Chebyshev's inequlity

John McCuan

December 11, 2023

Chebyshev's inequality asserts the following: If $\delta : \mathbb{R} \to [0, \infty)$ is an MDF (mass density function) for a probability measure, that is,

$$\int_{\mathbb{R}} \delta = 1, \tag{1}$$

with well-defined mean

$$\mu = \int_{\omega \in \mathbb{R}} \omega \delta(\omega)$$

and variance

$$\sigma^2 = \int_{\omega \in \mathbb{R}} (\omega - \mu)^2 \delta(\omega),$$

then for $k \geq 1$,

$$\int_{\omega \in [\mu - k\sigma, \mu + k\sigma]} \delta(\omega) \geq 1 - \frac{1}{k^2}.$$

1 Special case; strict inequality

Consider first the case δ is Riemann integrable, the mean is $\mu = 0$, and $k = 1$ so the basic inequality becomes

$$\int_{-\sigma}^{\sigma} \delta(\omega)\, d\omega \geq 0.$$

Of course this is vacuously true. In fact, for any fixed MDF δ in this class we must have strict inequality

$$\int_{-\sigma}^{\sigma} \delta(\omega)\, d\omega > 0. \tag{2}$$

To see this first note that $\sigma > 0$ since otherwise

$$\sigma^2 = \int_{-\infty}^{\infty} \omega^2 \, \delta(\omega) \, d\omega = 0,$$

which means the non-negative integrand $\omega^2 \, \delta(\omega)$ and hence the function $\delta(\omega)$ is zero almost everywhere resulting in the contradiction

$$1 = \int_{-\infty}^{\infty} \delta(\omega) \, d\omega = 1.$$

Now, let us call the quantity of interest in (2) the **central mass** M and assume by way of contradiction

$$M = \int_{-\sigma}^{\sigma} \delta(\omega) \, d\omega = 0.$$

It follows from this assumption that δ is zero almost everywhere on the interval $(-\sigma, \sigma)$. In particular,

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{-\sigma} \omega^2 \, \delta(\omega) \, d\omega + \int_{\sigma}^{\infty} \omega^2 \, \delta(\omega) \, d\omega \\ &< \sigma^2 \left(\int_{-\infty}^{-\sigma} \delta(\omega) \, d\omega + \int_{\sigma}^{\infty} \delta(\omega) \, d\omega \right) \qquad (3) \\ &= \sigma^2 \int_{-\infty}^{\infty} \delta(\omega) \, d\omega \\ &= \sigma^2 \qquad (4) \end{aligned}$$

which is a contradiction. The strict inequality in (3) deserves some comment. By virtue of the fact that under the assumption $\delta(\omega) = 0$ for $|\omega| < \delta$ there holds

$$1 = \int_{-\infty}^{-\infty} \delta(\omega) \, d\omega = \int_{-\infty}^{-\sigma} \delta(\omega) \, d\omega + \int_{\sigma}^{\infty} \delta(\omega) \, d\omega,$$

there is a point, a **Lebesgue point**, ω^* with $|\omega^*| > \sigma$ and

$$\delta(\omega^*) = \lim_{r \searrow 0} \frac{1}{2r} \int_{\omega^* - r}^{\omega^* + r} \delta(\omega) \, d\omega > 0. \qquad (5)$$

In fact the equality involving the limit in (5) holds at almost every point $\omega^* \in \mathbb{R}$ and at every Lebesgue point ω^*. This is a (somewhat difficult theorem from measure theory). If there is no such ω^* therefore, then one must have once again $\delta(\omega) = 0$ for almost every ω contradicting (1).

