Omega Analysis

University of Hawaii Mathematics Department Distinguished Lecture Series

The Geometry of Probability Distributions

February 26 2010

Dr. William F. Shadwick
Omega Analysis Limited
William.Shadwick@OmegaAnalysis.com

The Geometry of Probability Distributions

William F. Shadwick
(Joint work with Ana Cascon)
26 February 2010

If F is a cumulative distribution on $[A, B]$ (where either of A or B may be infinite) *and F has a finite mean μ* then the Omega function of F is defined as

$$\Omega(x) = \frac{I_1(x)}{I_2(x)} \tag{1}$$

where

$$I_{1F}(x) = \int_A^x F(z)dz \tag{2}$$

and

$$I_{2F}(x) = \int_x^B 1 - F(z)dz \tag{3}$$

The functions I_{1F} and I_{2F} are analogous to the values of a put and a call with strike x respectively.

$$I_{1F}(x) = E_F(Max(z - x, 0)) \tag{4}$$

and

$$I_{2F}(x) = E_F(Max(x - z, 0)) \tag{5}$$

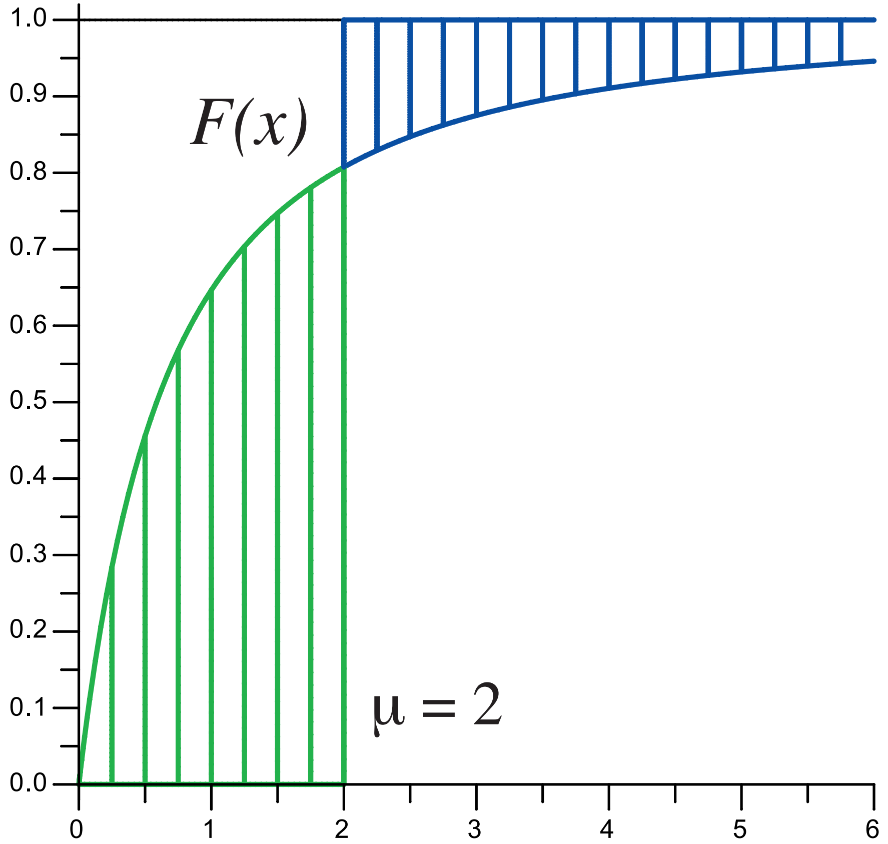
They satisfy a *virtual* ‘put-call parity’ relation

$$I_{2F}(x) - I_{1F}(x) = \mu - x \tag{6}$$

Ω_F may be interpreted as the ratio of the value of the downside to the upside.

The smaller this is, the better off you are.

