New Statistics from Omega Functions

Ana Cascon
(Joint work with W.F. Shadwick)
RiO Dec 2019

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

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

Omega Analysis

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

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

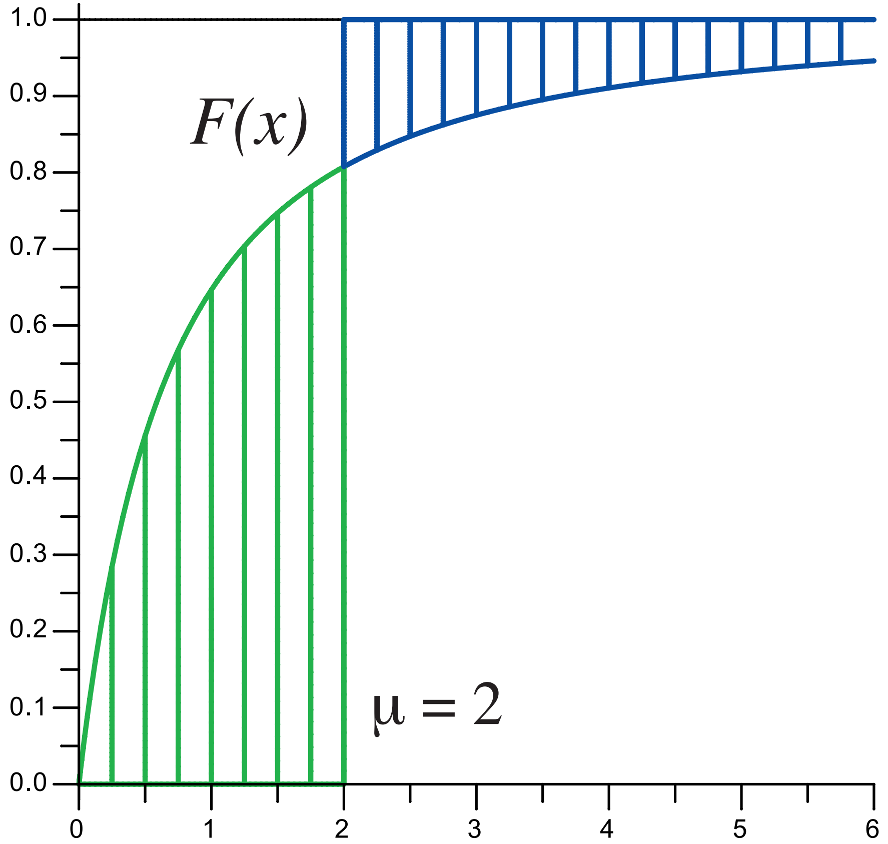
They satisfy a *virtual* ‘put-call parity’ relation

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

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:

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

$$\frac{d\Omega_F}{dx} > 0 \tag{8}$$

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

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

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

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

We can now see that

$$\Omega_F = \Omega_G \iff F = G, \qquad (12)$$

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

$$I_{2F} = I_{2G} \qquad (13)$$

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

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

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

Omega Analysis

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

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

For any non-negative random variable, we have $\omega \leq \mu$.

The following inequalities show that ω measures the spread around the mean.

$$probability(x - \mu > b) \leq \frac{\omega}{b} \ . \tag{17}$$

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

The smaller the value of ω, the more concentrated the distribution is around the mean.

The standard dispersion allows us to refine the simple Markov inequality:

For any non-negative random variable x with mean μ,

$$probability(x - \mu > b) \leq \frac{\mu}{\mu + b}. \tag{19}$$

Our version:

$$probability(x - \mu > b) \leq \frac{\omega}{b} \,. \tag{20}$$

This is always sharper than (19) for b 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,