Notice that (5) gives right away

$$\int_{\omega^*-r}^{\omega^*+r} \omega^2 \delta(\omega)\, d\omega \geq (|\omega^*| - r)^2 \int_{\omega^*-r}^{\omega^*+r} \delta(\omega)\, d\omega > \sigma^2 \int_{\omega^*-r}^{\omega^*+r} \delta(\omega)\, d\omega \tag{6}$$

whenever $r < |\omega^*| - \sigma$ is small enough. In particular fixing $r > 0$ small enough with $r < |\omega^*| - \sigma$, this means

$$\begin{aligned}
&\int_{-\infty}^{-\sigma} \omega^2\, \delta(\omega)\, d\omega + \int_{\sigma}^{\infty} \omega^2\, \delta(\omega)\, d\omega \qquad (7)\\
&\quad = \int_{-\infty}^{-|\omega^*|-r} \omega^2\, \delta(\omega)\, d\omega + \int_{-|\omega^*|-r}^{-|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega + \int_{-|\omega^*|+r}^{-\sigma} \omega^2\, \delta(\omega)\, d\omega \\
&\qquad + \int_{\sigma}^{|\omega^*|-r} \omega^2\, \delta(\omega)\, d\omega + \int_{|\omega^*|-r}^{|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega + \int_{|\omega^*|-r}^{\infty} \omega^2\, \delta(\omega)\, d\omega \\
&\quad \geq \sigma^2 \int_{-\infty}^{-|\omega^*|-r} \delta(\omega)\, d\omega + \int_{-|\omega^*|-r}^{-|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega + \sigma^2 \int_{-|\omega^*|+r}^{-\sigma} \delta(\omega)\, d\omega \\
&\qquad + \sigma^2 \int_{\sigma}^{|\omega^*|-r} \delta(\omega)\, d\omega + \int_{|\omega^*|-r}^{|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega + \sigma^2 \int_{|\omega^*|-r}^{\infty} \delta(\omega)\, d\omega. \qquad (8)
\end{aligned}$$

If $\omega^* < -\sigma < 0$, then

$$\begin{aligned}
\int_{-|\omega^*|-r}^{-|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega &> \sigma^2 \int_{\omega^*-r}^{\omega^*+r} \delta(\omega)\, d\omega \\
&= \sigma^2 \int_{-|\omega^*|-r}^{-|\omega^*|+r} \delta(\omega)\, d\omega
\end{aligned}$$

by (6), and we can take

$$\int_{|\omega^*|-r}^{|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega \geq \sigma^2 \int_{|\omega^*|-r}^{|\omega^*|+r} \delta(\omega)\, d\omega.$$

Similarly, if $\omega^* > \sigma > 0$, then

$$\begin{aligned}
\int_{|\omega^*|-r}^{|\omega^*|+r} \omega^2\, \delta(\omega)\, d\omega &> \sigma^2 \int_{\omega^*-r}^{\omega^*+r} \delta(\omega)\, d\omega \\
&= \sigma^2 \int_{|\omega^*|-r}^{|\omega^*|+r} \delta(\omega)\, d\omega
\end{aligned}$$

by (6), and we can take

$$\int_{-|\omega^*|-r}^{-|\omega^*|+r} \omega^2\,\delta(\omega)\,d\omega \geq \sigma^2 \int_{-|\omega^*|-r}^{-|\omega^*|+r} \delta(\omega)\,d\omega.$$

Either way, picking up from (7) and (8) we obtain

$$\begin{aligned}
&\int_{-\infty}^{-\sigma} \omega^2\,\delta(\omega)\,d\omega + \int_{\sigma}^{\infty} \omega^2\,\delta(\omega)\,d\omega \\
&\qquad > \sigma^2 \int_{-\infty}^{-|\omega^*|-r} \delta(\omega)\,d\omega + \sigma^2 \int_{-|\omega^*|-r}^{-|\omega^*|+r} \delta(\omega)\,d\omega + \sigma^2 \int_{-|\omega^*|+r}^{-\sigma} \delta(\omega)\,d\omega \\
&\qquad\quad + \sigma^2 \int_{\sigma}^{|\omega^*|-r} \delta(\omega)\,d\omega + \sigma^2 \int_{|\omega^*|-r}^{|\omega^*|+r} \delta(\omega)\,d\omega + \sigma^2 \int_{|\omega^*|-r}^{\infty} \delta(\omega)\,d\omega. \\
&\qquad = \sigma^2 \left(\int_{-\infty}^{-\sigma} \delta(\omega)\,d\omega + \int_{\sigma}^{\infty} \delta(\omega)\,d\omega \right)
\end{aligned}$$

as claimed/used in (3).