The upside and downside balance at the mean.



$$I_{1F}(\mu) = I_{2F}(\mu)$$

Properties of Ω_F:

$$\Omega(\mu) = 1 \tag{7}$$

$$\frac{d\Omega_F}{dx} = \frac{1}{I_{2F}} > 0 \tag{8}$$

$$\lim_{x \to A} \Omega_F = 0 \tag{9}$$

$$\lim_{x \to B} \Omega_F = \infty \tag{10}$$

It also follows from the definition and the put-call parity relation for I_1 and I_2 that

$$\Omega_F(x) = 1 + \frac{x - \mu_F}{I_{2F}(x)}. \tag{11}$$

We can now see that

$$\Omega_F = \Omega_G \iff F = G, \tag{12}$$

since $\Omega_F{=}\Omega_G$ implies $\mu_F{=}\mu_G$ and, from equation 11,

$$I_{2F} = I_{2G} \tag{13}$$

Differentiating equation 13 gives $F{=}G$.

We can also use equation 11 to produce an inverse formula for recovering F from Ω_F.

$$F = 1 + \frac{1}{\Omega_F - 1} + \frac{\mu - x}{(\Omega_F - 1)^2}\frac{d\Omega_F}{dx} \tag{14}$$

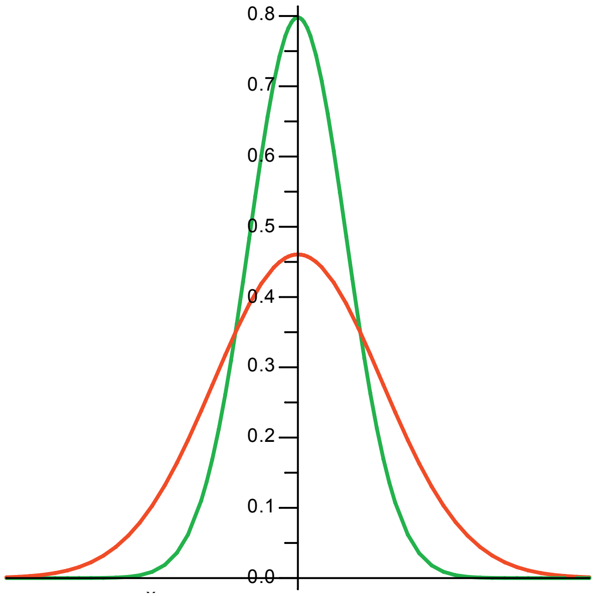
The Standard Dispersion (a.k.a.‘The om’).

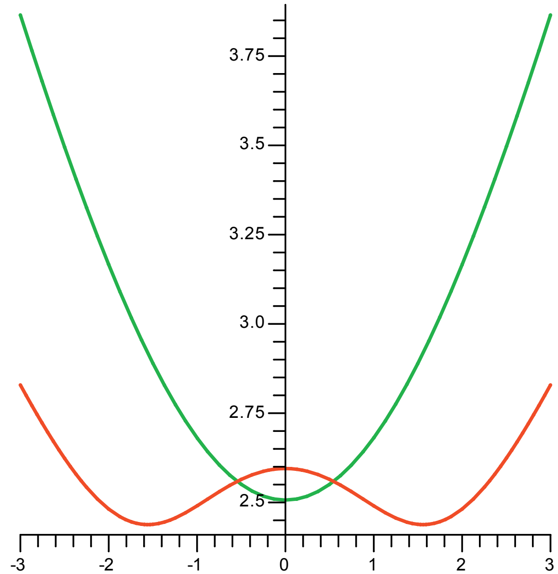
We define the om ω_F by

$$\frac{1}{\omega_F} = \frac{d\Omega_F(\mu)}{dx}. \tag{15}$$

We have

$$\omega_F = I_{1F}(\mu) = I_{2F}(\mu). \tag{16}$$





The om measures spread about the mean. The higher the concentration the higher the value of $\Omega'(\mu)$.

