Median Stable Distributions and Schröder's Equation

Gib Bassett

July 9, 2017

Abstract: Median stable distributions were recently introduced as an extension of the idea of (mean) stable distributions. The extension defines a stable distribution not in terms of sums of independent random variables, but instead relative to an estimator. For mean stable distributions the mean's (rescaled) sampling distribution is the same as the iid data on which the estimator is based. Median stable distributions replace the mean with the median in this definition. Since the sampling distribution of the median is a functional (instead of a weighted sum) of the data its stable distribution is the solution to a functional equation. It turns out that this defining functional equation is an instance of a famous equation due to Schröder from 1870. The renown of the equation is due to the way it incorporates iteration of functions, a key feature of what many years later would become dynamic systems analysis. The current paper reviews median stable distributions in light of the connection to Schröder's functional equation.

1 Introduction: Median Stable Distributions

Median stable distributions were recently introduced as an extension of the idea of (mean) stable distributions [1]. Such distributions are defined by recasting the standard (mean) definition of stability to refer to an estimator's sampling distribution: stability is defined so that an estimator's (rescaled) sampling distribution is the same as the distribution of the iid data on which the estimator is based.[1]

Let $\hat{X} = \hat{X}(F)$ denote the sample mean based on iid data, $X_i(F) = X(F)$ whose common cdf is F. A mean stable distribution is a cdf H so that (after recentering by μ and rescaling by λ) the distribution of the estimator and the data are the same:

$$\lambda(\hat{X}(H) - \mu) \stackrel{d}{=} X(H) - \mu \tag{1}$$

Familiar examples of mean stable distributions are the normal (with μ the expected value, λ=$\sqrt{n}$), and the Cauchy (with μ the median, and $\lambda = 1$). Replacing the sample mean with the sample median in (1) defines a median stable distribution.

[1] Identifying a stable distribution in terms of an is not standard. The usual presentation, motivated by the sample mean, is in terms of *sums* of i.i.d random variables; for example, [5]. Reinterpreting the standard definition in terms of the sample mean estimator has the advantage of generalization to new contexts.

To fix ideas consider the simplest case of the median with $n = 3$ i.i.d observations. Let $\hat{X} = \hat{X}(F)$ denote the sample median and its sampling cdf by $M(x) = M(x : F) = \Pr[\hat{X} < x]$, which is given by,

$$M(x) = G(F(x)) \tag{2}$$

where the "G-function" is[2],

$$G(w) = 3w^2 - 2w^3, \quad w \in [0,1]. \tag{3}$$

Note that $G(w)$ is the distribution of the sample median when the data is uniformly distributed on $[0,1]$.

The G-function is depicted in Figure 1 along with schematics illustrating how G turns F into M. The G-function is seen to be continuous, increasing, symmetric about $\frac{1}{2}$ with fixed points at $0, \frac{1}{2}, 1$. G is convex (concave) for $0 < w < \frac{1}{2}(\frac{1}{2} < w < 1)$. $G(w) > w$ for $\frac{1}{2} < w < 1$, $G(w) < w$ for $0 < w < \frac{1}{2}$. These features also hold for the median G-functions with $n > 3$. While the analysis of classical mean stability is about sums of random variables, median stability concerns the behavior of functionals like G.

It is readily verified that the median of M and F are always the same. The sample median is median unbiased similar to the sample mean being mean unbiased. Without loss of generality the data is hereafter centered so that either its unique median is 0, or 0 is in the interval of medians, $[\mu^-, \mu^+]$.

A median stable distribution therefore will be a cdf H that depends on a λ-scaling factor satisfying (1), or, $H(x) = M[x : H(\lambda^{-1}x)]$, or on substituting for M,

$$H(x) = G(H(\lambda^{-1}x)). \tag{4}$$

The solutions to this functional equation define the $n = 3$ version of median stable distributions.

Median stable distributions for $n = 2r+1$ are defined similarly: substitute G_r for G in (4) where G_r is the distribution of the sample median given iid uniformly distributed data. This $G_r(w)$, $w \in [0,1]$ can be viewed as a mapping from the G in (3) (whose derivative is g) to G_r given by;

$$G_r(w) = C_r \int_0^w g(t)^r dt. \tag{5}$$

where the constant C_r makes $G(1) = 1$; $C_r = 6^{-r}\frac{(2r+1)!}{r!r!}$; see, e.g., [4]. The sampling cdf of the median based on independent data with cdf F is, $G_r(F(x))$.

