Median Stable Distributions

Gib Bassett

Abstract: When i.i.d. data follows a stable distribution the sample mean has the same distribution as the rescaled data. The normal with a scaling factor of $\sqrt{n}$, and the Cauchy with scaling factor of 1 are well known examples of mean stable distributions. This idea is extended to *median* stable distributions by requiring that the sampling distribution of the median be identical to the distribution of the rescaled data. The median's sampling distribution is a functional of the data's cdf so that the analysis of median stability involves solutions to functional equations (as opposed to sums of random variables). A few properties of median stable distributions are presented including their relation to the limiting distribution of the remedian.

1 Introduction: The Sample Median with Small (n=3) Data

Consider the distribution of the sample median in the simplest case of $n = 3$ i.i.d observations. That is, X_1, X_2, X_3, are i.i.d with the common distribution function F; F will be sometimes referred to as the distribution of the data. The random variable with the cumulative distribution function (cdf) of the data is denoted by $X = X(F)$. Let $\hat{X} = \hat{X}(F)$ denote the sample median. The (sampling) cdf of $\hat{X}$ is denoted by $M(x) = M(x : F) = \Pr[\hat{X} < x]$.

The distribution of the sample median can be written as a functional depending on the F of the data,

$$M(x) = G(F(x)) \tag{1}$$

where the G-function is[1],

Gib Bassett
Department of Finance, University of Illinois at Chicago
2431 University Hall, 601 S. Morgan Street, Chicago, IL 60607
e-mail: gib@uic.edu

[1] $\hat{X} < x$ 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-

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

While the analysis of the sample mean is about sums of random variables, the analysis of the sample median has to do with the behavior of functionals like G.

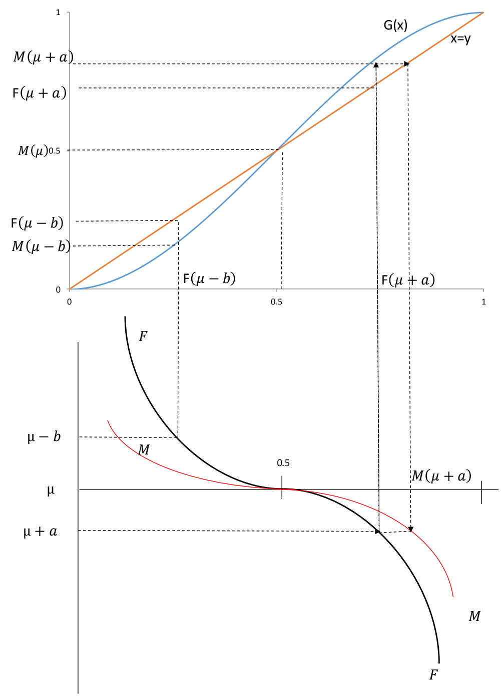

Fig. 1 Mapping from F to M, via G

The G function is depicted in Figure 1 along with schematics showing how G turns F into M. G is seen to be continuous, increasing, with fixed points at $G(1/2) = 1/2$, $G(0) = 0$, and $G(1) = 1$, and $G(w) > w$ for $1/2 < w < 1$, $G(w) < w$ for $0 < w < 1/2$. (These features hold for the median G-functions with $n > 3$ that will be considered later).

One consequence of the G properties is that the median of M and the median of F will always be the same; see appendix. If the median of the data F is an interval

three can occur in 1 way, which has probability, $F(x)^3$. So, $M(x) = F(x)^3 + 3F(x)^2(1-F(x)) = 3F(x)^2 - 2F(x)^3 = G(F(x))$. Note that $G(x)$ is the distribution of the median when the data is uniformly distributed on [0,1].

$[\mu^-, \mu^+]$, then the median of M will be the same interval. If the median of F is a point, $\mu = \mu^+ = \mu^-$, then the median of M will also be μ. The sample median being median-unbiased is analogous to the sample mean being mean-unbiased.

With the median of F and M the same, from now on without loss of generality the data will be centered so that either its unique median is 0, or 0 is in the interval of medians, $[\mu^-, \mu^+]$.