$$probability(|x-\mu| \geq b) \leq \frac{\sigma^2}{b^2}. \tag{21}$$

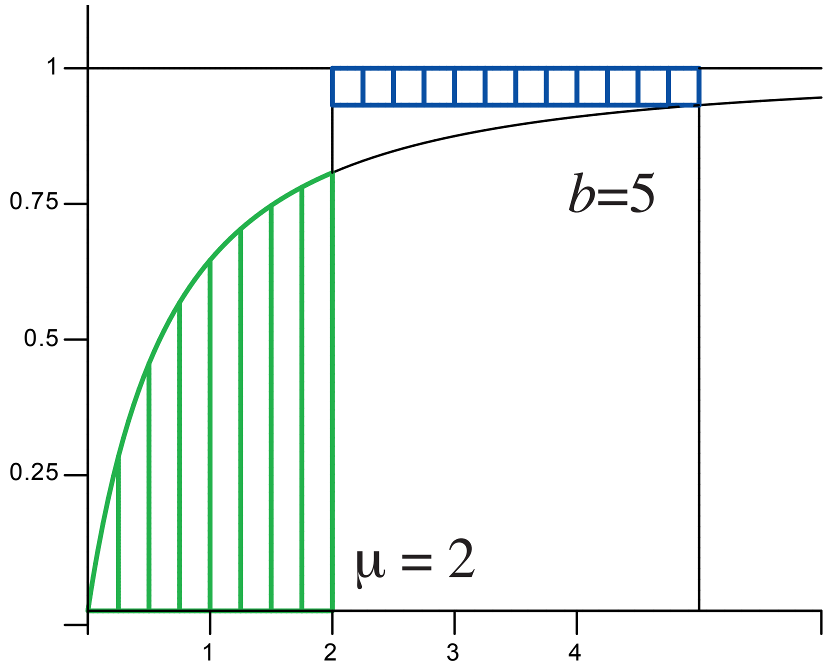
Our version bounds the upside and downside separately:

$$probability(x-\mu > b) < \frac{\omega}{b}\ . \tag{22}$$

and

$$probability(x-\mu < -b) < \frac{\omega}{b}\ . \tag{23}$$

This depends only on the existence of μ.



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

The proofs of our inequalities (and of the fact that for non-negative random variables $\omega \leq \mu$) follow directly from this picture.

Omega Analysis

Properties of the Standard Dispersion

It is easy to check that if

$$\phi : x \rightarrow ax + b \tag{24}$$

is a proper affine transformation and

$$F = \phi^* G \tag{25}$$

then

$$\Omega_F = \phi^* \Omega_G \tag{26}$$

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

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

If, in additon they have the same om then

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

and

$$\int_\mu^B (F - G)dx = 0. \tag{29}$$

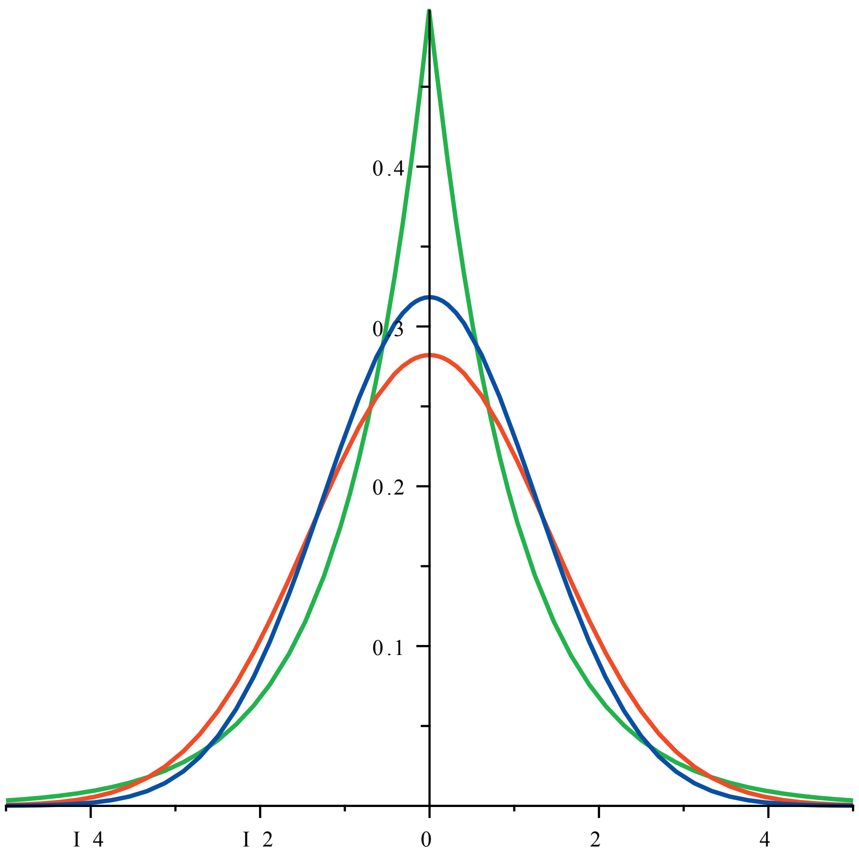
By contrast if they have the same standard deviation, we only know that

$$\int_A^B x(F - G)dx = 0. \tag{30}$$

This means we can get better fits by matching the mean and the om than by matching the mean and the standard deviation.

