Proceedings of the ICM2010 Satellite Conference
International Workshop on
Harmonic and Quasiconformal Mappings (HQM2010)
Editors: D. Minda, S. Ponnusamy, and N. Shanmugalingam
J. Analysis
Volume 18 (2010), 9–23

On Applications of Special Functions to Disparate Fields

Roger W. Barnard

Abstract. In this note we discuss several of our recent applications of Special Functions to disparate fields including Number Theory, Abstract Algebra, Probability and Asymptotic Statistics.

Keywords. Special Functions, Applications.

2010 MSC. 33C90.

1. Introduction

This is a survey article on the applications of Special Functions that the author has been able to provide to several members of his department in disparate fields in the last few years. The applications started a few years back (see [3]), where the author along with his former PhD student, Kendall Richards, proved that confluent hypergeometric functions, ${}_1F_1(a, b, x)$, are logarithmically concave functions of the first parameter. This answered a question of M. Gordy, economist with the Federal Reserve Board (USA), on the time threshold as to when banks should foreclose on mortgages.

We will give an outline as to what we will discuss in this paper.

In Number Theory we obtained, for C. Monico, with the help of the Wilf-Zeilberg algorithm, an analytic proof that

$$\sum_{k=0}^{n} \binom{n-3k}{k} \left(\frac{1}{2}\right)^{n-3k}$$

converges to 2/5 which enabled him to determine the density of base 3 happy numbers.

In Algebra we gave, for L. Christensen, an asymptotic estimate of when

$$\sum_{k=0}^{j+1} \binom{n-1}{k} \Big/ \binom{n}{k}$$

ISSN 0971-3611 © 2010

minimizes for $0 \le j \le n$ and a way to compute where the minimum occurs. This helped him to show that the Bass numbers for special Abelian Noetherian local rings are positive.

In Statistics we proved, with P. Hadjacostas, for the Whittaker function $W_{k,\mu}(x)$ and

$$f_1(x) = e^{2/(27x^2)} W_{1/2,1/6}(4/(27x^2))/(2x\sqrt{3\pi})$$

that $f_1(cx)/f_1(x)$ is increasing in x for $0 < c < 1$ and is decreasing in x for $c > 1$ which enabled us to show that the sample median of i.i.d. data is preferable to the sample mean from a symmetric stable distribution with an index of stability of 2/3 (using Banks' criterion).

In Asymptotic Statistics we proved, with A. Trindade, for

$$\mu_{r:n} = \frac{\sum_{k=0}^{n-r} \frac{(-1)^k}{r+k}\binom{n-r}{k}[\Psi(1+\theta(r+k)) - \Psi(1)]}{B(r, n-r+1)}$$

and

$$\lambda_{r:n} = -\log[1 - \left(\frac{r}{n+1}\right)^{1/\theta}]$$

that

$$\lim_{n\to\infty} \frac{\mu_{r:n}}{\lambda_{r:n}} = \frac{\Gamma(r+1/\theta)}{\Gamma(r)} \left(\frac{1}{r}\right)^{1/\theta}. \tag{1.1}$$

In [4] we used the precise estimate arising from the limit in (1.1) to access the large scale accuracy of David-Johnson's approximations when sampling from a generalized exponential distribution.

2. Number Theory

We begin with the definition of a happy number. Start with a positive integer. Replace the integer with the sum of the squares of its digits. Repeat the process until either you reach 1 or the process loops in a cycle which does not include 1. If the process reaches 1, the number is happy. If not (i.e., the process loops in a cycle), the number is unhappy (or sad).

As examples,

- $7 \to 49 \to 97 \to 130 \to 10 \to 1$

  $\Rightarrow$ 7 is happy

- $4 \to 16 \to 37 \to 58 \to 89 \to 145 \to 42 \to 20 \to 4$

$\Rightarrow$ 4 is unhappy

The happy numbers between 1 and 50 are 1, 7, 10, 13, 19, 23, 28, 31, 32, 44, 49. These were discussed in R. Guy's Book (2004) "Unsolved Problems in Number Theory." Guy noted that happy numbers were known in England in the 90's and probably originated in Russia in the 70's. There are several open questions discussed in Guy's book and in more than 20 papers in the last 5 years. Many of these involve generalizing to different bases. Common questions are: "How many happy numbers are there? How are they distributed? How dense?" Known results include,

- In base 10 the density is 1/7
- In base 2, all numbers are happy!
- In base 3 the density question is open.