Another feature following directly from the G-properties is that the sample median will always be more concentrated than the data around their common median[2]. That is, the probability that the sample median is in the interval, $[\mu^- - b, \mu^+ + a]$, $b > 0$, $a > 0$, is always greater than the probability that the data is in the interval; see appendix. This is indicated in the figure by, $M(x) > F(x)$, $F(x) \in (1/2, 1)$, and $M(x) < F(x)$, $F(x) \in (0, 1/2)$. Among other things, this means the sample median will always be a better estimate of the population median than the estimate that throws away all but one of the observations[3].

Finally, consider a rescaling of the sample median so that not only its location and scale, but also its shape is the same as the data. That is, for a rescaling factor $\lambda > 1$ there is an H for the data such that the rescaled distribution of the median is the same as H,

$$\lambda \hat{X}(H) \stackrel{d}{=} X(H). \tag{3}$$

Or, in terms of the cdfs, $M[x : H(\lambda^{-1}x)] = H(x)$; hence from (1), H solves the functional equation,

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

Such an H will be analogous to the normal distribution for the sample mean. With normal data (and $n = 3$) the sample mean (scaled by $\sqrt{3}$) has the same normal distribution. In a similar fashion, with H-data the sample median has the same H distribution. A picture illustrating the relationship between H and G is shown in Figure 2.

In the case of the sample mean the set of distributions that reproduce themselves in this fashion define the symmetric stable distributions; see e.g., [3] p.169. The normal is best known. Other scaling factors lead to different distributions; for example, a Cauchy distribution for the data results in a Cauchy distribution for the sample mean, where the scaling factor is, $\lambda = 1$.

Similarly, a distribution H that satisfies (4) will be referred to as median stable. In this paper the idea of median stable distributions is introduced and an initial investigation is begun into what they look like. The simplest, $n = 3$ case is continued in Section 2, and serves as a template for $n > 3$ considered later.

Discussion. Motivated by the sample mean, the classical presentation of symmetric stable distributions is in terms of sums of i.i.d random variables such that the

---

[2] Except in the trivial case in which there is no variation in the data around the median in which case $F = M$; see appendix.

[3] Note that this is not true for the sample mean. Unless tail conditions on the data are ruled out, averaging the data can be worse than the estimate based on a single observation; for examples, see, [3] p.172 . If averaging is the alternative, the advice to never put all your eggs in one (observation) basket is a mistake.

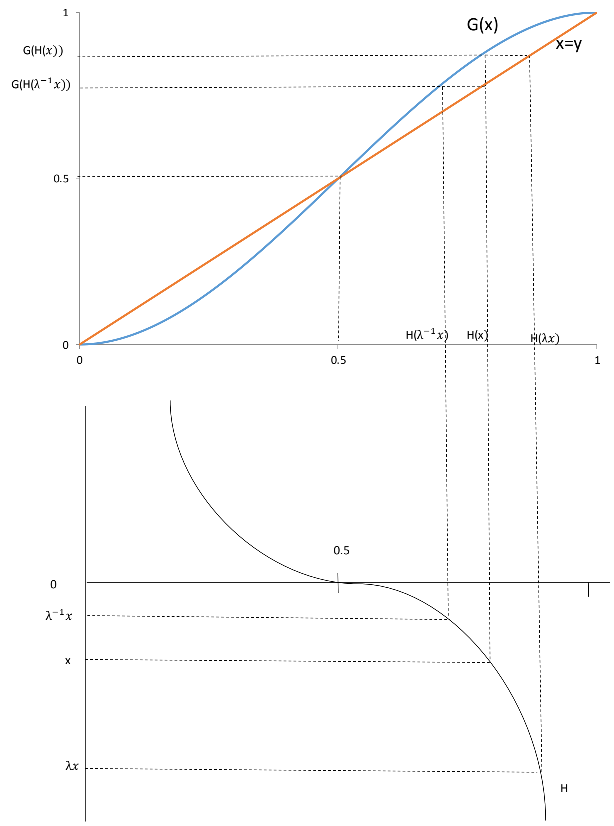

Fig. 2 The relationship between H and G

rescaled sum has the same distribution as the common distribution of the summands. Rather than a definition in terms of ”sums”, a definition in terms of the sample mean lends itself to extensions of the idea of stable distribution. That is, when $\hat{X}(H)$ is the sample mean (given i.i.d data H), the definition of a (mean) stable distribution is (3). Substituting the sample median for the sample mean in this definition leads to the median stable distributions considered here.