This is a general property. The standard dispersion allows us to refine the Markov inequality:
For any non-negative random variable x with mean μ,

$$probability(x > a) \leq \frac{\mu}{a}. \tag{17}$$

Our version:

$$probability(x - \mu > b) < \frac{\omega}{b}\ . \tag{18}$$

This is always sharper than (17) for a sufficiently large (and there are distributions where 'sufficiently large' is arbitrarily close to the mean.)

Chebychev inequality: For any random variable x with mean μ and variance σ^2,

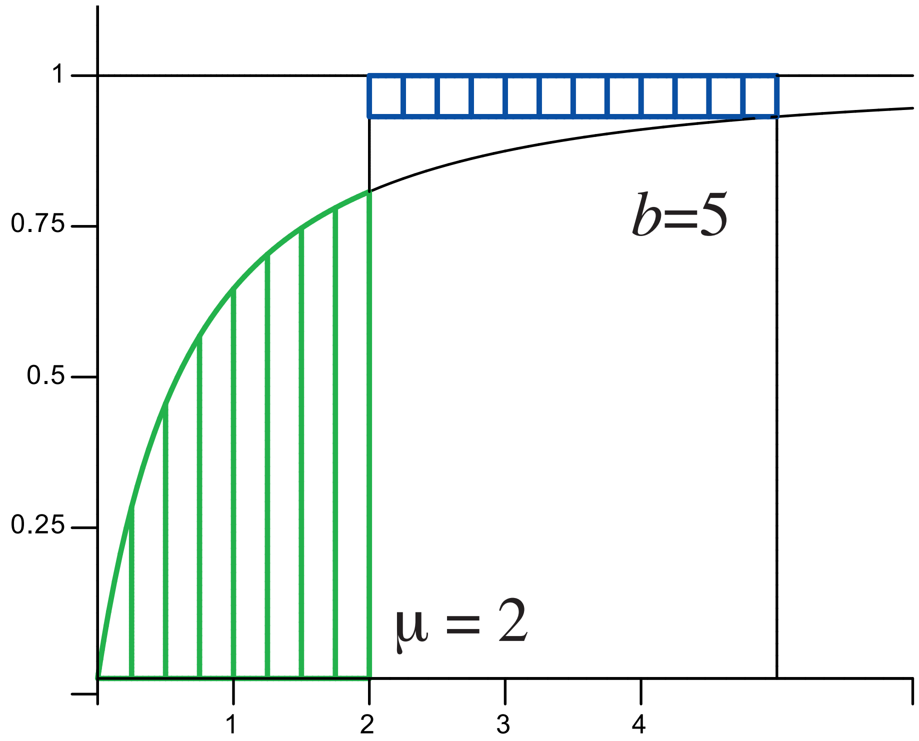
$$probability(|x - \mu| > a) \leq \frac{\sigma^2}{a^2}. \tag{19}$$

Our version:

$$probability(x > \mu + b) < \frac{\omega}{b} \ . \tag{20}$$

and

$$probability(x < \mu - b) < \frac{\omega}{b} \ . \tag{21}$$



$$(1 - F(b))b < \omega \text{ so } probability(x - \mu > b) < \frac{\omega}{b}$$

Properties of the Standard Dispersion

It is easy to check that if

$$\phi : x \to ax + b \tag{22}$$

is a proper affine transformation and

$$F = \phi^* G \tag{23}$$

then

$$\Omega_F = \phi^* \Omega_G \tag{24}$$

It follows from this that ω is translation invariant and scales like the mean.

