X_n^*)^4] = 1.$$

Example 3.2. (Following [HZ15], discussion after Theorem 1.1.) Letting X_n be the length statistic on S_n, the probability generating functions are $f_n(q) = [n]_q!/n!$, which satisfy the hypotheses of the theorem (well, we have to fiddle with the degrees). A "straightforward computation" gives

$$\mu_{4;n}^* = 3 - \frac{9(6n^2 + 15n + 16)}{25n(n-1)(n+1)} \to 3.$$

Hence we again see that length on S_n is asymptotically normal.

Example 3.3. (Following [HZ15, §4.1].) One may check that $\mu_{4;n}^* \to 3$ if and only if $\frac{\kappa_{4;n}}{\kappa_{2;n}^2} \to 0$, which is frequently easier to check. For instance, suppose we have a sequence of combinatorial generating functions

$$f_n(q) = \frac{\prod_{i=1}^{N(n)}(1 - q^{b_i})}{\prod_{i=1}^{N(n)}(1 - q^{a_i})} \in \mathbb{N}[q].$$

Then

$$\kappa_{d;n} = (-1)^d \frac{B_d}{d} \sum_{1 \leq i \leq N(n)} (b_i^d - a_i^d),$$

so the asymptotic normality condition becomes

$$\lim_{n\to\infty} \frac{\kappa_{n;4}}{\kappa_{n;2}^2} = \frac{144}{120} \lim_{n\to\infty} \frac{\sum_{i=1}^{N(n)}(b_i^4 - a_i^4)}{(\sum_{i=1}^{N(n)}(b_i^2 - a_i^2))^2} = 0.$$

Now instead take $f_n(q) = \binom{n}{\alpha^{(n)}}_q$ for a sequence of partitions $\alpha^{(n)}$. Using analogous computations, the corresponding coefficients are asymptotically normal if

$$\alpha_1^{(n)} \to \infty \qquad \text{and} \qquad \sum_{i \geq 2} \alpha_i^{(n)} \to \infty.$$

Hwang-Zacharovas say these conditions appeared earlier in [CJZ12a], though only the second one actually appears. Canfield-Janson-Zeilberger note that the second condition is necessary for asymptotic normality. See also the corrigendum [CJZ12b] for more of the history of this result.

4. Local Limit Theorems

Definition 4.1. (Informal.) A sequence of real-valued random variables $X_1, X_2, \ldots$ satisfies a *local limit theorem* if the we have some explicit, uniform upper bound on the difference between the density function of X_n and a corresponding normal density function.

We'll continue the theme of analyzing coefficients of $\binom{n}{\alpha}_q$ and discuss the local limit theorem in [CJZ12a] in detail.

Definition 4.2. Pick a (weak) composition $\alpha \vDash N$. Let W_α denote the set of words of content α. Let M_α be the inversion statistic on W_α. The corresponding probability generating function is $f_\alpha(q) := \binom{n}{\alpha}_q / \binom{n}{\alpha}$.

Let a^* be the maximum of α and let $N_* := N - a^*$. For instance, if $\alpha_1 \geq \alpha_2 \geq \cdots$, then $a^* = \alpha_1$ and $N_* = \alpha_2 + \alpha_3 + \cdots$.

Canfield et al. give a conjectured local limit theorem for arbitrary M_α.

Conjecture 4.3. *Uniformly for all α and all integers k,*

$$\mathbb{P}[M_\alpha = k] = \mathcal{N}(\mu, \sigma^2)(k) + O(1/N_*)$$

where μ, σ^2 are the mean and variance of M_α, and $\mathcal{N}(\mu, \sigma^2)$ is corresponding normal density function.

They are able to prove the following slightly weaker result.