One difference between the median and mean stable distributions concerns the way they depend on n, the number of observations. Mean stable distributions do not depend on–are the same–for all n: for $n = 3$ the normal is stable, and it is also stable for $n > 3$. In contrast, as discussed in Section 3, there are different median stable distributions for each n.[4]

Another contrast between mean and median stable distributions is that there are mean stable distributions with the λ-scaling factor less than one; the sample mean with fat-tailed data has to be scaled by a $\lambda < 1$ because it is more dispersed than the data. As mentioned above, since the sample median is more concentrated than the data, its scaling factor will always be greater than one.

2 More Small Data

When noting the dependence of the cdf H defined by (4) on λ, write, $H(x:\lambda)$. The next result summarizes general properties of $H(x:\lambda)$ that follow directly from the shape of G.

Theorem 1. *For each $\lambda > 1$, $H(x:\lambda)$ is symmetric, continuous, increasing, with $0 < H(x:\lambda) < 1$, and $H(1/2:\lambda) = 0$. An $H(x:\lambda)$ determines a scale family of cdfs: if $H(x:\lambda)$ solves (4) then so does $H^*(x:\lambda) = H(\sigma x:\lambda)$, $\sigma > 0$. An $H(x:\lambda_0)$ determines all ($\lambda = \lambda_0^{1/\alpha}, \alpha > 0$), solutions to (4) via, $H(x:\lambda_0^{1/\alpha}) = H(x^\alpha:\lambda_0)$.*

For the last part, let $L(x) = H(x^\alpha:\lambda_0)$ so $L(x) = G(H(\lambda_0^{-1}x^\alpha:\lambda_0)) = G(H((\lambda_0^{-1/\alpha}x)^\alpha:\lambda_0)) = G(L(\lambda_0^{-1/\alpha}x))$, but this says $L(x)$ solves the functional equation with $\lambda = \lambda_0^{1/\alpha}$.

Note that since H is increasing the stable H's all have a unique median.

The derivative of H is; $h(x:\lambda) = g(H(\lambda^{-1}x))h(\lambda^{-1}x)\lambda^{-1}$, which at $x = 0$ says, $h(0:\lambda) = g(0)h(0)\lambda^{-1} = sh(0)\lambda^{-1}$, so if the density is positive at the median the scaling parameter for the stable distribution will be $\lambda = s$. Since the solutions of (4) for any λ can be obtained from, $H(x:\lambda_0)$, we focus mostly on the case $\lambda_0 = s$ where the density of H is positive at $x = 0$.

It would be nice to have a way of going from the G function of (2) to an explicit formula for the cdfs implicitly defined by the functional equation (4). An explicit solution of the defining functional equation (4), even for this simplest $n = 3$ case, however is not evident. So as a starting point to understanding what H looks like, an approximation for $h(x:s)$ near zero is presented and compared to the normal distribution. This is followed by expressions for $H(x:s)$ in terms of functional composition, and for $H(x:s)$ as the limit of a sequence of cdfs that are related to the distribution of the remedian.

[4] A median stable distribution with n observations, call it H_n, however, is a limiting distribution but, rather than the median, for the median-like (base n) remedian estimator; see section 3 below and [5]

2.1 $H(x:s)$, x near zero

Write the derivative of $G(w)$ as $g(w) = s[1-4(w-1/2)^2]$. The log of the derivative of the functional equation is then given by, $\log h(sx)) - \log(h(x)) = \log(1-4(H(x)-1/2)^2)$. Dividing by x^2 and taking limits as x goes to zero means the RHS is $-4h(0)^2$, while for the LHS,

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

So,

$$\lim_{x\to 0} \frac{\log(h(x))}{x^2} = \frac{-4h(0)^2}{s^2-1} = -3.2h(0)^2$$

Hence,

Theorem 2. $h(x:s) = \exp(-3.2x^2 h(0)^2 + o(x))$.

This density whose value at $x = 0$ is 1 can be compared to the normal density whose value at 0 is also 1, or $n(x:0,\sigma = 1/\sqrt{2\pi}) = \exp(-\pi x^2)$. Hence the stable median and normal densities near zero are both proportional to x^2 with the normal exponent being $\pi \approx 3.14$ whereas the median stable exponent is 3.20.