Schröder. After median stable distributions were introduced I discovered that the functional equation defining median stablity is an instance of a famous functional equation considered by Schröder in 1870, [11]. As in our case, the givens are G and λ, and the Schröder problem is to solve the functional equation (4) for the unknown H.

The fame of the equation is due its important role in what would later become complex dynamic systems. It arises because of the neat way that the equation involves iteration of functions.

[2] $\hat{X} < x$ if: (i) two-out-of-three, or (ii) three-out-of-three of the X_i are less than x. Two-out-of-three has probability, $F(x)^2(1-F(x))$, and can occur in three-choose-two equals 3 ways. Three-out-of-three can occur in 1 way, which has probability, $F(x)^3$. So, $M(x) = F(x)^3 + 3F(x)^2(1-F(x)) = 3F(x)^2 - 2F(x)^3 = G(F(x))$.

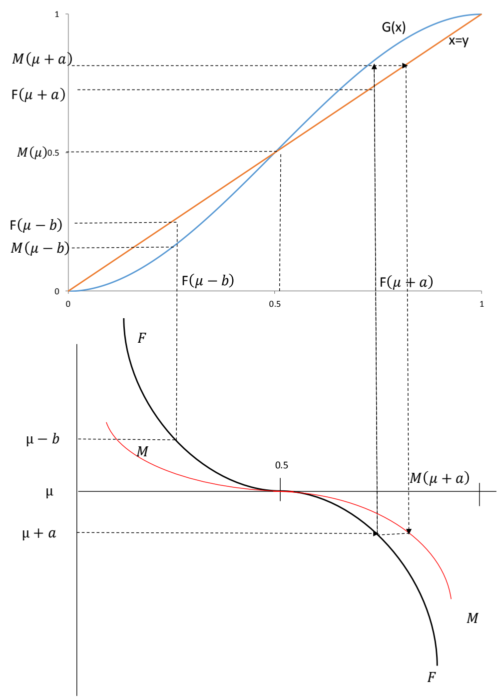

Figure 1: Mapping from F to M, via G

To illustrate, suppose H has an inverse (in our case, quantile) function H^{-1} so that G in (4) can be expressed in terms of H and H^{-1}:

$$G(w) = H(\lambda H^{-1}(w)). \tag{6}$$

Notice that in composing $G(w)$ 2-times, terms cancel so that,

$$G^{(2)}(w) = G(G(w)) = H(\lambda H^{-1}(H(\lambda H^{-1}(w)) = H(\lambda^2 H^{-1}(w))$$

and continuing,

$$G^{(k)}(w) = H(\lambda^k H^{-1}(w)). \tag{7}$$

This k-composition works when k is a positive integer. But it also works for any real k thus defining composition for non integers. For example, $G^{(\frac{1}{2})}(w) = H(\lambda^{\frac{1}{2}} H^{-1}(w))$ (because $G^{(\frac{1}{2})}(G^{(\frac{1}{2})}(w)) = G(w))$. Alternatively: $G^{(\sqrt{2})}(w) = H(\lambda^{\sqrt{2}} H^{-1}(w))$, and $G^{(\sqrt{-1})}(w) = H(\lambda^{\sqrt{-1}} H^{-1}(w))$.

The context that led Schröder to his functional equation was Newton's iteration method for approximating the roots of a polynomial. A polynomial determines an iterating Newton G-function whose fixed point is a root of the polynomial. Under the right conditions on the polynomial and a starting w_0 the iteration $G^{(k)}(w_0)$ converges (quickly) to a root of the polynomial.

Schröder was the first to use H-functions as a tool for understanding iteration of functions; see [3]. For example, consider the sequence generated by a $G(w)$ process that starts at a $w_0 \in (0,1)$, where $w_{k+1} = G(w_k) = G^{(k)}(w_0)$. The associated H describes the evolution of the sequence, as well as the interpolated (continuous as a function of k), $G^{(k)}(w_0) = w_k$.