In base 10, the first ten base 3 happy numbers are:

$$1, 3, 9, 13, 17, 23, 25, 27, 31, 35.$$

In base 3, the first ten happy numbers are:

$$1_3, 10_3, 100_3, 111_3, 122_3, 212_3, 221_3, 1000_3, 1011_3, 1022_3.$$

Chris worked on density questions. He used counting techniques to show that if

$$T(n) := \sum_{k=0}^{n} \binom{n-3k}{k} \left(\frac{1}{2}\right)^{n-3k}$$

then $T(n)$ converges to the probability of occurance as $n \to \infty$! Numerical estimates on a computer suggested that $T(n) \to 0.4$. He needed an analytic proof.

In [2] we provided this. First, using Pochhammer notation

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1)$$

it follows that

$$\binom{n}{k} = \frac{(-1)^k(-n)_k}{k!} \text{ and } (a)_{nk} = n^{nk} \prod_{j=0}^{n-1} \left(\frac{a+j}{n}\right)_k.$$

Second, the Generalized Hypergeometric Function has the form

$${}_pF_q\left(\begin{matrix} a_1, a_2, \cdots, a_p \\ b_1, b_2, \cdots, b_q \end{matrix}; x\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k x^k}{(b_1)_k \cdots (b_q)_k k!}.$$

Several manipulations give

$$T(n) = \frac{1}{2^{4n}} {}_4F_3\left(\begin{matrix} -n, -n+1/4, -n+2/4, -n+3/4 \\ -4n/3, -4n/3+1/3, -4n/3+2/3 \end{matrix}; \frac{2^{11}}{3^3}\right). \tag{2.1}$$

The difficulty is the non-uniform convergence. I mentioned to a colleague, P. Hadjicostas, that I wished I could find a finite recurrence for $T(n)$. He suggested that we could use the Wilf-Zeilberger Algorthim [7] since (2.1) is a "proper hypergeometric type" with rational coefficients. We used the W-Z algorithm in Maple to produce a recurrence relation.

The idea is to obtain a W-Z pair, i.e., a pair of discrete functions $(F(n,k), G(n,k))$ such that

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k).$$

To prove $\sum_k f(n,k) = r(n), n \geq n_0$ if $r(n) \neq 0$ we can divide by $r(n)$. Hence, we need to show

$$\sum_k F(n,k) = 1 \text{ where } F(n,k) = f(n,k)/r(n). \tag{2.2}$$

Let $\sum_k F(n,k) = S(n)$. To show $S(n) = 1$ it suffices to show

$$S(n+1) - S(n) = 0 \text{ for } n \geq n_0, \quad S(n_0) = 1.$$

To show (2.2), we create a "nice" function G such that

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k).$$

Then, sum for all k and imply $S(n+1) - S(n) = 0$.

W-Z proved if a nice G exists, then it has the form $G(n,k) = R(n,k)F(n,k)$ where $R(n,k)$ is a rational function of n and k. If G exists, it can found by a computer algorithm of W. Gosper.

To find a closed form for a finite recursion formula, we proceed as follows: If $F(n,k)$ is of "proper hypergeometric type", then there exists $G(n,k)$ such that

$$\sum_{j=0}^{J} a_j(n)F(n+j,k) = G(n,k+1) - G(n,k)$$

where the coefficients $a_j(n)$ are polynomials in n. Summing over k yields

$$\sum_{j=0}^{J} a_j(n)S(n+j) = 0.$$

J is the order of the recurrence.

Thus, knowing initial conditions$\{S(0), \cdots, S(J-1)\}$, we can use difference equation methods to solve this system. We can now determine to what $T(n)$ converges as $n \to \infty$. The W-Z algorithm gives, upon handling several parameters, that

$$16T(n) - 33T(n-1) + 24T(n-2) - 8T(n-3) + T(n-4) = 0. \tag{2.3}$$

Using difference equation methods write the characteristic equation

$$(x-1)(16x^3 - 17x + 7x - 1) = 0 = \prod_{j=0}^{3}(x - r_j)$$

where

$$r_0 = 1,\ r_1 \approx 0.29822,\ r_2, r_3 \approx 0.382139 \pm 0.252081i.$$

Then, because (key step)

$$T(n) = c_0 1^n + c_1 r_1^n + c_2 r_2^n + c_3 r_3^n$$

and in this case, $|r_j| < 1$, this implies $\lim_{n\to\infty} T(n) = c_0$.

