A combinatorial identity for a problem in asymptotic statistics

Citation for published version (APA):

DOI:
10.2298/AADM0901064A

Document status and date:
Published: 01/01/2009

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 11. Sep. 2020
A COMBINATORIAL IDENTITY FOR A PROBLEM IN ASYMPOTIC STATISTICS

Hansjörg Albrecher, Jozef L. Teugels, Klaus Scheicher

Let \((X_i)_{i \geq 1}\) be a sequence of positive independent identically distributed random variables with regularly varying distribution tail of index \(0 < \alpha < 1\) and define

\[ T_n := \frac{X_1^2 + X_2^2 + \cdots + X_n^2}{(X_1 + X_2 + \cdots + X_n)^2}. \]

In this note we simplify an expression for \(\lim_{n \to \infty} E(T_n^k)\), which was obtained by Albrecher and Teugels: Asymptotic analysis of a measure of variation. Theory Prob. Math. Stat., 74 (2006), 1–9, in terms of coefficients of a continued fraction expansion. The new formula establishes an unexpected link to an enumeration problem for rooted maps on orientable surfaces that was studied in Arquès and Béraud: Rooted maps of orientable surfaces, Riccati's equation and continued fractions. Discrete Mathematics, 215 (2000), 1–12.

1. INTRODUCTION

Let \((X_i)_{i \geq 1}\) be a sequence of positive independent identically distributed (i.i.d.) random variables with distribution function \(F\). Assume that \(F\) satisfies

\[ 1 - F(x) \sim x^{-\alpha} \ell(x) \quad \text{for} \quad x \uparrow \infty, \]

where \(\alpha > 0\) and \(\ell(x)\) is slowly varying, i.e.

\[ \lim_{x \to \infty} \frac{\ell(tx)}{\ell(x)} = 1 \quad \text{for all} \quad t > 0 \]

(cf. e.g. Bingham, Goldie and Teugels [4]). Relation (1) appears as the essential condition for the domain of attraction problem in extreme value theory. Note that the expectation \(E(X_1^{\beta})\) is finite if \(\beta < \alpha\) but infinite whenever \(\beta > \alpha\), so...
that distributions of type (1) play a crucial role for modeling extremely heavy-tailed data sets in statistics. Now define

\[ T_n := \frac{X_1^2 + X_2^2 + \cdots + X_n^2}{(X_1 + X_2 + \cdots + X_n)^2}. \]

This statistic turns out to have interesting properties in particular for the case \(0 < \alpha < 1\) (i.e. when \(E(X_i)\) does not exist). Utilizing Karamata theory of regularly varying functions, the following asymptotic limit for arbitrary moments of \(T_n\) was shown in Albrecher and Teugels [1]:

**Theorem 1.** If \(F\) satisfies (1) with \(0 < \alpha < 1\), then for all \(k \geq 1\)

\[ \lim_{n \to \infty} E(T_n^k) = \frac{k!}{(2k - 1)!} \sum_{r=1}^{k} \frac{\alpha^{r-1}}{r^r \Gamma(1 - \alpha)^r} G(r, k), \]

where \(G(r, k)\) is the coefficient of \(x^k\) in the polynomial

\[ \left( \sum_{j=1}^{k-r+1} \frac{j! (2j - \alpha)}{(2j-\alpha)! x^j} \right)^r. \]

The first few moments are given by

\[ \lim_{n \to \infty} E(T_n) = (1 - \alpha), \]
\[ \lim_{n \to \infty} E(T_n^2) = \frac{1}{3} (1 - \alpha)(3 - 2\alpha), \]
\[ \lim_{n \to \infty} E(T_n^3) = \frac{1}{15} (1 - \alpha)(15 - 17\alpha + 5\alpha^2), \]
\[ \lim_{n \to \infty} E(T_n^4) = \frac{1}{105} (1 - \alpha)(105 - 155\alpha + 79\alpha^2 - 14\alpha^3) \quad \text{etc.} \]

As the right-hand side is finite for each \(k\), this result gives rise to a convenient and simple method to both estimate the extreme value index and the finiteness of the mean of a distribution in the domain of attraction of a stable law from a data set of independent and identically distributed observations. Moreover, \(T_n\) is closely connected to the study of the sample coefficient of variation and the sample dispersion (cf. Albrecher, Ladoucette and Teugels [2]).

Given the structure of formula (3), it is natural to ask for a simpler representation of its right-hand side through generating functions. The purpose of this note is to establish such a relationship by identifying the right-hand side as a polynomial in \(\alpha\) with coefficients determined by a bivariate generating function of continued fraction type. Surprisingly, the result turns out to be intimately connected to the solution of an enumeration problem for rooted maps on orientable surfaces as dealt with in Arquès and Béraud [3].
2. AN ALTERNATIVE REPRESENTATION

Theorem 1 can be reformulated in the following way:

**Theorem 2.** If $F$ satisfies (1) with $0 < \alpha < 1$, then for all $k \geq 1$

\[
\lim_{n \to \infty} E(T_n^k) = \frac{1}{k!} \sum_{\ell=1}^{k} (\prod_{\ell=1}^{k} (2\ell - 1)) \sum_{j=0}^{k} (-1)^j a_{jk} \alpha^j,
\]

where $a_{jk}$ is the coefficient of $t^j z^k$ in the expansion of the continued fraction

\[
M(t, z) = \frac{1}{1 - \frac{(t+1)z}{1 - \frac{(t+2)z}{1 - \frac{(t+3)z}{1 - \ldots}}}}.
\]