2.2 H as the composition of functions

The equation (4) "at" $s^{-1}x$ on the RHS says, $H(s^{-1}x) = G(H(s^{-2}x))$ which on substituting into the LHS gives $H(x) = G(G(H(s^{-2}x))) = G^{(2)}(H(s^{-2}x))$ where $G^{(2)}(x)$ is the composition of G two-times. Continuing in this fashion, $H(x) = G^{(k)}(H(s^{-k}x))$, $k = 1,2,\ldots$. Further, since G is increasing its inverse is well-defined so that (4) says, $G^{-1}(H(x)) = H(s^{-1}x)$, and substituting as above gives, $H(x) = G^{(-k)}(H(s^k x))$, where $G^{(-k)}(w)$ is the composition of $G^{(-1)}(w)$, k-times. Hence, for $k = \pm 1, \pm 2, \ldots$, a median stable distribution satisfies, $H(x) = G^{(k)}(H(s^{-k}x))$, or $H(s^k x) = G^{(k)}(H(x))$. This can be, in turn, extended to rational and then real values of k by defining fractional composition via, for example, $G^{(1/2)}$ as $G^{(1/2)}(G^{(1/2)}(w)) = G(w)$.

This maps out all solutions to the functional equation. Pick w_0 in $(1/2,1)$, and $x_0 > 0$ so that

$$H(s^k x_0) = G^{(k)}(w_0).$$

This identifies the $H(x), x > 0$ in the scale family of cdfs satisfying (4)with $\lambda = s$ such that $H(x_0) = w_0$. (For $x < 0$, H is determined by symmetry, $H(x) = 1 - H(-x)$).

The tail of $H(x)$, $x \to \infty$, is thus seen to depend on the rate at which $G^{(k)}(w_0) \to 0$, $k \to \infty$. This rate can be determined from a G-inequality involving monomials, which can be used to compute and bound $G^{(k)}(w_0)$; see appendix. It gives,

Theorem 3.

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

where $0 < c < \infty$*, and* $s^\alpha = 2$.

Proof. See appendix.

This limit means a median stable distribution has exponential, not fat, tails.

2.3 *H* as the limit of $H_k(x)$

Consider $H(x) = G^{(k)}(H(s^{-k}x)$, since $H(s^{-k}x) = 1/2 + h(0)s^{-k}x + o(s^{-k}x)$, define the sequence of cdfs, $H_k(x) = G^{(k)}(1/2 + s^{-k}x)$, $|s^{-k}x| < 1/2$, $k = 1, 2, ...,$. The limit of the sequence of cdfs is is the median stable distribution in which, $h(0) = 1$. [5].

Table 1 in the appendix shows the values of $H_k(x)$ for various x and k. For comparison, it also shows the values of the normal distribution whose density at zero is 1. The comparison shows $H_k(x)$ is very close to normal even for moderate k. (But the limit is not a normal distribution because the normal does not satisfy the functional equation (4)).

3 Stable Median Distributions and the Remedian

A stable distribution for the mean does not depend on–is the same–given any number of observations. In addition, the limiting distribution of the sample mean with arbitrary, not necessarily stable, data will necessarily be a stable distribution. The property of the mean that makes its associated stable distributions the same for any number of observations is its recursive property in which the mean of means is the mean. For the regular median, as is well-known, the median of medians is not the median, and as a result median stable distributions are n-dependent. The recursive property for the mean however does hold for the median-like, remedian estimator[6]. As a result, the stable distribution for the (base b) remedian (which is the median stable distribution for b observations) is the same for all $n > b$.

The remedian is defined recursively as the median of medians. Its base-3 version is the ordinary median given n=3 observations. For $n = 3^2$ write the data as a 3×3 array X_{ij}, $i = 1, ..., 3$, $j = 1, ..., 3$. The base-3 remedian for $n = 3^2$ is given by, $\hat{X}_{3^2} = \hat{X}_3(\hat{X}_{1.}, ..., \hat{X}_{3.})$, where $\hat{X}_{i.} = \hat{X}_3(X_{i1}, ..., X_{i3})$, $i = 1, ..., 3$. That is, the $n = 3^2$ version is the same as $\hat{X}_3$, but on a set of "data" that are themselves medians. In a similar fashion the estimate for $n = 3^k$ is recursively defined as the median of the "data", $(\hat{X}_{1\hat{X}_{(k-1)}}, \hat{X}_{2\hat{X}_{(k-1)}}, \hat{X}_{3\hat{X}_{(k-1)}})$.