To summarize, Chebyshev's inequality in the case $k = 1$ says that for a Riemann integrable MDF $\delta : \mathbb{R} \to [0, \infty)$ with (well-defined) mean

$$\mu = \int_{-\infty}^{\infty} \omega\,\delta(\omega)\,d\omega = 0$$

and well-defined variance

$$\sigma^2 = \int_{-\infty}^{\infty} \omega^2\,\delta(\omega)\,d\omega,$$

there holds

$$\int_{-\sigma}^{\sigma} \delta(\omega)\,d\omega \geq 0, \tag{9}$$

and we have obtained the following assertion:

Theorem 1 *Under the assumptions on the MDF just given, there holds $\sigma > 0$, and*

$$\int_{-\sigma}^{\sigma} \delta(\omega)\,d\omega > 0.$$

This means the equality in (9) cannot be attained. Thus, it is natural to ask the question: How close can one come to attaining the equality in (9)? The answer, it turns out, is that we can come as close as we like.

2 Initial examples

Theorem 2 *Given $\epsilon > 0$, there exists an MDF $\delta : \mathbb{R} \to [0, \infty)$ having the following properties:*

(i) *δ is Riemann integrable with*

$$\int_{-\infty}^{\infty} \delta(\omega)\, d\omega = 1,$$

(ii) *δ is even, i.e., $\delta(-\omega) = \delta(\omega)$ for $\omega \in \mathbb{R}$ and the mean*

$$\mu = \int_{-\infty}^{\infty} \omega\, \delta(\omega)\, d\omega$$

is well-defined and therefore zero,

(iii) *δ has well-defined finite variance $\sigma > 0$, and*

for which

$$\int_{-\sigma}^{\sigma} \delta(\omega)\, d\omega < \epsilon.$$

It is of some interest to find an MDF having the properties of Theorem 2 with form as simple as possible. The simplest family of MDFs I have found with these properties are given by

$$\delta(\omega) = \chi_{[-\epsilon/2,\epsilon/2]}(\omega) + \frac{1-\epsilon}{\epsilon}\, \chi_{[-1-\epsilon/2,-1]}(\omega) + \frac{1-\epsilon}{\epsilon}\, \chi_{[1,1+\epsilon/2]}(\omega) \tag{10}$$

for $0 < \epsilon \le 1/2$ as indicated on the left in Figure 1.

A slight variation of of the MDF given in (10) is given by

$$\delta(\omega) = \chi_{[1-\epsilon/2,1]}(\omega) + \chi_{[1,1+\epsilon/2]}(\omega) + \frac{1-\epsilon}{\epsilon}\, \chi_{[-2,-2-\epsilon/2]}(\omega) + \frac{1-\epsilon}{\epsilon}\, \chi_{[2,2+\epsilon/2]}(\omega)$$

as indicated on the right in Figure 1. The calcuations are slightly more complicated for this family of MDFs, but the result is essentially the same. These MDFs illustrate that some variety of measures having the properties required by Theorem 2 is possible. Both of these families, however, limit to atomic measures with atoms at $\pm a \in \mathbb{R}$ of measure $\pi(\{\pm a\}) = 1/2$, and the Wikipedia page on Chebyshev's inequality has a