To find c_0 with initial conditions we solve the system

$$T(k) = c_0 + c_1 r_1^k + c_2 r_2^k + c_3 r_3^k, \quad k = 0, 1, 2, 3.$$

This gives $c_0 = 2/3$. Also, it is easy to do this by hand, since c_0 is a combination of the coefficients in (2.3) Q.E.D.

3. Algebra

We now provide some background terminology to discuss our assistance to L. Christensen with his paper [5].

Let R be an abelian Noetherian local ring with field k. Let E be an injective envelope of k as a summand in every degree starting from the depth of R. The number of copies of E in degree i equals the k-vector space dimension of the cohomology module $Ext_R^i(k, R)$. Form exact sequences

$$0 \to A \to B \to C \to 0.$$

For Hom = Linear Transformations $k \to A$ form

$$0 \to Hom_R(k, A) \to Hom_R(k, B) \to Hom_R(k, C) \to ?$$

Insert into ?

$$\to Ext_R^1(k, A) \to Ext_R^1(k, B)) \to Ext_R^1(k, C) \to ?$$

Insert again

$$\to Ext_R^2(k, A) \to Ext_R^2(k, B)) \to Ext_R^2(k, C) \to \cdots$$

(infinite sequence)

Then, the functors Ext determine how far off from being exact the sequences generated by the cohomology groups are. Define $\mu^i(R) = \text{rank} Ext_k^i(k, R)$. These are known at the Bass numbers (related to Betti numbers). They form an infinite sequence of <u>invariants</u> for R. Very little is know about Bass numbers. There are several open questions! E.g., Are they all positive for general R's? Christensen proved that they eventually grow exponentially if R is a non-trivial fiber-product.

Using counting methods he showed the n^{th} term to be

$$T(n,j) = \sum_{k=0}^{j+1} \binom{n-1}{k} \Big/ \binom{n}{j} \text{ for } j = 0, 1, 2, \cdots, n.$$

Define

$$T(n) = \min_j T(n,j).$$

Note, $T(n,0) = n, T(n,n) = 2^{n-1}, T(2,1) = 1$.

For $n \geq 3$,

$$T(n,1) \leq \frac{n}{2}$$
$$T(n,j) > 1 \text{ for } 1 \leq j \leq \left[\frac{n}{2}\right]$$
$$T(n,j) \leq T(n,j+1) \text{ for } \left[\frac{n}{2}\right] \leq j \leq n-1.$$

Direct computation shows

$$T(n) = T(n, \left[\frac{n}{2}\right]) \text{ for } n \leq 9 \text{ but } T(10) = T(10,4).$$

<u>Problems</u>

Find the asymptotic value for when $T(n,j)$ takes its minimum.

Find a reasonable way to compute the minimum!

<u>Outline of Solutions</u>

First, for asymptotics, using Pochhammer notation,

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)},$$

we have

$$\binom{n}{k} = \frac{(-1)^k(-n)_k}{k!}.$$

Rewriting, we have

$$T(n,j) = \sum_{k=0}^{j+1} \frac{-(n-1)_k(-1)^k}{k!} \Big/ \frac{\Gamma(n+j)}{\Gamma(a+1)\Gamma(n-j+1)}.$$

Interchanging $k \to k-(j+1)$ and using $(a)_{n-k} = (-1)^k (a)_n/(1-a-n)_k$ yields

$$T(n,j) = \frac{(n-j)(n-j+1)}{n(j+1)} \sum_{k=0}^{\infty} \frac{\Gamma(2+j)\Gamma(n-j+1)}{\Gamma(2+j-k)\Gamma(n-j-1+k)}. \tag{3.1}$$

Use $\frac{\Gamma(n+a)}{\Gamma(n+b)} \approx n^{a-b}$ and let $j = nx$ to obtain as $j \to \infty$,

$$\begin{aligned} T(n,j) &\approx \frac{(1-x)((1-x)-1/n)}{x+1/n} \sum_{k=0}^{\infty} \left(\frac{x+2/n}{(1-x)-1/n} \right)^k \\ &\approx \frac{(1-x)^3}{x(1-2x)} = S(x) \end{aligned}$$

This gives $S'(x) = 0$ for $x = \frac{\sqrt{3}-1}{2} \approx 0.366$.