**Proof.** Define $\tau_k := \lim_{n \to \infty} E(T_n^k)$ . From Theorem 1 we know that

\[
\frac{(2k - 1)!}{k!} \tau_k = \frac{1}{\alpha} \sum_{r=1}^{k} \frac{1}{r} W_{\alpha}(r, k)
\]

where $W_{\alpha}(r, k)$ is the coefficient of $y^k$ in the expansion of

\[
\left( \sum_{s=1}^{\infty} \frac{\alpha \Gamma(2s - \alpha)}{\Gamma(1 - \alpha)} \frac{y^s}{s!} \right)^r.
\]

Turning to generating functions, we obtain

\[
\sum_{k=1}^{\infty} \frac{(2k - 1)!}{k!} \tau_k x^k = \frac{1}{\alpha} \sum_{r=1}^{\infty} \frac{1}{r} \sum_{k=r}^{\infty} W_{\alpha}(r, k)x^k
\]

\[
= \frac{1}{\alpha} \sum_{r=1}^{\infty} \frac{1}{r} \left( \sum_{s=1}^{\infty} \frac{\alpha \Gamma(2s - \alpha)}{\Gamma(1 - \alpha)} \frac{x^s}{s!} \right)^r
\]

\[
= \frac{1}{\alpha} \ln \left( \sum_{n=1}^{\infty} \frac{\alpha \Gamma(2n - \alpha)}{\Gamma(1 - \alpha)} \frac{x^n}{n!} \right)
\]

\[
= \frac{1}{t} \ln \left( \sum_{n=0}^{\infty} \frac{\Gamma(2n + t)}{\Gamma(t)} \frac{x^n}{n!} \right),
\]

where $\alpha$ is replaced by $-t$. Now we would like to identify the coefficients of $x^k$ on the left-hand side of (7) as polynomials in $t$ ($\alpha$, respectively). For that purpose, we guess that

\[
\tau_k = \frac{1}{k} \prod_{\ell=1}^{k} \frac{\prod_{\ell=1}^{k} (2\ell - 1)}{2\ell - 1} \sum_{j=0}^{k} (-1)^j t^j a_{jk} \alpha^j
\]

with $a_{jk} = \frac{1}{j!k!} \mu_{jk}$.
where the terms $\mu_{jk}$ denote double partial derivatives (for $j \neq k$ twice the double partial derivatives, respectively), evaluated at zero, of some bivariate generating function $M(t, z)$, i.e.

$$M(t, z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} \sum_{j=0}^{k} \frac{t^j}{j!} \mu_{jk}$$

with $z = 2x$. If we take (8) for granted, then the generating function for the sequence $\frac{(2k-1)!}{k!} \tau_k$ can be rewritten in the form

$$\sum_{k=1}^{\infty} \frac{(2k-1)!}{k!} \tau_k x^k = \sum_{k=1}^{\infty} \frac{1}{2k} \frac{(2x)^k}{k!} \sum_{j=0}^{k} \frac{t^j}{j!} \mu_{jk},$$

From (7), (9) and (10) it follows that

$$\frac{1}{2} \int_0^x M(t, y) \frac{dy}{y} = \sum_{k=1}^{\infty} \frac{1}{2k} \frac{(2x)^k}{k!} \sum_{j=0}^{k} \frac{t^j}{j!} \mu_{jk} = \frac{1}{t} \ln \left( \sum_{n=0}^{\infty} \frac{\Gamma(2n+1) \ (x/2)^n}{\Gamma(t) \ n!} \right).$$

This finally leads to

$$tM(t, z) := M^*(t, z) = 2z \frac{\partial}{\partial z} \ln \left( \sum_{n=0}^{\infty} \frac{\Gamma(2n+y) \ (z/2)^n}{\Gamma(y) \ n!} \right).$$

But, by algebraic techniques, $M^*(t, z)$ was identified in Jackson and Visentin [5, Prop. 3.6] as the generating function for all rooted maps on orientable surfaces, without regard to genus, with respect to edges and vertices. In [3], using a topological approach, Arquès and Beraud alternatively identified this generating function as the solution of the Riccati differential equation

$$(1 - z (2t + 1)) M^*(t, z) = z M^*(t, z)^2 + z (t^2 + t) + 2z^2 \frac{\partial}{\partial z} M^*(t, z),$$

for which they gave the solution in terms of the continued fraction

$$M^*(t, z) = \frac{t}{1 - \frac{(t+1)z}{1 - \frac{(t+2)z}{1 - \frac{(t+3)z}{1 - \cdots}}}} - t.$$
which appeared by collecting all asymptotically relevant terms of the multinomial expansion of an integral representation of $E(T_k^n)$ (see [1] for details). It is somewhat surprising that the resulting counting problem in (4) is intimately connected with the problem of counting all possible orientable rooted maps of any genus for a given number of edges and vertices.

REFERENCES


Hansjörg Albrecher
Institute of Actuarial Science, Faculty HEC, University of Lausanne, Extrane Building, CH-1015 Lausanne, Switzerland
E-mail: hansjoerg.albrecher@unil.ch

Klaus Scheicher
University of Natural Resources and Applied Life Sciences, Gregor Mendel Strasse 33, 1180 Vienna, Austria
E-mail: klaus.scheicher@boku.ac.at

Jozef L. Teugels
E-mail: jef.teugels@wis.kuleuven.be

\^1Supported by the Austrian Science Fund Project P18392
\^2Supported by the FWF research projects P18079-N12 and P20089