Solving (4) for the unknown H in terms of G is not straightforward. In fact, despite introducing the functional equation "Schröder never succeeded in finding methods that guaranteed solutions.... He more or less admitted defeat, settling instead for what one might call the "Opposite Approach", wherein one begins with [H] and λ, and then defining [G] [via (6)], thereby obtaining ready made examples of "pre-solved" Schröder equations"([3], p.214).

It was not until 1884 that Koenigs [7] derived H-solutions for a given G. For most G functions there is no closed form solution and H is expressed as the limit of a functional composition operation involving the known G function.

Discussion of Schröder's Equation and median stable distributions is continued in the next section. Section 3 applies the results to the G functions that correspond to the median. Properties of H in terms of G are not obvious, and results are derived for H around zero and in the tails.

Section 4 discusses connections between median stable distributions and the remedian, an estimator defined recursively as the median of medians; see [10]. Many well known features of mean stable distributions have direct extensions to the remedian. This is a consequence of: (i) the recursive feature of the remedian is shared by the sample mean (the mean of means is the mean) and (ii) many features of mean stable distributions derive from the recursive property. An example is the domain of attraction: the limit distribution of the remedian (for any data) is a remedian stable distribution.

Section 5 notes that median G_r-functions can be composed with different r-values, thus suggesting varieties of remedians with mixed bases, as well as their stable distributions.

The final section departs from the median problem with its G_r and takes up Schröder's "Opposite Approach" in which G-functions are derived from known H distributions. In particular, the G functions corresponding the the Pareto and Laplace distribution are derived.

Discussion: Mean stable distributions do not depend on the number of observations. The normal, Cauchy, or any other mean stable distribution is stable for any n. The number of observations only determines the appropriate λ_n scaling factor; for the normal, $\lambda_n = \sqrt{n}$; for the Cauchy, $\lambda_n = 1$.

Another difference between mean and median stable distributions concerns the scaling factor λ. For mean stable distributions there are λ-scaling factors less than one; with fat-tailed data the sample mean is worse than (more dispersed) than the data so that stability requires $\lambda < 1$. The sample median in contrast is always (any F) more concentrated than the data, so that its scaling factor is always greater than one.

2 Schröder's Equation

A median G-function is continuous, increasing, symmetric about $\frac{1}{2}$ with fixed points at $(0, \frac{1}{2}, 1)$. It is convex (concave) for $0 < w < \frac{1}{2}(\frac{1}{2} < w < 1)$, and $G(w) > w$ for $\frac{1}{2} < w < 1$, $G(w) < w$ for $0 < w < \frac{1}{2}$. Given such a G and a $\lambda > 1$, an H solution to (4) inherits these features and will be: increasing, symmetric about 0 with $H(0) = \frac{1}{2}$, convex (concave) for $0 < w < 1/2$ $(1/2 < w < 1)$.

An (H, λ) that solves (4) determines a scale family of distributions. If $(H(x), \lambda)$ is a solution to the function equation then so is, $H(\sigma x) = G(H(\lambda^{-1}\sigma x)$, $\sigma > 0$.

In addition, an $(H(x), \lambda_0)$ for a given λ_0 determines solutions for all $\lambda > 1$. That is, if $(H(x), \lambda_0)$ solves (4) then so does $(H(x^\alpha), \lambda_0^{\alpha^{-1}})$, all $\alpha > 0$.

(Let $H(x|\lambda_0)$ denote the solution with λ_0. Consider $L(x) = H(x^\alpha|\lambda_0)$ so that, $L(x) = G(H(\lambda_0^{-1}x^\alpha)) = G(H((\lambda_0^{-1/\alpha}x)^\alpha)) = G(L(\lambda^{-1/\alpha}x))$, which says, $L(x)$ with $\lambda_0^{-1/\alpha}$ solves the functional equation).

Since a solution for any $\lambda > 1$ determines solutions for all λ we focus on the case, $\lambda = s$ where s denotes the slope of $G(w)$ at the fixed point, $w = \frac{1}{2}$.

The derivative of (4) is: $h(x) = g(H(\lambda^{-1}x)h(\lambda^{-1}x)\lambda^{-1}$, so $h(0) = g(\frac{1}{2})h(0)\lambda^{-1}$. If $0 < h(0) < \infty$, then $\lambda = g(\frac{1}{2}) = s$.