Theorem 4.4 ([CJZ12a, Thm. 4.5])**.** *There exists a positive constant c such that for every C, the following is true. Uniformly for all α such that $a^* \leq Ce^{cN_*}$ and all k, the estimate in the conjecture holds.*

Example 4.5. When $\alpha = (1^N)$, M_α is the inversion statistic on S_N, $a^* = 1$, and $N_* = N - 1$. The exponential inequality trivially holds, giving an $O(1/N)$ error bound on the normal estimate for counting inversions. The exponential inequality seems to be a rather mild constraint in most situations.

Remark 4.6. The proof of the theorem is lengthy and quite technical in places (at least to a non-probabalist). Nonetheless it is instructive since some of the techniques are very common when proving local limit theorems.

Observation 4.7. (Cauchy integral formula.) Suppose $f(z) = \sum_{k=0}^n a_k z^k$. Then

$$a_k = \frac{1}{2\pi i} \int_C f(z) \frac{dz}{z^{k+1}},$$

where C is a positively oriented contour around 0. More concretely,

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-ik\theta} \, d\theta.$$

Hence good estimates for the integral give good estimates for a_k.

Remark 4.8. Suppose $f(z)$ is the probability generating function of a discrete random variable X on a finite set. Then $f(e^{it}) = \mathbb{E}[e^{iXt}]$ is the characteristic function of X defined last time. Recall that the characteristic function is in general the Fourier transform of the probability density function (if it exists). In this way the Cauchy integral formula is allowing us to use Fourier inversion and knowledge of the characteristic function to deduce properties of the distribution X and its density function.

Observation 4.9. By continuity, $f(e^{i\theta}) \approx 1$ for $\theta \approx 0$. Moreover, one expects $f(e^{i\theta})$ to suffer significant "destructive interference" away from $\theta = 0$. So, one may break the integral up into the small θ regime and the large θ regime, using separate estimates in each case.

For asymptotically normal distributions, one expects the "dominant" small θ contribution to give rise to the normal behavior, while deviations from normality arise from the more subtle and smaller large θ contributions.

Example 4.10. By Taylor expansion with the product formula for $f_\alpha(e^{i\theta})$, we have [CJZ12a, (3.4)]:

$$\text{if } |\theta| \leq \frac{1}{N}, \qquad \text{then} \qquad f_\alpha(e^{i\theta}) = \exp\left(i\mu\theta - \frac{\sigma^2\theta^2}{2} + O(N^4 N_* \theta^4)\right).$$

Recall that if $X = \mathcal{N}(\mu, \sigma^2)$, then

$$\mathbb{E}[e^{itX}] = \exp\left(i\mu\theta - \frac{\sigma^2\theta^2}{2}\right).$$

Thus the characteristic function of M_α agrees with that for $\mathcal{N}(\mu, \sigma^2)$ for small θ. This estimate is actually powerful enough to show that the characteristic function of M_α^* tends to that of $N(0,1)$ as $N_* \to \infty$, which is the heart of their proof of asymptotic normality (see the end of their proof of Theorem 1.2).

They proceed to break the integral into two regimes as follows.

Lemma 4.11 ([CJZ12a, Lemma 4.1])**.** *There exists a constant $\tau > 0$ such that for all α,*

$$\textit{if } |\theta| \leq \frac{\tau}{N} \qquad \textit{then} \qquad |f_\alpha(e^{i\theta})| \leq \exp\left(-\frac{\sigma^2\theta^2}{4}\right).$$

Using this estimate, they are able to bound the error in the normal approximation as follows.

Lemma 4.12 ([CJZ12a, Lemma 4.2])**.** *Uniformly for all α and k,*

$$|\mathbb{P}[M_\alpha = k] - \mathcal{N}(\mu, \sigma^2)(k)| \leq \int_{\tau/N}^{\pi} |f_\alpha(e^{i\theta})|\, d\theta + O\left(\frac{1}{\sigma N_*}\right).$$

Remark 4.13. Hence the local limit theorem reduces to finding a sufficiently tight bound on the modulus of $f_\alpha(e^{i\theta})$ away from 0. They conjecture that

$$f_\alpha(e^{i\theta}) \leq O\left(\frac{1}{\sigma^3\theta^3}\right), \qquad 0 < \theta \leq \pi,$$

at least for $N_* \geq 6$, which is sufficient to prove the bound in the local limit theorem. They are unable to prove this, unfortunately, and instead show the following.