[5] $H_K(x)$ is the distribution of the remedian (base 3) with $n = 3^K$ data that are uniform on $[0, 1]$; see citeRousseeuw

[6] This contrasts with generalizing the median (the .5 quantile) via its M-estimator representation as in [1], [4]

The cdf of the remedian given i.i.d. data with cdf F is, $Pr[\hat{X}_{3^k}(F) < x] = G^{(k)}(F(x))$, see [5].

Similar to the definition of mean and median stable distributions, define a remedian stable distribution so that the scaled-by-λ sampling distribution of the estimate is the same as the data. Let $H_n(x : \lambda_n)$ be notation for such a remedian stable distribution.

Since the remedian of remedians is the remedian, an H_n that is $\hat{X}_n$-stable with n observations will be also stable with n^2 observations. That is, $H_{n^2}(x : \lambda_{n^2}) = H_n(x : \lambda_n^2)$. The stable distribution is the same, and the scaling parameter for n^2 observations is just $\lambda_{n^2} = \lambda_n^2$, namely the squared value of the scaling parameter with n observations; see appendix.

4 Median Stable for $n > 3$

Let $\hat{X}_r$ denote the median given $n = 2r+1$, i.i.d F-distributed observations. The sampling distribution of $\hat{X}_r$, like the $r = 1$ case discussed above, is a functional of the data, $G_r(F)$. It is convenient to write this G_r function as a transformation of its $G = G_1$ version in equation (2).

Write the G function of (2) as, $G(w) = \int_0^w g(t)dt$, and consider the transformation to a new G function by raising the integrand to the power r, and scaling the result so that it integrates to 1:

$$G_r(w) = \frac{\int_0^w g(t)^r dt}{\int_0^1 g(t)^r dt} = C_r^{-1} \int_0^w g(t)^r dt$$

where $C_r = \int_0^1 g^r(t)dt = 4^r s^r \int_0^1 [t(1-t)]^r dt = 4^r s^r \frac{\Gamma^2(r+1)}{\Gamma(2r+2)} = \frac{r!r!}{(2r+1)!}$. The sampling cdf of the median is given by, $M_r(x) = G_r(F(x))$; for this and many additional results about the median see e.g., [2].

Let the median stable distributions for $n = 2r+1$ be denoted by $H_r(x)$; as in section 1 they are defined by a functional equation given by $H_r(x : \lambda_r) = G_r(H_r(\lambda_r^{-1} x))$.

Consider the stable cdfs with positive density at the median. As in the $r = 1$ case this entails $\lambda_r = s_r$ so that the functional equation is: $H_r(x : s_r) = G_r(H(s_r^{-1} x))$ where $s_r = g_r(1/2)$, where $g_r(x)$ denotes the derivative of $G_r(x)$.

The previous results regarding H near zero and H in the tails extend readily to the $r > 1$ case:

Theorem 4.

$$h(x : s) = \exp(-\alpha(r) x^2 h(0)^2 + o(x))$$

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

Proof. See appendix.

Theorem 5.

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

where, $0 < c < \infty$*, and* α *satisfies,* $s^{\alpha} = r + 1$.

Proof. See appendix.

Finally, given data with a positive density at the median we know the limiting distribution of the median;

$$\lim_{n \to \infty} M_n(n^{-1/2}x : F(x)) = G_n(F(n^{-1/2}x)) = N(x : 0, \sigma_M)$$