This $\lambda = s$ particular case of the functional equation is:

$$H(x) = G(H(s^{-1}x)). \tag{8}$$

Schröder Solutions. Repeated substitution of the LHS of (8) at $s^{-1}x$ into the RHS, gives, $H(x) = G^{(k)}(H(s^{-k}x)) = G^{(k)}(\frac{1}{2} + h(0)s^{-k}x) + o(s^{-k}x))$. Koenigs [7] showed that the solution to (8) (with $h(0) = 1$) is,

$$H(x) = \lim_{k\to\infty} G^{(k)}(\tfrac{1}{2} + s^{-k}x) \tag{9}$$

For an $H(x)$ with $(h(0) = 1)$ the associated $G(w)$ with $g(\frac{1}{2}) = s$ is given by,[3]

$$G(w) = H(sH^{-1}(w)) \tag{10}$$

3 Properties of Median Stable Distributions

Computations reveal that H is close to normal. Some evidence about the relation to the normal is in the behavior of the density $h(x)$ around zero which is also proportional to e^{-x^2}.

[3]Analogous to the *standard* normal distribution, $h(0) = 1$ defines the *standard* cdf solution to the functional equation.

3.1 $H_r(x)$: x near zero

The behavior of $H(x)$ around $x = 0$ depends not surprisingly on $G(w)$ around $w = \frac{1}{2}$. The next result shows that, while H is not normal, its log density at zero, like the normal, is proportional to x^2.

Theorem 1. *$h_r(x) = \exp(-\alpha(r)x^2 h_r(0)^2 + o(x))$*
where, $\alpha(r) = \frac{-4r}{s^2-1}$.

Proof. The derivative of (8) (suppressing the r-subscript) is $h(sx) = s^{-1}g(H(x))h(x)$, where $g(w) = s[1-4(w-1/2)^2]^r$, so: $\log(h(sx)) - \log(h(x)) = r\log(1-4(H(x)-1/2)^2)$. Dividing by x^2 and taking limits, $x \to 0$ gives, for the RHS:$-4rh(0)^2$, while the LHS is:

$$\lim_{x\to 0} \tfrac{s^2\log(h(x))}{x^2} - \tfrac{\log(h(sx))}{s^2x^2} = (s^2-1)\lim_{x\to 0} \tfrac{\log(h(x))}{x^2}.$$

So,

$$\lim_{x\to 0} \tfrac{\log(h(x))}{x^2} = -\alpha(r)h(0)^2,$$

where $\alpha(r) = \frac{-4r}{s^2-1}$. $\square$

The normal (with density equal to one at the origin) has density, $e^{-\pi x^2}$. The density of the median stable distribution (with density equal to one at the origin) is $e^{-\alpha(r)x^2}$ for x near the origin. With $n = 3$ ($r = 1$), the median stable coefficient on x^2 is $\alpha(1) = 3.2$ close to the normal value of π. As $r \to \infty$, $\alpha(r)$ decreases to π.

3.2 $H_r(x)$: $x \to -\infty$

The tail of H depends on $G(w)$ near $w = 0$ (or what is the same, via symmetry, near $w = 1$). To motivate, consider the simplest case of $r = 1$, $n = 2r+1 = 3$ where $G(w) = H(sH^{-1}) = 3w^2 - 2w^3$ or

$$\frac{H(sH^{-1})}{3w^2} = 1 - \frac{2}{3}w$$

or changing to $x = H^{-1}$ and taking limits,

$$\lim_{x\to -\infty} H(sx)(3H(x)^2)^{-1} = 1$$

and this means

$$\lim_{x\to -\infty} -|x|^{-\alpha}\log 3H(x) = 1$$

where $s^\alpha = 2$. With $s = \frac{3}{2}$, $\alpha \approx 1.71$, which compares to the Normal tail rate of $\alpha = 2$.

For $r > 1$ the leading term of $G_r(w)$ $\frac{(2r+1)!}{r!(r-1)!}w^{r+1}$, $s_r = 4^r\frac{(2r+1)!}{r!r!}$ so

Theorem 2.

$$\lim_{x\to -\infty} -|x|^{-\alpha}\log cH_r(-x) = 1$$

where, $c = \frac{(2r+1)!}{r!(r-1)!}$, and $s_r^\alpha = r+1$.

