CORPUS PUBLISHERS

Advance Research in Sciences (ARS)

ISSN: 2837-5777

Volume 2, Issue 2, 2024

Article Information

Received date : August 28, 2024

Published date: September 16, 2024

*Corresponding author

Eugene Pinsky, Department of Computer Science, Metropolitan College, Boston University, 1010 Commonwealth Avenue, Boston, 02215, MA, USA

Keywords: Eugene Pinsky, Department of Computer Science, Metropolitan College, Boston University, 1010 Commonwealth Avenue, Boston, 02215, MA, USA

DOI: 10.54026/ARS/1021

Research Article

Mean Absolute Deviation (About Mean) Metric for Kurtosis

Eugene Pinsky*

Department of Computer Science, Metropolitan College, Boston University, 1010 Commonwealth Avenue, Boston, 02215, MA, USA

Abstract

In this paper, we introduce a MAD (about mean) alternative metric to classical kurtosis. Using the formula for mean absolute deviation, we identify two special points and construct the corresponding distributions with these reference points as their means. The proposed alternative is computed from mean absolute deviations for these distributions. These measures are bounded, require only the existence of first-order moments, and are easy to interpret. We illustrate the computation of these measures for several well-known distributions.

Preliminaries

We start with some preliminaries. Consider a real-valued random variable X on a sample space $\Omega \subseteq R$ with density f(x) and cumulative distribution function F (x). We require the existence of the first moment only and denote the mean of X by $\mu = E(X)$. If X is a discrete random variable, then Ω is some countable sample space, and f (x) is the probability mass function (discrete density function). Let F^{-1} (.) be the quantile function defined by $F^{-1}(t) = \inf\{x: F(x) \geq t\}$ with $t \in [0, 1]$. In particular, the median $M = F^{-1}(1/2)$ and the quartiles are $Q_1 = F^{-1}(1/4)$ and $Q_3 = F^{-1}(3/4)$.

We will find it convenient to use the indicator function: for any subset $U \subseteq \Omega$ define

$$1_U = \begin{cases} 1 & if\ x \in U \\ 0 & otherwise \end{cases}$$

Define the left sub-space of Ω by $\Omega_L = \{w \in \Omega \mid x(w) \leq \mu\}$ and right sub- space of Ω by $\Omega_R = \{w \in \Omega \mid x(w) > \mu\}$. Also, define random variables $X_L = X1_{\Omega L}$ and $XR = X1_{\Omega R}$. Finally, define the following auxiliary integral

$$I(Z) = \int_{x \leq z} x f(x) dx \qquad (1)$$

We can consider I(z) as a partial mean of X computed over all $x \leq z$. The mean absolute deviation (MAD) of X about μ is

$$H = \int_{\Omega} |x-\mu| f(x) dx = 2\mu F(\mu) - 2I(\mu) \qquad (2)$$

Defining MAD (about Mean) Alternative to Kurtosis

In this paper, we focus on introducing a MAD-based measure of kurtosis. In probability theory, the classical Pearson's kurtosis K is defined as Johnson & Kotz [1]: $K = E(X-\mu)^4/\sigma^4$. A very useful interpretation of kurtosis has been suggested in Moors [2]. Consider a standardized variable $Z = (X-\mu)/\sigma$ and let $W = Z^2$. Then E (W) = 1 and Var (W) = $E(W^2) - E^2(W) = K-1$ and therefore K = Var (W) + 1. The kurtosis K can be viewed as related to the dispersion of W = Z2 around its mean 1. Equivalently, K is associated with the dispersion of Z around -1 and 1 [2]. This implies that kurtosis is associated with the concentration of X around points $(\mu - \sigma)$ and $(\mu + \sigma)$ respectively. With this analogy, an quantile alternative T_Q to kurtosis was suggested in Moors [3] in terms of the inter-quartile ranges X_L and X_R around corresponding medians Q_1 and Q_3 using octiles $O_i = F^{-1}(i/8)$. In Pinsky & Klawansky [4], the analogy to T_Q was pursued to define MAD-based (around median) alternative T_M to kurtosis by measuring the mean absolute deviations of X_L and X_R from their corresponding medians Q_1 and Q_3, playing the role of $(\mu \pm \sigma)$ as in classical statistics. The corresponding formulas for T_Q and T_M are

$$T_Q = \frac{(O_3 - O1) + (O_7 - O_5)}{Q_3 - Q_1}, T_{M=} \frac{E(|X-Q_1| 1_{X \leq M}) + E(|X-Q_3| 1_{X \geq M})}{E(|X-M|)} \qquad (3)$$

Classical kurtosis K, quantile alternative T_Q and MAD (about median) alternative T_M can all be interpreted as measures of dispersion around the special "reference" points like $(\mu \pm \sigma)$ or quartiles Q_1 and Q_3. Note that in the case of MAD about median, in general, we have $E(|X - M| 1_{x \leq M}) \neq E(|X-M|\ 1x>M)$ and we can use this as measurements for tails as suggested in Habib [5]. However, in the case of a mean, for any distribution, we have $E(|X-\mu|\ 1_{\Omega L}) = E(X - \mu\ 1_{\Omega R})$, requiring us to use some other reference point(s) rather than μ. But what should be the corresponding reference points if we define an alternative to kurtosis using mean absolute deviation around the mean? Moreover, we would like to define such points using only the first moments. We suggest defining the left and right derived distributions for XL and X_R directly from MAD (about mean) of X and using their means μ_L and μ_R as the corresponding reference points. We proceed as follows: From equation (2), we can write $H = H(X, \mu)$ as follows:

$$H = F(\mu)\left(\mu - \frac{I(\mu)}{F(\mu)} + (1-F(\mu))\left(\frac{\mu - I(\mu)}{1-F(\mu)} - \mu\right)\right) \quad (4)$$