where $\sigma_M = (2f(0))^{-1}$. In terms of the G function this means

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

(Verify noting

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

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

Appendix

1. Median unbiased for any F.
   μ is a median of F if; (i) $F(\mu + 0) \geq 1/2$ and, (ii) $F(\mu - 0) \leq 1/2$. Since G is increasing with fixed point at $1/2$, $F(\mu + 0) \geq 1/2$, implies $G(F(\mu + 0)) \geq G(1/2) = 1/2$, or $M(\mu + 0) \geq 1/2$. On the other side, $F(\mu - 0) \leq 1/2$ implies $G(F(\mu - 0)) < G(1/2) = 1/2$, or $M(\mu - 0) \leq 1/2$. Hence μ will also be a median of M.
2. No variation in the data around the median.
   The sample mean and data have the same cdf in the trivial case in which there is no variability in the data; there is probability 1 point mass at 0; $\Pr(X = 0) = 1$. The analogous situation for the median occurs when the data is equal to "the" median with probability 1. As in the mean case, this trivially occurs when $\Pr(X = 0) = 1$. But it also occurs when $\Pr(X = \mu^{-}$or $X = \mu^{+}) = 1$ which means, $\Pr(X = \mu^{-}) = \Pr(X = \mu^{+}) = 1/2$. In this case $F = M$, the cdfs are the same.
3. The sample mean is more concentrated around the median than the data.
   Let $a > 0$ such that, $1/2 < F(\mu^{+} + a + 0) < 1$. Since $G(x) > x$ for $1/2 < x < 1$, $1/2 < F(\mu^{+} + a + 0) < G(F(\mu^{+} + a + 0)) < 1$. But $G(F(\mu^{+} + a + 0)) = M(\mu^{+} + a + 0)$, so, $1/2 < F(\mu^{+} + a + 0) < M(\mu^{+} + a + 0) < 1$. Similarly, let $b > 0$ is such that $0 < F(\mu^{-} - b - 0) < 1/2$. Since $G(x) < x$ for $x < 1/2$, $0 < G(F(\mu^{-} + b - 0)) < F(\mu^{-} + b - 1) < 1/2$. Combining gives: for any $a > 0$ such that $1/2 < F(\mu^{+} + a + 0) < 1$, $b > 0$ is such that $0 < F(\mu^{-} + b - 0) < 1/2$:

$$\Pr(\mu^{-} - b < \hat{X} < \mu^{+} + a) > \Pr(\mu^{-} - b < X < \mu^{+} + a).$$

4. Table 1

	$K=1$	$K=2$	$K=5$	$K=10$	$K=15$	Normal
$x=0.00$	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
$x=0.10$	0.59941	0.59915	0.59896	0.59894	0.59894	0.59896
$x=0.20$	0.69526	0.69330	0.69192	0.69179	0.69179	0.69193
$x=0.30$	0.78400	0.77800	0.77395	0.77359	0.77358	0.77397
$x=0.40$	0.86207	0.84981	0.84196	0.84126	0.84124	0.84199
$x=0.50$	0.92593	0.90669	0.89492	0.89389	0.89387	0.89495
$x=0.60$	0.97200	0.94818	0.93368	0.93241	0.93239	0.93371
$x=0.70$	0.99674	0.97538	0.96033	0.95899	0.95896	0.96034
$x=0.75$	1.00000	0.98435	0.96995	0.96862	0.96860	0.96994
$x=0.80$		0.99076	0.97755	0.97628	0.97626	0.97753
$x=0.90$		0.99769	0.98799	0.98692	0.98691	0.98796
$x=1.00$		0.99976	0.99394	0.99312	0.99311	0.99391
$x=1.10$		1.00000	0.99712	0.99655	0.99654	0.99709
$x=1.50$			0.99992	0.99986	0.99986	0.99992

Table 1 $H_K(x)$ for alternative x and K

5. Stability for n^2.
   The LHS of the stability condition for n^2 is, $\lambda_{n^2}\hat{X}_{n^2}(H_{n^2})$. Make the substitution, $H_{n^2} = H_n$, $\lambda_{n^2} = \lambda_n^2$, or, $\lambda_n^2\hat{X}_{n^2}(H_n)$. Now use, $\hat{X}_{n^2} = \hat{X}_n(\hat{X}_{1.},\ldots,\hat{X}_{n.})$ and linear homogeneity so that,

$$\lambda_n^2\hat{X}_{n^2}(H_n) = \lambda_n^2\hat{X}_n(\hat{X}_{1.}(H_n),\ldots,\hat{X}_{n.}(H_n)) = \lambda_n\hat{X}_n(\lambda_n\hat{X}_{1.}(H_n),\ldots,\lambda_n\hat{X}_{n.}(H_n)).$$