Proof. The proof is similar to the above $r = 1$ case and is omitted. $\square$

A final property for H follows from the well known limiting distribution of the median: as $n \to \infty$, the ($\sqrt{n} = \sqrt{2r+1}$-scaled) median given iid uniform data is normal with mean zero and standard deviation $\frac{1}{2}$. In our notation,

$$= \lim_{r\to\infty} G_r(\tfrac{1}{2} + \frac{1}{\sqrt{2r+1}} x) = N(0:0, \tfrac{1}{2}).$$

In terms of our functional equation this means,

$$\lim_{r\to\infty} G_r(N(s_r^{-1} x : 0, \tfrac{1}{2})) = N(x : 0, \tfrac{1}{2})$$

(Verify noting

$$\lim_{n\to\infty} s_n^{-1}\sqrt{n} = \sqrt{\pi/2}$$

and $N((\pi/2)^{1/2} x : 0, \sigma_M) = N(x : 0, \frac{1}{2})$). Hence the normal distribution solves the functional equation as $n = 2r+1 \to \infty$.

4 Remedian Stable Distributions

The remedian is defined as the median of medians; see [10]. Its base-r version-(with $(2r+1)^k$ observations) is defined recursively as the median of "data", each of which is itself a median, $(\hat{X}_{1\hat{X}_{(k-1)}}, \hat{X}_{2\hat{X}_{(k-1)}}, ..., \hat{X}_{(2r+1)\hat{X}_{(k-1)}})$. Because of the recursive definition, the sampling distribution is the k^{th} iterate of the median G_r function; the sampling distribution of the base $2r+1$ remedian with $n = (2r+1)^k$ observations is, $G_r^{(k)}(F(x))$, [10].

Consider the scaled-by s_r^k-distribution of the remedian; $G_r^{(k)}(F(s_r^{-k}x))$. If the data has median stable distribution, H_r, then the cdf of the remedian is $G_r^{(k)}(H_r(s_r^{-k}x)) = H_r(k)$. This says the (rescaled) remedian with H_r data will have the same sampling distribution. In other words, the base r remedian has remedian stable distribution equal to the median stable distribution H_r. This remedian stable distribution H_r is stable for all $n = (2r+1)^k$, $k = 1, 2, ...$

Domain of Attraction. The distribution of the sample mean with normal data is also normal. This is the defining property of stability. Since mean stability holds for all n, the limiting distribution of the sample mean given normal data will also be normal. But even when the data is not normal, the limiting distribution will be normal (under certain conditions). (And if the conditions for a normal limit do not hold, the limit will be some other mean stable distribution). The mean stable distributions are all the possible limiting distributions of the sample mean.

This does not work for the sample median because–among other things–median stable distributions vary with n.

But it does work for the sample remedian. The remedian, base r, stable distributions (which is a r-median stable distribution) are all possible limits of the sample remedian, base r, given any data.

Let the data have distribution, $F(x)$, where suppose initially, $f(x) > 0$ in a neighborhood of zero. The distribution of the remedian is, $G_r^{(k)}(F(s_r^{-k}x))$, or $G_r^{(k)}(\frac{1}{2} + f(\omega s_r^{-k}x)x)$, for $|\omega| < 1$, which as $k \to \infty$ is $H_r(f(0)w)$, the median stable distribution with scale, $1/f(0)$.

(More generally, let $F(x) = \frac{1}{2} + c(x)x^\alpha + o(x)$, where $c(x) > 0$ in a neighborhood of 0. Then, as in the above, the limiting distribution of the remedian is, $H_r(c(0)x^\alpha))$.

5 Mixing G functions

The remedian was defined in [10] as the median of medians where the medians all had the same "base"–number of observations. Consider a modification that mixes different bases, which corresponds to sampling distributions that compose different G functions.

For example, consider, $G_1(G_2(w))$. The corresponding estimator can be described as follows. Let $n = 15$ and put the data into an $5x3$ array. The estimator defined as the median of the 5 row medians has distribution, $G_1(G_2(F(x)))$.

Consider, instead, the median of the column medians. The distribution of this estimator is $G_2(G_1(F(w)))$ (which is not the same as the first estimator). While these two composite G functions are different, their derivatives at the $w = \frac{1}{2}$ fixed point are equal, $(s_1 s_2 = s_2 s_1)$. Which estimator is better? [4] Should you do the median of 3 then 5 or 5 then 3?