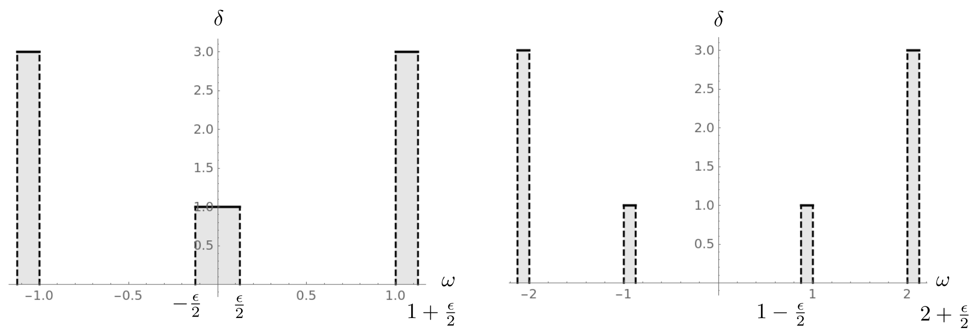

Figure 1: MDFs with small central mass.

claim that these atomic measures are essentially the unique probability measures for which equality is attained.[1]

It is relatively easy to check that the MDF $\delta : \mathbb{R} \to [0, \infty)$ defined in (10) is a probability measure. We will give the computation of the variance and the central mass corresponding to one standard deviation.

$$\begin{aligned}
\sigma^2 &= 2\int_0^{\epsilon/2} \omega^2 \, d\omega + 2\int_1^{1+\epsilon/2} \omega^2 \, \frac{1-\epsilon}{\epsilon} \, d\omega \\
&= \frac{\epsilon^3}{12} + \frac{2}{3} \, \frac{1-\epsilon}{\epsilon} \, \left[\left(1+\frac{\epsilon}{2}\right)^3 - 1 \right] \\
&= \frac{\epsilon^3}{12} + \frac{2}{3} \, \frac{1-\epsilon}{\epsilon} \, \left[\frac{3\epsilon}{2} + \frac{3\epsilon^2}{4} + \frac{\epsilon^3}{8} \right] \\
&= \frac{\epsilon^3}{12} + \frac{1-\epsilon}{12} \, \left[12 + 6\epsilon + \epsilon^2 \right] \\
&= 1 + \frac{1}{12} \left[6\epsilon + \epsilon^2 \right] - \frac{1}{12} \, \left[12\epsilon + 6\epsilon^2 \right] \\
&= 1 - \frac{1}{12} \left(6\epsilon + 5\epsilon^2 \right).
\end{aligned}$$

Thus, it is obvious that $\sigma^2 < 1$ and $\sigma < 1$. It is also very easily checked that $\sigma^2 > \epsilon^2/4$. This is equivalent to $(\epsilon + 3)(\epsilon - 2) < 0$. Consequently,

$$\frac{\epsilon}{2} < \sigma < 1,$$

[1]A more general class of measures including atomic measures is of course included in the exposition on the Wikipedia page and consequently a more general form of Chebyshev's inequality.

so the central mass is simply $M = \epsilon$, and this can be made arbitrarily small by making ϵ small.

3 Other examples

Of course, another example much simpler than the one above is the centrally symmetric uniform MDF $\delta = \chi_{[-a,a]}/(2a)$ illustrated in the middle plot in Figure 2. For this MDF

$$\sigma^2 = \frac{1}{a}\frac{a^3}{3} = \frac{a^2}{3}$$

and

$$M = \frac{1}{\sqrt{3}}.$$

Next, we consider more genrally the piecewise affine MDFs indicated in Figure 2. These are given by

$$\delta(\omega) = \left(\frac{1}{2a}(1 - ma^2) + m|\omega|\right) \chi_{[-a,a]}(\omega)$$

for $-1/a^2 \leq m \leq 1/a^2$. The variance is

$$\sigma^2 = a^2 \left(\frac{1}{3} + \frac{ma^2}{6}\right)$$

which it may be observed increases with m, so that

$$\sigma\left(-\frac{1}{a^2}\right) = \frac{a}{\sqrt{6}} < \sigma(0) = \frac{a}{\sqrt{3}} < \sigma\left(\frac{1}{a^2}\right) = \frac{a}{\sqrt{2}}$$

with

$$\sigma(m_1) < \sigma(m_2) \quad \text{for } -\frac{1}{a^2} \leq m_1 < m_2 \leq \frac{1}{a^2}.$$

The central mass on the other hand, also indicated in Figure 2 for various values of m, is found to decrease with m, so that among these piecewise affine MDFs with maximum at $\omega = 0$, there holds

$$M(m) \geq M(0) = \frac{1}{\sqrt{3}} \qquad \text{for } m \leq 0.$$

This perhaps suggests the following conjecture.