If two distributions F and G defined on $[A, B]$ have the same mean μ then

$$\int_A^B (F - G)dx = 0. \tag{25}$$

If, in additon they have the same om then

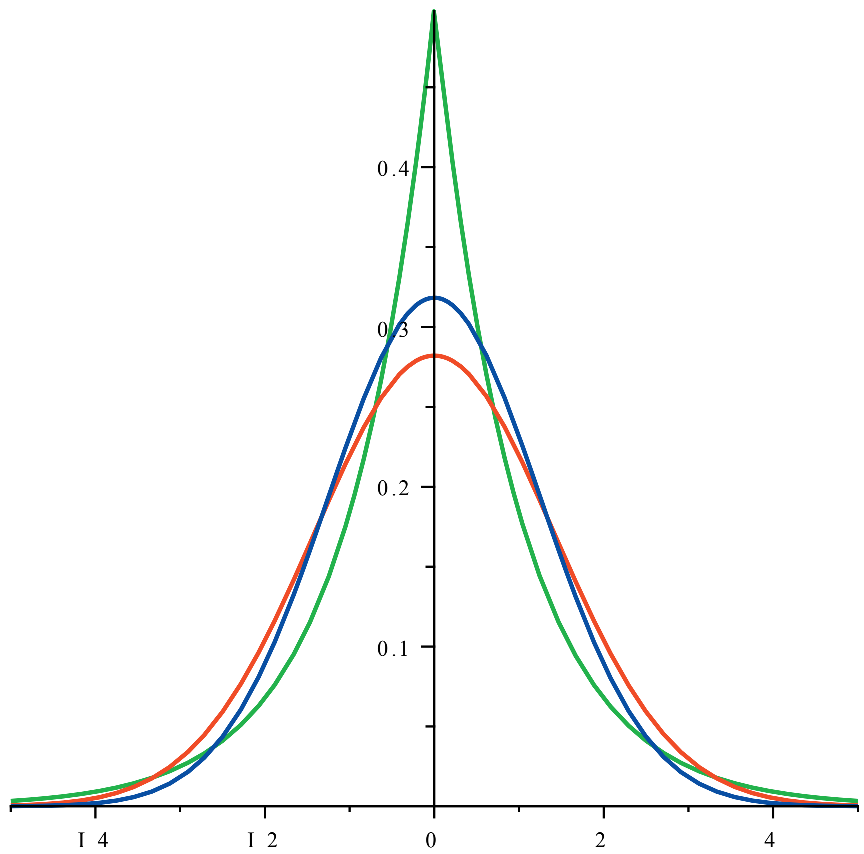
$$\int_A^\mu (F - G)dx = 0 \tag{26}$$

and

$$\int_\mu^B (F - G)dx = 0. \tag{27}$$

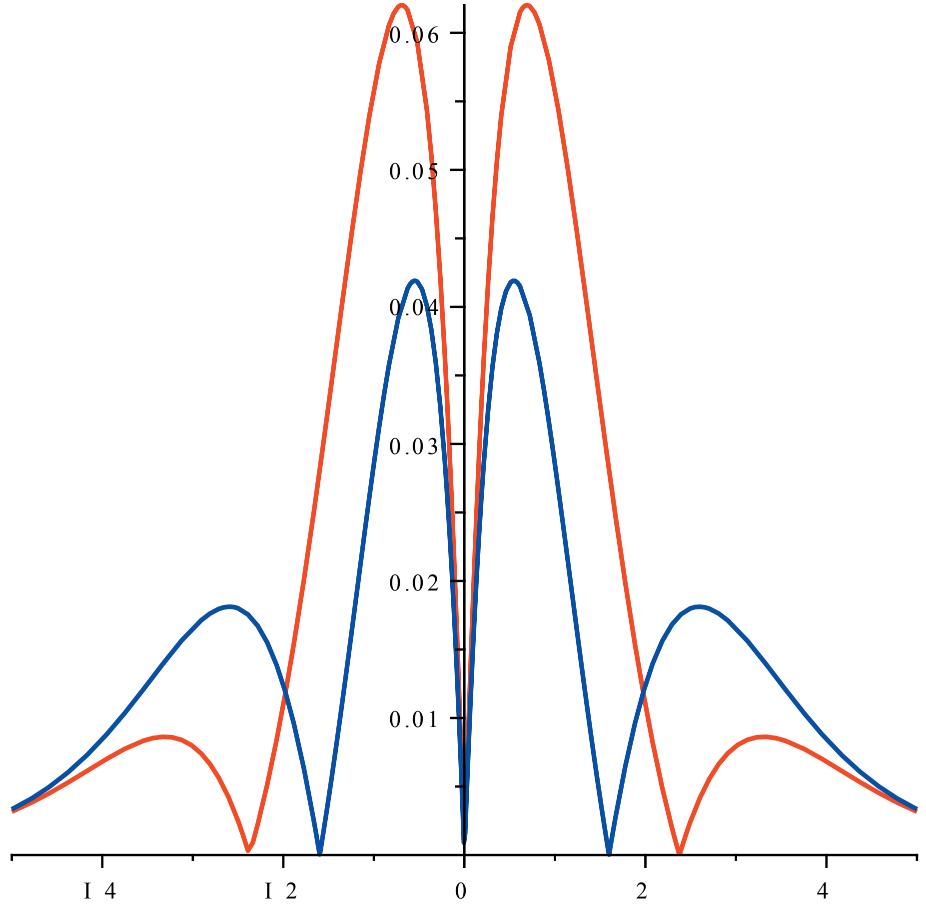
By contrast if they have the same standard deviation then