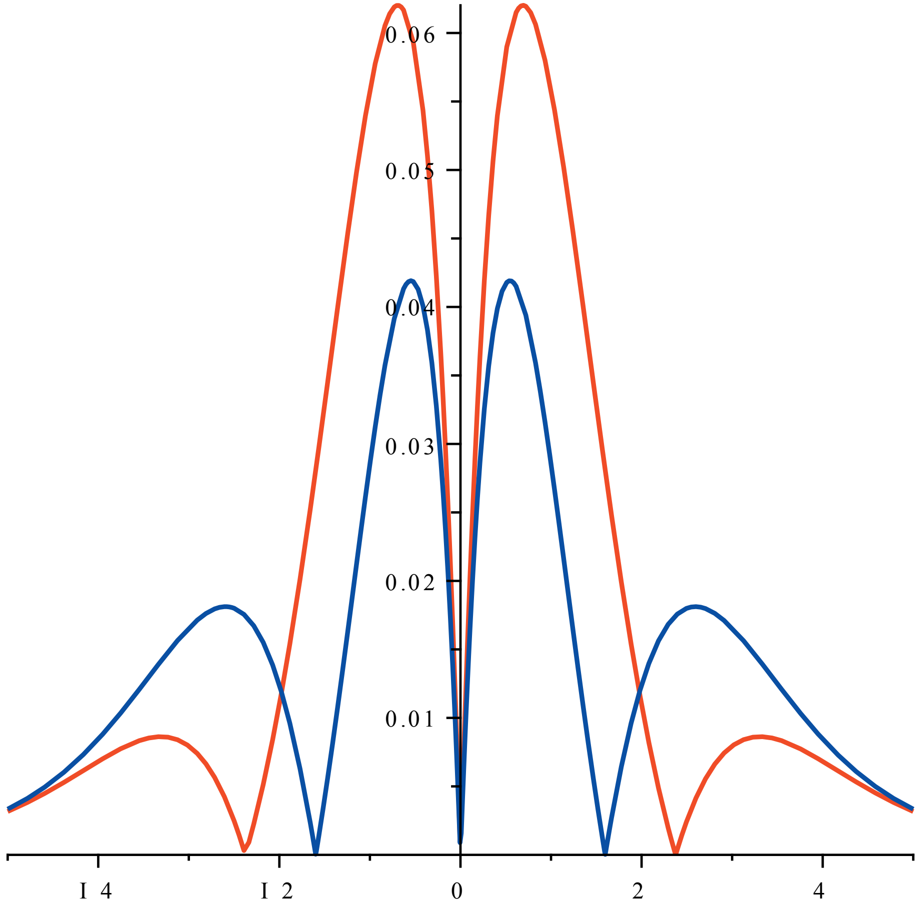
We illustrate this by fitting a normal distribution to a Laplace distribution in both ways.

Fitting with the om gives a better match to the increased concentration around the mean.



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).

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: $CS_1 = 2\sqrt{2}$

C-S Characters are remarkably robust statistics. Small data samples give results close to population values.

For example, even with a sample of 250 data points, 90% of the time the value for a normal distribution will be between 2.44 and 2.58.

The population value is $CS_1 = \sqrt{2\pi} \approx 2.51$.

C-S Characters are remarkably robust statistics. Small data samples give results close to population values.

Normal Distribution	population value 2.51		
Number of Draws	250	500	1500
lowest 5% cutoff	2.44	2.46	2.48
highest 5% cutoff	2.58	2.56	2.53
lowest 10% cutoff	2.45	2.47	2.48
highest 10% cutoff	2.56	2.54	2.53
Laplace Distribiution	population value 2.83		
Number of Draws	250	500	1500
lowest 5% cutoff	2.68	2.73	2.77
highest 5% cutoff	2.97	2.92	2.88
lowest 10% cutoff	2.71	2.75	2.78
highest 10% cutoff	2.92	2.90	2.87

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.

Properties of the Standard Dispersion Continued.

We discovered the om using the geometry of Omega Functions. But there's a classical way to describe it.

The om is half the mean absolute deviation:

$$\omega_F = \frac{1}{2} E_F(|x - \mu_F|). \tag{31}$$

This follows from the fact that $I_1(\mu) = \omega = I_2(\mu)$ and integration by parts.

We can use the geometry of Omega Functions to generate additional useful statistics.

We illustrate this with a Shape Score which can be applied to empirical financial returns data.

Unlike the skewness, which requires 3 finite moments, our statistic is defined whenever the mean exists.

Omega Analysis

Let

$$\Omega(r, \mu, \lambda) = [\frac{(1+r)(1-\mu)}{(1-r)(1+\mu)}]^{\lambda}. \qquad (32)$$

This is a family of Omega Functions defined on the interval $[-1, 1]$ with mean μ.