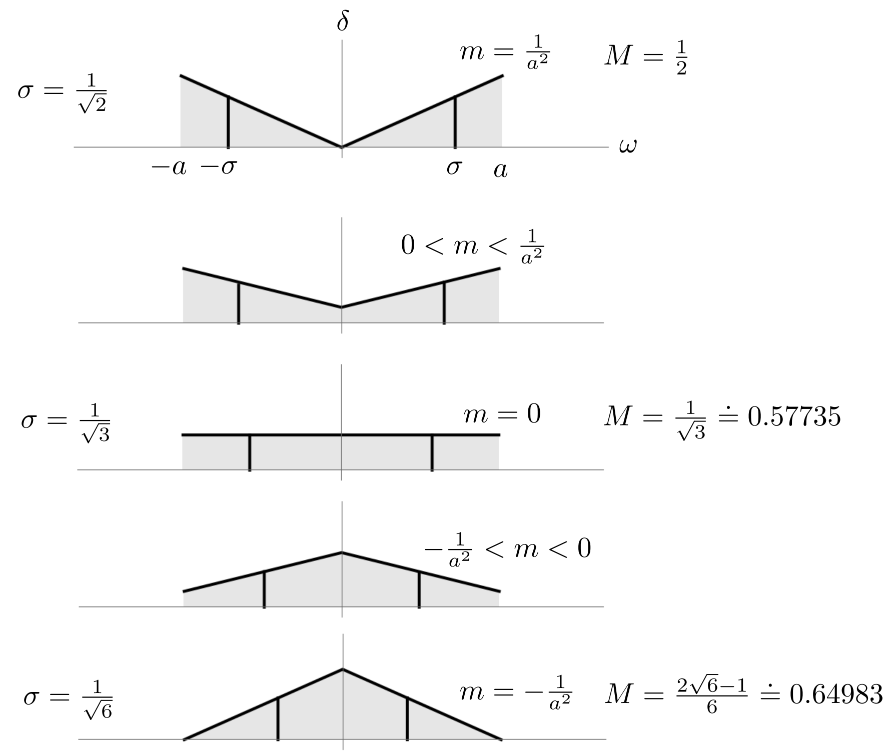

Figure 2: Piecewise affine MDFs.

4 Conjecture

Conjecture 1 *There exists some positive number $M_0 > 0$ for which the following holds: If $\delta : \mathbb{R} \to [0, \infty)$ is an MDF satisfying properties* **(i)**, **(ii)**, and **(iii)** *of Theorem 2 above and the additional condition that*

$$\delta(\omega_1) \geq \delta(\omega_2) \qquad \text{for} \qquad 0 \leq \omega_1 \leq \omega_2,$$

then

$$\int_{-\sigma}^{\sigma} \delta(\omega)\, d\omega \geq M_0.$$

We could add the assertion that $M_0 = 1/\sqrt{3}$, but this is likely incorrect, though I do not know a counterexample at the moment.

Let us say an MDF having the properties given as hypotheses in Conjecture 1 is **symmetric-monotone**.

I will first carry out the calculation of the central mass associated with the piecewise affine MDFs given above and verify the monotonicity leading to Conjecture 1. Then I will consider some other examples.