$$\int_A^B x(F - G)dx = 0. \tag{28}$$



Laplace and Normal Distributions with the same standard deviation (red) and same om (blue).

Omega Analysis



Absolute Errors in Cumulative Distributions same sigma(red) and same om (blue).

Omega Analysis

Standard deviation gives *much* more weight to outliers than the om does.

Example: Returns on the S&P 500 for the last 10,361 days and the same period *excluding* 19 October 1987.

$$\Delta\sigma = 2.23\% \text{ but } \Delta\omega = 0.29\%.$$

It follows from the scaling property of ω that, when the standard deviation is defined, the ratio of standard deviation to standard dispersion is an affine invariant. We have called this the first C-S Character.

Examples:
Uniform distribution: $CS_1 = \frac{4}{\sqrt{3}}$
Normal distribution: $CS_1 = \sqrt{2\pi}$
Laplace distribution: $CS_1 = 2\sqrt{2}$

When distributions have different C-S characters, standard deviation is not a common unit of measurement.

In financial data or instrument measurements subject to noise, a common number of oms as a threshold for outliers or noise can be much more appropriate than using thresholds in standard deviations.

C-S Characters are remarkably robust statistics. Even small data samples give results close to population values. They provide a much more practical measure of 'fat tails' than kurtosis. For the S&P example already cited, Δ kurtosis = 240% but $\Delta CS_1 = 2.8\%$. (Incidentally the CS_1 value of 2.79 *proves* the returns aren't normal.)

These statistics are currently being used in investment risk management and to remove noise from measurements of atmospheric data.
(see e.g. *Cascon and Shadwick IMCA Journal of Investment Consulting Summer 2006 and Summer 2007)*

Properties of the Standard Dispersion Continued

The om is half the mean absolute deviation:

$$\omega_F = \frac{1}{2} E_F(|x - \mu_F|). \tag{29}$$

Addition formula: If $F = \sum_{i=1}^{n} a_i F_i$ where the a_i are positive and sum to 1, then $\mu_F = \sum_{i=1}^{n} a_i \mu_i$ and

$$\omega_F = \sum_{i=1}^{n} a_i \omega_i + \sum_{i=1}^{n} a_i \int_{\mu_i}^{\mu_F} F_i dx \tag{30}$$

The Geometry Determined by the Affine Group

The proper affine group on the line is the largest subgroup of the diffeomorphism group that preserves the property of having a finite first moment.

This group is responsible for the two parameters which appear in almost all the standard textbook distributions.

The assignment $F \to log(\Omega_F)$ commutes with the action of the affine group.

Thus the geometry defined by Omega functions under this group induces geometry on the space of distributions with finite first moment.

This turns out to have some remarkable structure.