Lemma 4.14 ([CJZ12a, Lemma 4.4])**.** *There exists a constant $c > 0$ such that*

$$\textit{if } \frac{\tau}{N} \leq |\theta| \leq \pi, \qquad \textit{then} \qquad |f_\alpha(e^{i\theta})| \leq \exp(-cN_*).$$

Remark 4.15. The proof is rather involved. It uses a two-variable analogue of $f_\alpha(z)$; estimates of exponentials in terms of rational functions; various trig identities and inequalities; the Cauchy integral formula; Stirling's approximation; various estimates of pieces of the product formula for $f_\alpha(z)$; and more. The local limit theorem above follows very quickly from these pieces.

Remark 4.16. Local limit theorems can feel more satisfying than asymptotic normality results. They can also prove (often partial) statements of independent interest. The preceding local limit theorem implies the following log-concavity result.

Theorem 4.17 ([CJZ12a, Thm. 4.6])**.** *Let*

$$c_k := [q^k]\binom{2n}{n}_q = \binom{2n}{n}\mathbb{P}[M_{(n,n)} = k].$$

Then for each constant C we have some n_0 such that for all $n \geq n_0$ and $|k - \mu| \leq C\sigma$,

$$c_k^2 \geq c_{k-1}c_{k+1}.$$

Remark 4.18. That is, the number of inversions of $W_{(n,n)}$ is log-concave for n sufficiently large sufficiently close to the mean. In fact, log-concavity for the c_k fails sometimes, so some hypothesis is necessary.

Remark 4.19. Louchard-Prodinger [LP03] give more refined local limit theorems for the number of inversions on S_n using the "Saddle point method." This involves using Cauchy's integral formula with a contour passing through a saddle point of the probability generating function. The argument breaks up the analysis into different regimes, e.g. near the peak and in the tails.

References

[BÍ5] Miklós Bóna, editor. *Handbook of enumerative combinatorics*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, 2015.

[CJZ12a] E. Rodney Canfield, Svante Janson, and Doron Zeilberger. Corrigendum to "The Mahonian probability distribution on words is asymptotically normal" [Adv. in Appl. Math. 46 (1–4) (2011) 109–124] [mr2794017]. *Adv. in Appl. Math.*, 49(1):77, 2012.

[CJZ12b] E. Rodney Canfield, Svante Janson, and Doron Zeilberger. Corrigendum to "The Mahonian probability distribution on words is asymptotically normal" [Adv. in Appl. Math. 46 (1–4) (2011) 109–124] [mr2794017]. *Adv. in Appl. Math.*, 49(1):77, 2012.

[EL41] Paul Erdös and Joseph Lehner. The distribution of the number of summands in the partitions of a positive integer. *Duke Math. J.*, 8:335–345, 1941.

[Fel45] W. Feller. The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.*, 51:800–832, 1945.

[Gon44] V. L. Goncharov. From the realm of combinatorics. *Izv. Akad. Nauk SSSR*, 8:3–48, 1944.

[Har67] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[HZ15] Hsien-Kuei Hwang and Vytas Zacharovas. Limit distribution of the coefficients of polynomials with only unit roots. *Random Structures Algorithms*, 46(4):707–738, 2015.

[LP03] Guy Louchard and Helmut Prodinger. The number of inversions in permutations: a saddle point approach. *J. Integer Seq.*, 6(2):Article 03.2.8, 19, 2003.

[Pet75] V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[Pit97] Jim Pitman. Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combin. Theory Ser. A*, 77(2):279–303, 1997.