Each estimator has its associated stable distribution H_{12}, H_{21}. The functional equation for the first is, $H_{12}(x) = G_{12}(H_{12}((s_1 s_2)^{-1}x))$, whereas for the second, $H_{21}(x) = G_{21}(H_{21}((s_1 s_2)^{-1}x))$.

6 Pareto and Laplace G functions

This section presents two examples of presolved functional equations. It is reminiscent of Schröder's opposite approach approach in which, being unable to solve for H in terms of G, he starts with an H and derives its G. The first example is the Pareto distribution for which the resulting G is piecewise linear. The second example has H as the Laplace distribution and derives its associated G.

The G functions for the examples are compared with each other as well as the median G function ($r = 1, n = 3$) in Figure 6.

Pareto. Consider the polyhedral G function,

$$\begin{aligned} G(\tfrac{1}{2} + w) &= \tfrac{1}{2} + sw & 0 \le w < w_0 < \tfrac{1}{2} \\ &= \tfrac{1}{2} + sw_o + b(w - w_o) & w_o \le w \le \tfrac{1}{2} \end{aligned}$$

where $s > 1$, $b = \frac{1-2sw_o}{1-2w_o}$, and $G(\frac{1}{2} - w)$ is given by symmetry.

The associated $H(x)$ is a polyhedral–linear in-between the kinks–cdf whose value at the kink, $x_i = s^i w_o$, is:

$$H(s^i w_o) = G^{(i)}(\tfrac{1}{2} + w_o) = \tfrac{1}{2} + sw_o + (s-1)w_o \tfrac{b-b^i}{1-b}$$

[4] Answer: $G_2 G_1$ is better than $G_1 G_2$

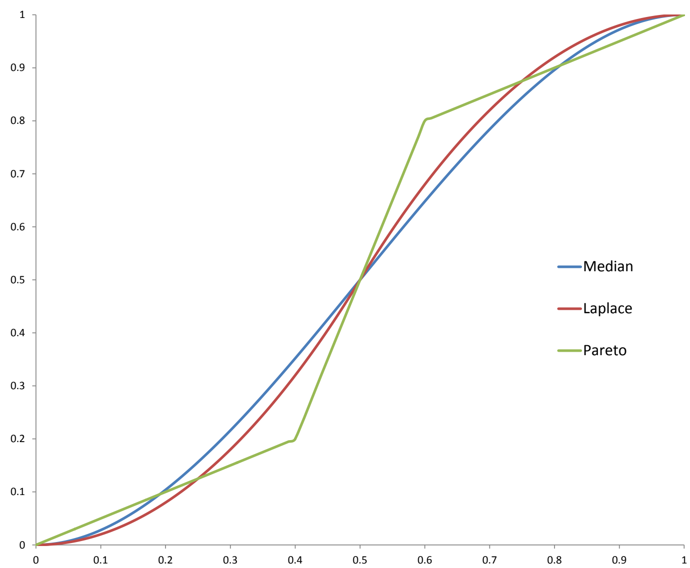

Figure 2: G functions

(Verify using (9), and $H_k(x) = G^{(k)}(\frac{1}{2} + s^{-k}x)$, where $H_k(s^i w_o) = G^{(i)}(\frac{1}{2} + w_o) = \frac{1}{2} + sw_o + (s-1)w_o\frac{b-b^i}{1-b})$. Evaluating the density/slope of $H(x)$ gives

$$h(x) = (\frac{b}{s})^i$$

for, $s^i w_o < x < s^{i+1} w_o$, a linear interpolation of the Pareto distribution; that is, $h(x) = Ax^{-\beta}$, $\beta = 1 - \frac{\log b}{\log s}$.