The Equivalence Problem

Coframe on the affine group

$$\theta_1 = y dx \tag{31}$$

$$\theta_2 = \frac{dy}{y} \tag{32}$$

Coframe adapted to the (invariant) $I = log(\Omega(x))$

$$\omega_1 = dI = \frac{\Omega_x}{\Omega} dx \tag{33}$$

$$\omega_2 = \frac{dy}{y} \tag{34}$$

We have

$$\omega_1 = J\theta_1 \tag{35}$$

where

$$J = \frac{\Omega_x}{y\Omega} \tag{36}$$

Any diffeomorphism that preserves both co-frames also preserves J so we have a second functionally independent invariant.

Now

$$dI = \omega_1 \tag{37}$$

and

$$\frac{dJ}{J} = H\omega_1 - \omega_2, \tag{38}$$

where

$$H = \frac{\Omega_{xx}\Omega}{\Omega_x^2} - 1. \tag{39}$$

But H depends only on x so it follows that the remaining information is in the functional dependence of H on I.

The exceptional equivalence classes are the ones for which H is a constant function c.

So far we have not used the information that there is a distribution F for which $I = log(\Omega_F)$. This condition limits the possible constant values of H.

When $I = log(\Omega_F)$, we can evaluate H at μ:

$$H(\mu) = 2F(\mu) - 1 \tag{40}$$

from which it follows that $-1 < c < 1$.

The normal forms corresponding to the exceptional cases $c = 0$, $-1 < c < 0$ and $0 < c < 1$ are

$$\Omega = e^x \tag{41}$$

$$\Omega = x^{\frac{-1}{c}} \tag{42}$$

and

$$\Omega = (-x)^{\frac{-1}{c}} \tag{43}$$

respectively.

Omega Analysis

We denote the corresponding exceptional distributions by CSS_0 and CSS_λ respectively, where $\lambda = \frac{-1}{c}$. (These first appeared in a paper with B.A.Shadwick, before we solved the equivalence problem.)

CSS_0 has $\mu = 0$ and $\omega = 1$. It has finite moments of all orders. Its standard deviation is $\pi\sqrt{\frac{2}{3}}$, which provides a probabilistic definition of π.

The‘Cascon-Shadwickian’ CSS_0 is a global attractor under a natural action on the space of Omega functions. This results in a new Central Limit Theorem.

For any integer $N > 1$, the assignment $\Omega_F \to (\Omega_F)^N$ commutes with the natural action of the affine group.

The inversion formula (15) shows (after a long calculation) that Ω_F is the Omega function of a distribution F_N. As N tends to ∞, F_N tends to $CSS_0(\mu_F, \frac{\omega_F}{N})$.

The proof of this result depends only on the behaviour of the invariant H for Ω^λ.

$$H(\Omega^\lambda) = \frac{1}{\lambda} H(\Omega) \tag{44}$$

It is an exercise to show that H is a bounded function, so as $\lambda \to \infty$, $H(\Omega_\lambda) \to 0$.

The convergence is remarkably quick as the following movie shows.

Because everything commutes with the affine group action, we may as well pass to the quotient in which:

Any distribution defined on R is replaced by its normal form with $\mu = 0$ and $\omega = 1$

Any distribution defined on a half line bounded below is replaced by a normal form on $[0, \infty)$ with $\mu = 1$ (We have to use both affine parameters to translate the lower bound to 0 and re-scale the mean to 1 so we cannot choose the om.)

Any distribution defined on a half line bounded above is replaced by a normal form on $(-\infty, 0]$ with $\mu = -1$ (We have to use both affine parameters to translate the lower bound to 0 and re-scale the mean to -1 so we cannot choose the om.)

Any distribution defined on a compact interval is replaced by a normal form on $[-1, 1]$. (This uses both affine parameters so we cannot choose the mean or the om.)

In this quotient, the Cascon-Shadwickian may be replaced by a δ-function with $\mu = 0$ and $\omega = 0$.

(Note that this is quite distinct from a Gaussian δ-function as it is the limit of distributions whose CS character is $\pi\sqrt{\frac{2}{3}}$ and not $\sqrt{2\pi}$.)

There is nothing (aside from some details in the proof!) to restrict the powers of Omega functions to integers (as the movie illustrated.)

In particular one may take fractional powers, which leads to a new affine invariant: the age of a distribution.