As mentioned above for the piecewise affine MDFs we have

$$\sigma = \frac{a}{\sqrt{6}}\sqrt{2+ma^2} \qquad \text{for} \qquad -\frac{1}{a^2} \le m \le \frac{1}{a^2}.$$

Thus, the central mass is given by

$$\begin{aligned} M &= 2\int_0^\sigma \left(\frac{1}{2a}(1-ma^2) + m\omega\right)\, d\omega \\ &= \frac{1-ma^2}{a}\ \sigma + m\sigma^2 \\ &= \frac{1}{\sqrt{6}}(1-ma^2)\ \sqrt{2+ma^2} + \frac{ma^2}{6}\ (2+ma^2). \end{aligned}$$

Notice we may also write

$$M = -\frac{1}{\sqrt{6}}(2+ma^2)^{3/2} + \frac{3}{\sqrt{6}}\ \sqrt{2+ma^2} + \frac{a^2}{3}\ m + \frac{a^4}{6}\ m^2$$

so that

$$\begin{aligned} \frac{dM}{dm} &= -\frac{3}{2\sqrt{6}}(2+ma^2)^{1/2}\ a^2 + \frac{3}{2\sqrt{6}}\ \frac{a^2}{\sqrt{2+ma^2}} + \frac{a^2}{3} + \frac{a^4}{3}\ m \\ &= \frac{3a^2}{2\sqrt{6}\ \sqrt{2+ma^2}}\ (-2-ma^2+1) + \frac{a^2}{3} + \frac{a^4}{3}\ m \\ &= \frac{a^2}{3}\left(-\frac{9}{2\sqrt{6}\ \sqrt{2+ma^2}}\ (1+ma^2) + 1 + a^2 m\right) \\ &= \frac{a^2}{3}\ A(m) \end{aligned}$$

where

$$\begin{aligned} A(m) &= -\frac{9}{2\sqrt{6}\ \sqrt{2+ma^2}}\ (1+ma^2) + 1 + a^2 m \\ &\le -\frac{3}{2\sqrt{2}}(1+ma^2) + 1 + a^2 m \\ &= a^2\left(1 - \frac{3}{2\sqrt{2}}\right)\ m + 1 - \frac{3}{2\sqrt{2}}. \end{aligned}$$

Setting

$$\phi(m) = a^2 \left(1 - \frac{3}{2\sqrt{2}}\right)\ m + 1 - \frac{3}{2\sqrt{2}},$$

we see

$$\phi'(m) = a^2 \left(1 - \frac{3}{2\sqrt{2}}\right) < 0.$$

It follows that $A(m) \le \phi(m) \le \phi(-1/a^2) = 0$ and

$$\frac{dM}{dm} \le 0$$

with equality only for $m = -1/a^2$. This implies

$$M(m_2) < M(m_1) \qquad \text{for} \qquad -\frac{1}{a^2} \le m_1 < m_2 \le \frac{1}{a^2}.$$

In particular,

$$\frac{1}{\sqrt{6}} = M\left(\frac{1}{a^2}\right) < \frac{1}{\sqrt{3}} = M(0) < M(m) \le M\left(-\frac{1}{a^2}\right) = \frac{2}{\sqrt{6}} - \frac{1}{6} = \frac{2\sqrt{6}-1}{6}$$

for

$$-\frac{1}{a^2} \le m < 0$$

as claimed.

Exercise 1 *Use mathematical software to plot the, variance, standard deviation, and central mass corresponding to one standard deviation for the affine MDFs above as a function of m for*

$$-\frac{1}{a^2} \le m \le \frac{1}{a^2}.$$

Do the functions σ^2 and M have any interesting propeties as functions of m?

5 Other examples

Another obvious example to try would be the Gaussian or normal MDF given by

$$\delta(\omega) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(\omega-\mu)^2}{2\sigma^2}}.$$

In this case one finds

$$M = \int_{\mu-\sigma}^{\mu+\sigma} \delta(\omega)\, d\omega \doteq 0.6826895 > \frac{1}{\sqrt{3}}.$$

This is consistent with Conjecture 1.

Exercise 2 *Consider the MDF obtained as follows:*

(a) *For $a > 0$, let $f : \mathbb{R} \to [0, \infty)$ by*

$$f(\omega) = \begin{cases} 1/a^2, & |\omega| \le a \\ 1/\omega^2, & |\omega| \ge a. \end{cases}$$

Find an appropriate constant $c > 0$ so that $\delta(\omega) = c\, f(\omega)$ is a probability MDF.