To show how to verify the numerical computations of the minimum, we show that $T(n,j)$ is a convex function of j for each n. Simplifying (3.1) gives

$$T(n,j) = \sum_{k=0}^{\infty} \frac{\Gamma(1+j)\Gamma(n-j+1)}{\Gamma(2+j-k)\Gamma(n-j-1+k)} = \sum_{k=0}^{\infty} H_n(k,j),$$

where $H_n > 0$, which implies $H_n(k,j) = e^{\log H_n(k,j)}$.

We can show $\frac{\partial^2 H_n(k,j)}{\partial^2 j} = H_n(k,j)S(n)$ where, for $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$,

$$\begin{aligned} S(n) = {} & [\Psi'(1+j) + \Psi'(n-j+1) - \Psi'(2+j-k) - \Psi'(n-j-1+k)] \\ & + [\Psi(1+j) - \Psi(n-j+1) - \Psi(2_j - k) + \Psi(n-j-1+k)]^2. \end{aligned}$$

Now use $\Psi(1+x+n) - \Psi(x) = \sum_{k=0}^{n} \frac{1}{x+k}$ and $\Psi(x-n) - \Psi(x) = -\sum_{k-0}^{n} \frac{1}{x-k}$ (take derivatives with respect to x).

$$\begin{aligned} \frac{\partial S(n)}{\partial n} &= -\sum_{m=0}^{k-3} \frac{2}{(n-j+1+m)^3} \\ &\quad - \sum_{m=0}^{k-3} \frac{2}{(n-j+1+m)^2} [\sum_{m=1}^{k-1} \frac{1}{1+j-m} + \sum_{m=1}^{k-3} \frac{1}{n-j+1+m}] \\ &\quad \leq 0. \end{aligned}$$

Hence, S is decreasing. Then, let $n \to \infty$.

$$S(\infty) = -\sum_{m-1}^{k-1} \frac{1}{(1+j-m)^2} + [\sum_{m=1}^{k-1} \frac{1}{1+j-m}]^2$$

which is greater than 0 since $\left(\sum_{j=0}^{\infty} x_j^t\right)^{1/t}$ is a decreasing function of t. Thus, $T(n,j)$ is a convex function of j for each n.

4. Statistics

In statistics we are working with P. Hadjicostas on our paper, "Banks' Criterion and Symmetric Stable Laws with Index of Stability Less Than One." (See [1]) We have proved that the sample median of i.i.d. data is preferable to the sample mean of samples from the symmetric stable distribution with index of stability of 2/3 (using Banks' Criterion).

We give some background notation. Let

- Ω = sample space of experiment (a measure space)
- $\mathcal{F}$ = σ-field of subsets of Ω = set of "events" (observed)
- $\underline{X} : \Omega \to \mathbb{R}^p$, a random vector (r.v.), the "data"
- Θ = parameter space = set of states of nature
- $\mathcal{A}$ = estimation space $\subset \mathbb{R}$
- $\delta(\underline{X}) : \Omega \to \mathcal{A}$ estimator.

Here

$$\delta_1(\underline{X}) := \tfrac{x_1+x_2+\cdots+x_n}{n} = \overline{X_n} = \underline{\text{mean}}$$

$$\delta_2(\underline{X}) := \begin{cases} \text{middle point, n odd} \\ \text{average of 2 middle points, n even} \end{cases} = \underline{\text{median}}$$

Assume P is a probability measure on $(\Omega, \mathcal{F})$. Given a function $X : \Omega \to \mathbb{R}$, its cumulative distribution function (cdf), $F_X : \mathbb{R} \to [0,1]$, is defined by

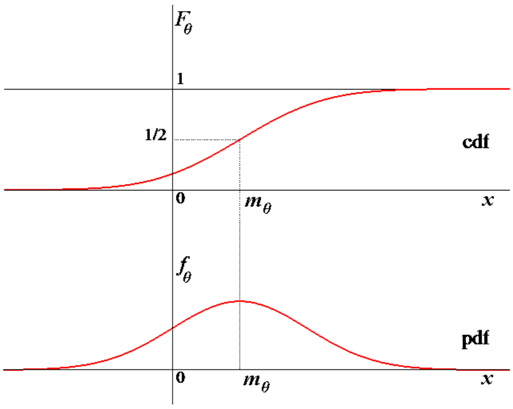
$$F_X(x) := P[X \le x] = P[\{\omega \in \Omega | X(\omega) \le x\}]$$