Laplace. Let $H(x)$ be the standard ($h(0) = 1$) Laplace distribution and $H^{-1}(w)$ the associated inverse/quantile function:

$$\begin{aligned} H(x) &= 1 - \tfrac{1}{2}e^{-2x} & x > 0 \\ H^{-1}(w) &= -\tfrac{1}{2}ln[2(1-w)] & w > \tfrac{1}{2} \end{aligned} \tag{11}$$

The distribution for $x < 0$ is given by symmetry, $(H(x) = 1 - H(-x))$, as is the inverse for $w < \frac{1}{2}$; $H^{-1}(w) + H^{-1}(1-w) = 1$. For the associated G-function:

$$\begin{aligned} G(w) &= 1 - \tfrac{1}{2}[2(1-w)^{\lambda}] & \tfrac{1}{2} < w < 1 \\ &= 2^{\lambda-1}w^{\lambda} & 0 < w < \tfrac{1}{2}. \end{aligned} \tag{12}$$

This pair, (G,H) can be used to create a new G associated with the same H, which has any prescribed $g(\frac{1}{2}) = \lambda$. This provides bounds for stable distributions that are otherwise difficult to discern.

To illustrate, let $G_1(w)$ denote the median G-function (2) with its $s_1 = g_1(\frac{1}{2}) = \frac{3}{2}$. We want to bound the unknown $H_1(x)$ relative to the Laplace.

Let $H_L(x)$ denote the Laplace distribution (11), and its G function (12), which is denoted by, $G_L(w|\lambda)$. Select $\lambda = s_1$, so that $G_1(w)$ and $G_L(w|s_1)$ have the same derivative at the fixed point: $g_1(\frac{1}{2}) = g_L(\frac{1}{2}|s_1) = s_1$.

The two G functions can be compared to infer a bound on the unknown $H_1(x)$ relative to the known $H_L(x)$. In particular, for $\frac{1}{2} < w < 1$, verify the inequality:

$$G_1(w) = 3w^2 - 2w^3 > 1 - \tfrac{1}{2}(1-w)^{\frac{3}{2}} = G_L(w|s_1)$$

which implies, $G_1^{(k)}(\frac{1}{2} + s_1^{-k}x) > G_L^{(k)}(\frac{1}{2} + s_1^{-k}x|s_1)$. Hence,

$$H_1(x) > H_L(x|s_1), \qquad x > 0 \tag{13}$$

While the stable distribution H_1 is not known precisely, (13) says it is less dispersed than the Laplace distribution.

References

[1] Bassett, G. W. (2016). "Median Stable Distributions", in Robust Rank-based and Nonparametric Methods, by Regina Liu and Joseph W. McKean. Springer, (2016), Chapter 14: 233-244.

[2] Carleson, Lennart; Gamelin, Theodore W. (1993). Complex Dynamics. Textbook series: Universitext: Tracts in Mathematics. New York: Springer-Verlag. ISBN 0-387-97942-5.

[3] Curtright, Thomas, and Cosmas Zachos (2009). "Evolution profiles and functional equations." J. Phys. A: Math. Theor. 42.ANL-HEP-PR-09-80 (2009).

[4] David H. A. (1981). Order Statistics. Wiley, New York (1981).

[5] Feller, W. (1971). An Introduction to Probability Theory and Its Applications. Wiley, New York (1971).

[6] Gamelin, T. W. (1996). "A History of Complex Dynamics, from Schrder to Fatou and Julia. By Daniel S. Alexander. Wiesbaden (Vieweg)." Historia Mathematica 23.1 (1996): 74-84.

[7] Koenigs, G. (1884). "Recherches sur les intgrales de certaines quations fonctionelles". Annales Scientifiques de l'cole Normale Suprieure 1 (3, Supplment): 341.

[8] Kuczma, Marek (1968). Functional equations in a single variable. Monografie Matematyczne. Warszawa: PWN Polish Scientific Publishers. ASIN: B0006BTAC2

[9] Kuczma, Marek (1964). "Note on Schrder's functional equation." Journal of the Australian Mathematical Society 4.02 (1964): 149-151.

[10] Rousseeuw, P.J., Bassett, G.W. (1990). The Remedian: A Robust Averaging Method for Large Data Sets. J. Amer. Statist. Assoc. **85**, 97-104 (1990).

[11] Schrder, Ernst (1870). ”Ueber iterirte Functionen”. Math. Ann. 3 (2): 296322. doi:10.1007/BF01443992.

[12] Shapiro, Joel H. (1998). ”Composition operators and Schrder's functional equation.” Contemporary Mathematics 213 (1998): 213-228.

[13] Szekeres, G. (1958). ”Regular iteration of real and complex functions”. Acta Mathematica 100 (34): 361376. doi:10.1007/BF02559539. [1]