The age is defined as the reciprocal of the smallest power of Ω_F that is still the Omega function of a distribution. This is clearly an affine invariant and only the exceptional distribution CSS is ‘ageless’.

Examples:
Normal distribution: $Age = \frac{\pi}{\sqrt{\pi(\pi-3)}}$
Uniform distribution: $Age = 2$
Exceptional distributions:
CSS_λ: $Age = \frac{1}{\lambda}$
CSS_0: $Age = \infty$.

Distributions age in different ways.

Some unimodal distributions (like the normal distribution) start off as bimodal and age to unimodal.

Some originally tri-modal age to unimodal (like the Laplace distribution).

Others (like the exceptional distributions) remain unimodal throughout.

Now we turn to the '3 Types Theorem' of Fisher and Tippett. If X is the maximum value of a sample of size n i.i.d draws of a random variable with distribution G than $Prob(X < r) = G^n(r)$.

The Statistician's *Stability Postulate* is that if there is a limiting distribution G as $n \to \infty$ (up to an action of the proper affine group) then the limiting distribution must be its own attractor so

$$G^n(x) = G(a_n x + b_n) \tag{45}$$

The Geometer's version of the Stability postulate is that for all $\alpha \geq 1$

$$G^{\alpha}(x) = G(g_{\alpha}x) \tag{46}$$

where g_{α} is in the (proper) Affine group and $g_1 = Identity$.

We have already solved the Equivalence Problem for an invariant function I, which we will now take to be $I = log(G)$ under the action of this group.

We know that the solution depends on the functional relation $H = H(I)$ where (Equation 39)

$$H = \frac{G_{xx}G}{G_x^2} - 1 = \frac{G_{xx}G - G_x^2}{G_x^2} \tag{47}$$

If we write this in terms of derivatives of $I = log(G)$ we have

$$H(I) = \frac{I_{xx}}{I_x^2} \tag{48}$$

We know that $I(G^\alpha) = log(G^\alpha) = \alpha I(G)$ and the Stability Postulate ensures that raising G to the power α does not change the equivalence class.

But we also have (Equation 44)

$$H(\alpha I) = \frac{1}{\alpha} H(I). \tag{49}$$

Differentiating this with respect to α and evaluating at the Identity gives

$$H(I) + I \frac{dH}{dI} = 0. \tag{50}$$

So $H(I) = \frac{c}{I}$ for some constant c.

Combining this with Equation (48) we have

$$\frac{I_{xx}}{I_x} = c\frac{I_x}{I} \tag{51}$$

or

$$\frac{d[log(I_x)]}{dx} = \frac{d[log(I^c)]}{dx} \tag{52}$$

which we solve to find the normal forms corresponding to the Stability Postulate. Integrating once gives

$$\frac{I_x}{I^c} = c_1 \tag{53}$$

for another constant c_1.

So far we have not used the fact that $I = log(G)$ where G is a probability distribution. This means $0 \leq G(x) \leq 1$ so $log(G)$ is negative as is $log(-log(G)) = log(-I)$.

Using these restrictions, in the case that $c = 1$ Equation 53 gives $I = -exp(-c_1 x + c_2)$ and finally

$$G(x) = exp(-e^{-c_1 x + c_2}) \tag{54}$$

This is the Gumbel distribution and its Equivalence Class is given by

$$H(I) = \frac{1}{I} \tag{55}$$

The remaining cases which produce the Fréchet and Weibull distributions are left as exercises. Their Equivalence Classes are given by

$$H(I) = (1 + \frac{1}{\alpha})I \tag{56}$$

and

$$H(I) = (1 - \frac{1}{\alpha})I \tag{57}$$

respectively.

As $\alpha \to \infty$ both the Fréchet and Weibull Equivalence Classes approach the Gumbel Equivalence Class.

These three cases include all of the relations of the form

$$H(I) = \frac{c}{I} \tag{58}$$

so these are all the distribution equivalence classes that satisfy the Stability Postulate.