Note, the cdf is increasing, right continuous and satisfies

$$\lim_{x\to-\infty} F_X(x) = 0 \text{ and } \lim_{x\to\infty} F_X(x) = 1.$$

We assume for each $\theta \in \Theta$ and for each $n \in (N)$ the random variables are i.i.d. (independent identically distributed), i.e., $P_\theta[X \le x] = F_\theta(x)$ is independent of order i and has the same cdf, $F_\theta(x)$, for all $x \in \mathbb{R}$,

Here $m_\theta = F_\theta^{-1}(1/2)$; so m_θ is the theoretical median. Let $f_\theta : \mathbb{R} \to \mathbb{R}$ be a pdf, the probability density function for θ. We note that $f_\theta = \frac{dF_\theta}{dx}$. Thus, $F_\theta(x) = \int_{-\infty}^{x} f_\theta(y)dy$.



Symmetric Probability Distribution

We define the characteristic function of F_θ, $\Phi_{F_\theta} : \mathbb{R} \to \mathbb{C}$, by

$$\Phi_{F_\theta}(t) = \int_{-\infty}^{\infty} e^{itx} dF_\theta(x) = \int_{-\infty}^{\infty} e^{itx} f_\theta(x) dt$$

Note, if $y = x + a$, then $\Phi_y(t) = e^{ita}\Phi_x(t)$.

A stable distribution means that the characteristic function preserves the distribution properties, i.e.,

$$\Phi_{\sum\limits_i X_i}(t) = \prod_i \Phi_{X_i}(t).$$

Petros had proved in [6] that if, with all of the above notation and if we define a pdf, $k_\theta : \mathbb{R} \to \mathbb{R}$ where

$$k_\theta(x) = f_\theta(m_\theta + x),$$

with the assumption that for each $\theta \in \Theta$, there exists a function $q_\theta : [1, \infty) \to \mathbb{R}$ with $q_\theta(1) = 1$ and

a) $\Phi_\theta^n(t/n) = \Phi_\theta(q_\theta(n)t)$ for all $t \in \mathbb{R}$ and all odd $n \in \mathbb{N}$

b) $\Gamma(2b)q_\theta(2b-1) > 4^{b-1}\Gamma^2(b)$ for all $b > 1$

c) for each $y \geq 1, \lambda_{\theta:y} : [0, \infty) \to \mathbb{R}$ defined by

$$\lambda_{\theta:y}(s) := \frac{k_\theta(\frac{s}{q_\theta(y)})}{k_\theta(s)}$$

is non-decreasing, then, for $\epsilon > 0$,

$$P_\theta(|\overline{X}_n - m_\theta| < \epsilon) < P_\theta(|\tilde{X}_n - m_\theta| < \epsilon),$$

(This is Banks' Criterion) i.e., $\tilde{X}_n$ is preferable.

If for $0 < y < 1, \lambda_{\theta:y}$ is non-increasing, then the inequality reverses.

Now we consider a characteristic exponent $\alpha \in (0, 2]$ and a stable symmetric distribution (i.i.d.) with characteristic function $t \to e^{im_\theta t}\phi_\theta(t)$ where $\phi_\theta(t) = e^{-|t|^\alpha}, t \in \mathbb{R}$, then

a) and b) are straightforward to prove,

c) is difficult,

$\alpha = 1$ and $\alpha = 2$, Petros proved earlier, and

$\alpha = 2/3$ is a special case, since we can obtain a reasonable closed form from integration.

In the 2/3 case, the pdf, $g : (-\infty, \infty) \to \mathbb{R}$, is defined by

$$g(x) = \begin{cases} f_1(-x), x < 0 \\ \frac{3}{4\sqrt{\pi}}, x = 0 \\ f_1(x), x > 0 \end{cases},$$

where $f_1(x) = \frac{1}{2x\sqrt{3\pi}} e^{(\frac{2}{27x^2})} W_{-\frac{1}{2},\frac{1}{6}}(\frac{2}{27x^2})$ and

$$W_{\kappa,\mu}(z) = e^{-\frac{1}{2}} z^{\mu+\frac{1}{2}} U(\mu - \kappa + \frac{1}{2}, 2\mu + 1; z)$$

and $U(.,.;x)$ is a confluent hypergeometric function of the 2nd kind, i.e.,

$$U(a, b; x) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-xt} t^{a-1} (1 + t)^{b-a-1} dt.$$