   But $\lambda_n\hat{X}_{i.}(H_n) \stackrel{d}{=} X(H_n)$, so,

$$\lambda_n\hat{X}_n(\lambda_n\hat{X}_{1.}(H_n),\ldots,\lambda_n\hat{X}_{n.}(H_n)) \stackrel{d}{=} \lambda_n\hat{X}_n(X_1(H_n),\ldots,X_n(H_n)) \stackrel{d}{=} X(H_n).$$

   The last step following from the fact that H_n is stable.
6. H, x near zero.
   Suppressing the r-subscript, let G denote the G-function with $n = 2r+1$ (as in Section 4), and $H(x)$ a median stable cdf with $\lambda = s$, s denoting the derivative of G at $1/2$. The derivative of the functional equation is $h(sx) = s^{-1}g(H(x))h(x)$, where $g(w) = s[1-4(w-1/2)^2]^r$. Take the log of both sides so, $\log(h(sx)) - \log(h(x)) = r\log(1-4(H(x)-1/2)^2)$. Divide both sides by x^2 and take limits as x goes to zero. The RHS is just $-4rh(0)^2$ and the LHS is

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

   So,

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

where $\alpha(r)=\frac{-4r}{s^2-1}$. When $r=1$, $\alpha(1)=3.2$ as in Theorem 2. As r increases, $\alpha(r)$ decreases and converges to π, the coefficient for the normal density.

7. Tails: $H(x), x\to-\infty$.
Suppressing the r-subscript, let G denote the G-function with $n=2r+1$ (as in Section 4), and $H(x)$ a median stable cdf with $\lambda=s$, s denoting the derivative of G at $1/2$.
Consider the following result for the tail of G. For $0<t<1/2$, write $t(1-t)=\omega t$ where, $1/2<\omega<1$. So,

$$G(w)=C^{-1}\int_0^w[6t(1-t)]^r dt=C^{-1}(6\omega)^r\int_0^w t^r dt=$$

$$C^{-1}(6\omega)^r\frac{w^{(r+1)}}{(r+1)}.$$

So, $G^{(k)}(w)=(A(r)w^{r+1})^{(k)}$ where $A(r)=\frac{C^{-1}(6\omega)^r}{r+1}$. Now verify,

$$(bx^{\alpha})^{(k)}=b^{\frac{1-\alpha^k}{1-\alpha}}x^{\alpha^k}$$

so that,

$$\lim_{k\to\infty}\log\frac{(bx^{(\alpha+1)})^{(k)})}{\alpha+1}=\log x+\alpha^{-1}\log b$$

So,

$$\lim_{k\to\infty}\frac{\log G^{(k)}(w)}{(r+1)^k}=\lim_{k\to\infty}\frac{\log(A(r)w^{r+1})^{(k)}}{(r+1)^k}=\log x+r^{-1}\log A(r)$$

To prove the result, write $x=s^k x_0$, so

$$\lim_{x\to-\infty}x^{-\alpha}\log H(x)=\lim_{k\to\infty}(s^k w_0)^{-\alpha}\log H(-s^k x_0)=\lim_{k\to\infty}s^{-\alpha k}x_0^{-\alpha}\log G^{(k)}(w_0)$$

but, $s^{\alpha}=r+1$ so that

$$=\lim_{k\to\infty}\frac{x_0^{-\alpha}\log(A(r)w^{r+1})^{(k)}}{(r+1)^k}=x_0^{-\alpha}(\log x+r^{-1}\log A(r))=c.$$

References

1. Bassett, G.W., Koenker, R.: The Asymptotic Theory of Least Absolute Error Regression. J. Amer. Statist. Assoc. **73**, 618-622 (1978).
2. David H. A.: Order Statistics. Wiley, New York (1981).
3. Feller, W.: An Introduction to Probability Theory and Its Applications. Wiley, New York (1971).

4. Koenker, R., Bassett, G.W. (1978) Regression Quantiles, Econometrica. 46:33-50.
5. Rousseeuw, P.J., Bassett, G.W.: The Remedian: A Robust Averaging Method for Large Data Sets. J. Amer. Statist. Assoc. **85**, 97-104 (1990).