It is immediate from the definition that $\omega = \frac{1-\mu^2}{2\lambda}$.

When $\lambda = 1$ this is the Omega Function of the Bernoulli distribution.

Let F_λ be the distribution corresponding to $\Omega(r, \mu, \lambda)$.

Theorem 1. Let F be a distribution on a closed interval $[A, B]$ and let ϕ be the (unique) affine transformation from $[A, B]$ to $[-1, 1]$. Then there is a unique value of $\lambda \geq 1$ such that $\phi^* F_\lambda$ has the same mean and om as F.

Corollary.For every empirical distribution given by a data sample $\{X_{min} \ldots X_{max}\}$ there is a unique affine transformation ϕ from $[X_{min}, X_{max}]$ to $[-1, 1]$ and a unique $\Omega(r, \mu, \lambda)$ in the family given by equation (32) so that $\phi^* F_\lambda$ has the sample mean μ_S and om ω_S.

Omega Analysis

An Affine Invariant 'Shape' Score

For the Omega Functions given by equation (32) we can define $Shape(\mu, \lambda)$ by

$$Shape(\mu, \lambda) = \frac{1}{2}\int_{-1}^{1} log\Omega(r, \mu, \lambda)dr = \lambda log(\frac{1-\mu}{1+\mu}). \quad (33)$$

This is zero for symmetric distributions, is positive when $\mu < 0$ (so the distribution is biased to the upside) and negative when $\mu > 0$ (so the distribution is biased to the downside).

Omega Analysis

Theorem 2. $Shape(\mu, \lambda)$ is an affine invariant.

Proof. This follows directly by computing the values of μ, ω as the images of the affine map from any interval $[A, B]$ to $[-1, 1]$ and the corresponding value of λ.

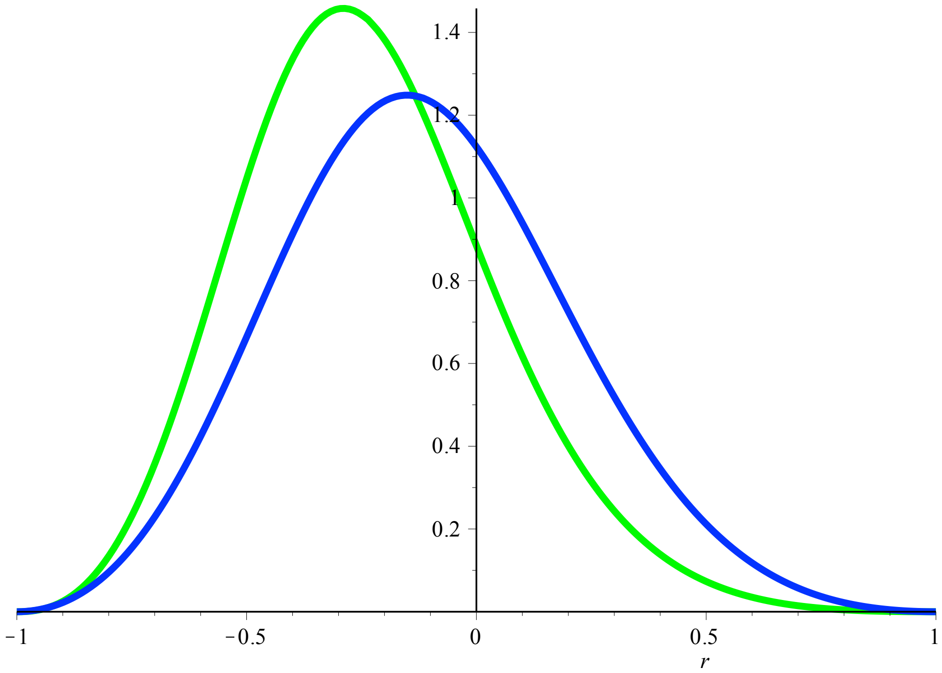
Corollary For a data sample $S = \{X_{min} \dots X_{max}\}$ with mean μ_S and om ω_S, $Shape_S$, given by

$$Shape_S = \frac{(\mu_S - x_{min})(x_{max} - \mu_S)}{\omega_S(x_{max} - x_{min})} log(\frac{x_{max} - \mu_S}{\mu_S - x_{min}}). \quad (34)$$

is an affine invariant.

Omega Analysis

Example: VIX (Green) and S&P 500 (blue) probability densities for returns data transformed to Omega Functions on $[-1, 1]$. Shape Scores VIX 3.2, S&P 500 0.9.



Data: Daily Returns, VIX, last 6 months, S&P500, 2008