Thus, for $x > 0$,

$$\begin{aligned} f_1(x) &= \frac{\sqrt[3]{2} x^{-7/3}}{9\sqrt{3\pi}} U(\frac{7}{6}, \frac{4}{3}; \frac{4}{27x^2}) \\ &= \frac{x^{-5/3}}{3\sqrt{3\pi}\sqrt[3]{2}} U(\frac{5}{6}, \frac{2}{3}; \frac{4}{27x^2}). \end{aligned}$$

Our goal is to show that $h_c : [0, \infty) \to \mathbb{R}$, defined by

$$h_c(x) = \frac{f_1(cx)}{f_1(x)}, \text{for } x > 0,$$

is increasing for $0 < c < 1$ and decreasing for $c > 1$. Set $x = \frac{2}{3\sqrt{3y}}$ and define $f_2(y) = \frac{3y^{5/6}}{4\sqrt{\pi}} U(\frac{5}{6}, \frac{7}{3}; y)$, and $K(y) = \frac{y f_2'(y)}{f_2(y)}$.

Then, noting that

$$\overline{h}_c(y) := \frac{f_2(cy)}{f_2(y)} = h_{\frac{1}{\sqrt{c}}}(\frac{2}{3\sqrt{3y}})$$

implies

$$h_c(x) = \overline{h}_{\frac{1}{c^2}}(\frac{4}{27x^2})$$

and

$$\begin{aligned}\overline{h}_c{}'(y) &= \frac{f_2(cy)}{yf_2(y)}\left[\frac{cyf_2'(cy)}{f_2(cy)} - \frac{yf_2'(y)}{f_2(y)}\right] \\ &= \frac{f_2(cy)}{yf_2(y)}\left[K(cy) - K(y)\right].\end{aligned}$$

Thus, we need only show $K(y)$ is strictly decreasing to obtain the monotonicity results for $0 < c < 1$ and $1 < c$, simultaneously.

Using identities for U,

$$K(y) = \frac{5}{6}\left[1 - \frac{yU(\frac{11}{6}, \frac{5}{3}; y)}{U(\frac{5}{6}, \frac{2}{3}; y)}\right] = \frac{5}{6}[1 - \lambda(y)].$$

So, it suffices to show $\lambda(y)$ is strictly increasing for $y > 0$.

Applying more identities for U and U', we obtain

$$\begin{aligned}\lambda'(y) &= \frac{7/6}{[U(\frac{7}{6}, \frac{4}{3}; y)]^2} \\ &\quad \times \left[U(\frac{7}{6}, \frac{1}{3}; y)U(\frac{13}{6}, \frac{7}{3}; y) - U(\frac{7}{6}, \frac{4}{3}; y)U(\frac{13}{6}, \frac{4}{3}; y)\right] \\ &= \frac{7/6}{[U(\frac{7}{6}, \frac{4}{3}; y)]^2}\mu(y).\end{aligned}$$

Then,

$$\begin{aligned}\mu(y) = \hat{c}_1 \Bigg[&\int_0^\infty e^{-yt}t^{1/6}(1+t)^{-11/6}dt \int_0^\infty e^{-ys}t^{7/6}(1+t)^{-5/6}ds \\ &- \int_0^\infty e^{-yt}t^{1/6}(1+t)^{-5/6}dt \int_0^\infty e^{-ys}t^{7/6}(1+t)^{-11/6}ds\Bigg].\end{aligned}$$

Applying Fubini's Theorem, a change of variable and switching to polar coordinates, we obtain

$$\mu(y) = c_2 \int_0^\infty e^{-yr^2} r^{\pi/3} dr \int_0^{\pi/2} \omega(y, r, \theta) d\theta dr,$$

where

$$\omega(y,r,\theta) = [(1+r^2\sin^2\theta)(1+r^2\cos^2\theta)]^{-11/6}$$
$$\times \sin^{1/3}\theta\cos^{7/6}\theta\cos 2\theta.$$

Then, halfing θ and making a change of variables we have

$$\int_0^{\pi/2}\omega(y,r,\theta)d\theta = \int_0^{\pi/2}[(1+r^2\frac{1+\cos\phi}{2})(1+r^2\frac{1-\cos\phi}{2}]^{-11/6}\sin^{1/3}\phi\cos^2\phi d\phi > 0$$

as claimed. Q.E.D.