(b) *What happens if you try to test the assertion of Conjecture 1 using the MDF from part* **(a)** *above?*

For $a > 0$, consider $g : \mathbb{R} \to [0, \infty)$ by

$$g(\omega) = \begin{cases} e^{-a^2}, & \omega^2 \le a \\ e^{-\omega^2}, & |\omega| \ge a \end{cases}$$

as indicated in Figure 3. Notice g is $\sqrt{\pi}\ \delta$ where δ is the normal MDF given above

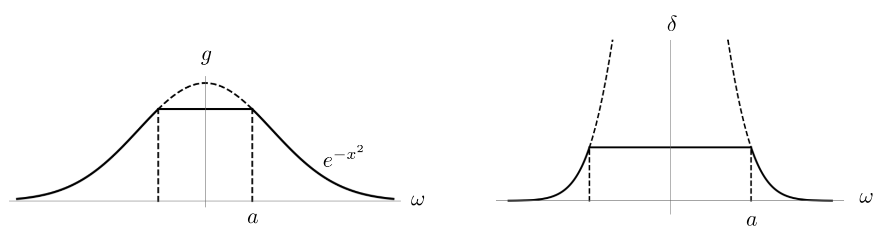

Figure 3: Exponentially decaying (Gaussian) MDFs.

with mean $\mu = 0$ and standard deviation $\sigma = 1/\sqrt{2}$. Thus, one may take c with

$$\frac{1}{c} = \int_{-\infty}^{\infty} g(\omega)\, d\omega$$

and then $\delta(\omega) = c\ g(\omega)$ defines a probability MDF. The function g with $a = 1/2$ is plotted on the left in Figure 3, and the scaled function $\delta = cg$ with $a = 1.5$ is plotted on the right in Figure 3. In this way, we obtain a one parameter family of MDFs with $\delta = c(a)\ g$ for $a > 0$. The value of $c = c(a)$ will, in principle, need to be calcuated numerically for each $a > 0$, but in practice, most mathematical software packages include a "standard" **Gaussian error function** which may be used to compute values of $c = c(a)$ as well as $\sigma = \sigma(a)$ and $M = M(a)$. If I have calculated correctly, then

$$\sigma(1.5) \doteq 0.639684 \qquad \text{and} \qquad M(1.5) \doteq 0.358367$$

which is already significantly lower than the conjectured value $M_0 = 1/\sqrt{3}$.

Exercise 3 *Check the value of $M(1.5)$ reported above, and plot the values of $\sigma(a)$ and $M(a)$ for $a > 0$ in general using the truncated Gaussian MDFs described above. What is the conclusion concerning Conjecture 1, i.e., what is the lowest value of M attained among these examples?*

If Conjecture 1 still survives Exercise 3 above, then another direction that may be of interest is the following: Let $\delta : [-a, a] \to [0, \infty)$ satisfy the conditions of the conjecture with

$$2 \int_0^a \delta(\omega)\, d\omega = 1.$$

Consider the family of MDFs $\tau : [-a, a] \to [0, \infty)$ given by

$$\tau(\omega) = (1 - t)\delta(\omega) + \frac{t}{2a}. \tag{11}$$

Exercise 4 *Show τ is a symmetric-monotone probability MDF for each fixed value of t with $0 \le t \le 1$.*

Conjecture 2 *If $M = M(\tau)$ denotes the central mass associated with the convex combination τ of MDFs given in* (11)*, then*

$$\frac{dM}{dt} \le 0 \qquad \text{for} \qquad 0 < t < 1.$$

If Conjecture 2 holds, then this establishes Conjecture 1 at least for symmetric-monotone MDFs with statistical range $[-a, a]$ and no (exponential) tails.

Final comment/project: The only other approach I considered on this topic was to give a direct proof of Conjecture 1 by going through the standard proof of the Chebyshev/Markov inequality and attempting to sharpen the estimates using the symmetric-monotone condition. It seemed like there might have been some way forward with this, but I certainly didn't see a way to prove the conjecture.