For $0 < \alpha < 1$, define

$$f_\alpha(x) = \begin{cases} g_\alpha(-x), x < 0 \\ \frac{1}{\pi}\Gamma(1+\frac{1}{\alpha}), x = 0 \\ g_\alpha(x), x > 0 \end{cases},$$

where

$$g_\alpha(x) = \frac{\alpha x^{\frac{1}{\alpha-1}}}{2(1-\alpha)}\int_0^1 V_\alpha(\phi)e^{-x^{\frac{\alpha}{\alpha-1}}V_\alpha(\phi)}d\phi,$$

with

$$V_\alpha(\phi) = \left(\frac{\sin(\frac{\pi}{2}\alpha\phi)}{\cos\frac{\pi}{2}\phi}\right)^{\frac{\alpha}{\alpha-1}}\frac{\cos(\frac{\pi}{2}(1-\alpha)\phi)}{\cos\frac{\pi}{2}\phi}.$$

Arguing as before, the problem reduces to showing that for

$$W_k = \int_0^1 e^{-V_\alpha(t)}V_\alpha(t)dt$$

that

$$\frac{W_3}{W_2} - \frac{W_2}{W_1} \le 1 \qquad (4.1)$$

We conjecture that (4.1) holds.

5. Asymptotic Statistics

We worked with A. Trindade in our submitted paper, "An Improved Asymptotic Approximation for 1st Moment of Order Statistics from a Generalized Exponential Distribution". (See [4]) Again we give some background notation:

Given a random sample

$$\underline{X} = (x_1, x_2, \cdots, x_n)$$

consider the order statistics

$$x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}.$$

Assume the underlying distribution with cdf, $F(x)$, has finite moments of all required orders, i.e., $\int x^n dF < \infty$, for all $n > 0$. By the probability integral transform we have for each r, $1 \leq r \leq n$,

$$x_{r:n} = F^{-1}(U_{r:n})$$

where $U_{1:n} \leq U_{2:n} \leq \cdots \leq U_{n:n}$ are the order statistics of a random sample with uniform distribution on $(0, 1)$.

Let $\mu_{r:n} := \mathbb{E}[X_{r:n}]$ be the mean of the r^{th} order statistic where

$$\mathbb{E}[U_{r:n}] = \frac{r}{n+1} \equiv p_{r:n}.$$

Using the Taylor approximation

$$\mu_{r:n} = \mathbb{E}[F^{-1}(U_{r:n})] \overset{(7)}{\approx} F^{-1}(\mathbb{E}[U_{r:n}]) = F^{-1}(p_{r:n}) := \lambda_{r:n}$$

(7) is called the David-Johnson Approximation (1954).

Our goal: To assess the large-scale accuracy of the approximation (7) when the sampling is from a Generalized Exponential Distribution (GED).

Let $X \in GE(\theta)$ denote the standard member of the GED with shape parameter $\theta > 0$ defined by a distribution with cdf

$$F(x) = (1 - e^{-x})^{\theta}, x > 0.$$

For $\theta = 1$, the GED is the standard exponential distribution. There are several generalizations depending on the shape parameter "θ". We will show that $\mu_{r:n}$ is asmyptotically equivalent to $\lambda_{r:n}$, i.e.,

$$\lim_{n\to\infty} \frac{\mu_{r:n}}{\lambda_{r:n}} = 1 \text{ iff } \theta = 1.$$

We use the formula given by Ragab and Ahsonulleah (2001), i.e., with $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$ we have

$$\mu_{r:n} = \frac{\Gamma(n+1)}{\Gamma(r)\Gamma(n-r+1)} \sum_{k=0}^{n-r} \frac{(-1)^k}{r+k} \binom{n-r}{k} [\Psi(1 + \theta(r+k)) - \Psi(1)].$$

The disadvantage here is that it requires very high precision arithmetic for modest accuracy! It follows easily that the D-J approximation (7) yields

$$\lambda_{r:n} = F^{-1}(p_{r:n}) = -\log[1 - \left(\frac{r}{n+1}\right)^{\frac{1}{\theta}}].$$

We prove

Theorem 5.1. *For $\mu_{r:n}$ and $\lambda_{r:n}$ defined above*

$$\lim_{n\to\infty}\frac{\mu_{r:n}}{\lambda_{r:n}}=\frac{\Gamma(r+1/\theta)}{\Gamma(r)}\left(\frac{1}{r}\right)^{1/\theta}.$$

Proof. Defining $T(r,n)=\Gamma(n+1)/[\Gamma(r)\Gamma(n-r+1)$ and using the identity $\binom{n-r}{k}=(-1)^k\frac{[-(n-r)]_k}{k!}$ and a series expansion for $\Psi(a)-\Psi(b)$ we obtain

$$\mu_{r:n}=T(r,n)\sum_{k=0}^{n-r}\frac{[-(n-r)]_k}{k!(n+k)}\sum_{l=0}^{\infty}\frac{\theta(r+k)}{(l+1+\theta(r+k))(l+1)}.$$

Now if we switch order of the sums and simplify

$$\mu_{r:n}=T(r,n)\sum_{l=0}^{\infty}\frac{\theta(r+k)}{(l+1+\theta(r+k))(l+1)}$$
$$\times\ {}_2F_1(-(n+r),\frac{l+1}{\theta}+r,\frac{l+1}{\theta}+r+1;1).$$

(Several manipulations needed justifying: using integral representation, switching order, change of variables twice). Let

$$\mu_{r:n}=T(r,n)\int_0^{\infty}e^{-(n-r+1)s}q(s)ds,$$

where $q(s)=(1-e^{-s})^{r-1}\{-\log[1-(1-e^{-s})^{\frac{1}{\theta}}]\}$. Making a change of variables $m=n-r+1$ and $t=ms$

$$\mu_{r:n}=\frac{T(r,n)}{m^{\frac{1}{\theta}+r}}\int_0^{\infty}e^{-t}q_m(t,\frac{1}{\theta},r-1)dt,$$

where

$$q_m(t,a,b)=-m^{a+b}(1-e^{-\frac{t}{m}})^b\log[1-(1-e^{-\frac{t}{m}})^a].$$

Using asymptotics of q_m, applying Watson's Lemma and the Dominated Convergence Theorem gives

$$\lim_{n\to\infty}\mu_{r:n}=\lim_{n\to\infty}\left\{\frac{T(r,n)}{(n-r+1)^{\frac{1}{\theta}+r}}\right\}\int_0^{\infty}e^{-t}t^{\frac{1}{\theta}+r}dt.$$

Thus,

$$\mu_{r:n} \sim T(r,n)\frac{\Gamma(\frac{1}{\theta}+r)}{(n-r+1)^{\frac{1}{\theta}+r}} \sim \frac{\Gamma(\frac{1}{\theta}+r)}{\Gamma(r)}\left(\frac{1}{n-r+1}\right)^{-\frac{1}{\theta}}$$
$$\sim \frac{\Gamma(\frac{1}{\theta}+r)}{\Gamma(r)}\left\{-\log[1-(\frac{1}{n-r+1})^{\frac{1}{\theta}}]\right\}.$$

which implies

$$\lim_{n\to\infty}\frac{\mu_{r:n}}{\lambda_{r:n}} = \frac{\Gamma(r+1/\theta}{\Gamma(r)}\left(\frac{1}{r}\right)^{1/\theta}.$$

■

Remark 5.2. Theorem 5.1 provides the useful result that

$$\mu_{r:n} \sim \lambda_{r:n} \ \text{ iff } \ \theta = 1.$$

This technique generalizes to several generalized distributions.

References

[1] R.W. Barnard and P. Hadjicostas, Bank's criterion and symmetric stable laws with index of stability less than one. (In preparation.)

[2] R.W. Barnard and C. Monico, How dense are the happy numbers base 3? (In preparation.)

[3] R.W. Barnard and K.C. Richards, A note on Turan-type and mean inequalities for the Kummer function, *J. Math. Anal. Appl.*, **349**(2009), 259–263.

[4] R.W. Barnard and A.A. Trindade, An improved asymptotic approximation for 1st order statistics from a generalized exponential distribution. (Submitted.)

[5] L. Christensen, J. Striuli, and O. Veliche, Growth in the minimal injective resolution of a ring, *J. London Math. Soc.*, **81**(2010), 24–44.

[6] P. Hadjicostas and S.I. Karanarantne, Comparison of location estimators using Bank's criterion, *Math. Inequal. Appl.*, **12**(2009), 455–472.

[7] A. Tefara, What is a Wilf-Zeilberger pair? *Notices Amer. Math. Soc.*, **57**(2010), no. 4, 508–509.

Roger W. Barnard
E-mail: `roger.w.barnard@ttu.edu`
Address:
Department of Mathematics and Statistics,
Texas Tech University,
Lubbock, TX 79409
U.S.A.