To interpret equation (4), consider the "left" and "right" functions defined for X_L and X_R:

$$\begin{cases} f_L(x) = f(x)/F(\mu) \\ F_L(x) = F(x)/F(\mu) \\ I_L(Z) = I(Z)/F(\mu) \end{cases} \quad and \quad \begin{cases} f_R(x) = f(x)/(1-F(\mu)), \\ F_R(x) = (F(x)/F(\mu))/(1-F(\mu)) \\ I_R(Z) = (I(Z)/I(\mu))/(1-F(\mu)) \end{cases} \quad (5)$$

It is easy to verify that $f_L(\cdot)$ and $F_R(\cdot)$ define probability densities for X_L and X_R with the corresponding cumulative distribution functions $F_L(\cdot)$ and $F_R(\cdot)$ respectively. The functions $I_L(\cdot)$ and $I_R(\cdot)$ are the corresponding auxiliary integrals. The corresponding means μ_L and μ_R for the corresponding distributions are given by $\mu_L = I(\mu)/F(\mu)$ and $\mu_R = (\mu - I(\mu))/(1 - F(\mu))$. Therefore, we can express H in equation (4) as

$$H = F(\mu)(\mu - \mu_L) + (1 - F(\mu))(\mu_R - \mu) \quad (6)$$

The mean absolute deviation from the mean H is the weighted average of the distances of the left and right means μL and μR from the mean μ taken with weights F (μ) and (1 − F (μ)) respectively. Once these reference points μ_L an μ_R are defined, we suggest to define MAD-based (around mean) kurtosis to measure the concentration of X around μ_L (playing the role of μ-σ) and around μ_R (playing the role of μ + σ) as in classical statistics. Specifically, we suggest to define MAD-based kurtosis T_μ by computing the total of mean absolute deviations of points in ΩL and ΩR around these reference points μ_L and μ_R respectively, and then normalizing this by total H:

$$T_\mu = \frac{1}{H}\left[\int_{\Omega L} |x - \mu_L| f(x)dx + \int_{\Omega R} |x - \mu_R| f(x)dx\right] \quad (7)$$

The above expression for T_μ can be interpreted as the ratio of two distances: the numerator is the average of distances from values in Ω_L and Ω_R from μ_L and μ_R respectively, whereas the denominator H is the average distance of values of Ω to its mean μ. To provide further interpretation for T_μ, let H_L denote the mean absolute deviation of X_L from its mean μ_L and let H_R denote the mean absolute deviation of X_R from its mean μ_R:

$$H_L = \int_{\Omega L} |x - \mu_L| f_L(x)\, dx \quad and \quad H_R = \int_{\Omega R} |x - \mu_R| f_R(x)\, dx$$

Then using equation (5) we can rewrite T_μ in equation (7) as

$$T_\mu = \frac{F(\mu)H_L + (1-F(\mu))H_R}{H} \quad (8)$$

We can interpret MAD-based kurtosis from equation (8) as the following ratio: the numerator is the weighted sum of mean absolute deviations of random variables X_L and X_R from their respective means μ_L and μ_R. The denominator H is the average absolute distance of X from its mean μ. It is more difficult to provide an intuitive explanation of the classical kurtosis K. The computation of MAD (around mean) kurtosis T_μ can be summarized in the following steps:

compute MAD H for $X : H = 2_\mu F(\mu) - 2I(\mu)$

compute MAD H_L for $X_L : H = 2_{\mu_L} F_L(\mu_L) - 2I_L(\mu_L)$

compute MAD H_R for) $X_R : H_R = 2_{\mu R} F_R(\mu_R) - 2I_R(\mu_R)$

compute MAD-based kurtosis $T_\mu = (F(\mu)H_L + (1-F(\mu))H_R)/H$

It can be shown that for some distributions such as log-normal, both classical kurtosis K and quantile kurtosis T_Q can be unbounded [4]. By contrast, T_μ is bounded and satisfies 0 ≤ Tμ ≤ 2. This is easily established by applying Cauchy-Schwartz inequality as follows:

$$T_\mu = \frac{1}{H}\left[\int_{\Omega L} |(x-\mu) + (\mu - \mu_L)| f(x)dx + \int_{\Omega R} |(\mu - x) + (\mu_R - \mu)| f(x)\right]$$

$$\leq \frac{1}{H}\left[\int_{\Omega} |x - \mu| f(x)dx\right] + \frac{(\mu - \mu_L)F(\mu) + (\mu_R - \mu)(1 - F(\mu))}{H} = 2$$

For MAD (around median) kurtosis T_M introduced in Pinsky & Klawansky [4], we have an even sharper bound $0 \leq T_M \leq 1$ Finally, we note that we only insist on the existence of the first-order moment to compute T_μ by contrast to classical kurtosis K that requires the existence of higher order moments.

Examples

We illustrate the computation of MAD-based kurtosis for the uniform, normal, exponential, and Laplace distributions. Since mean absolute deviations are shift-invariant and ratios are scale-invariant, we can consider standardized distributions.

Standard uniform distribution

Suppose X is distributed according to a uniform distribution in [0,1]. Its probability density function f (x) = 1 and its cumulative distribution function F (x) = x. For this distribution, mean μ = 1/2, σ² = 1/12 and kurtosis K = 1.8. The auxiliary integral I(z) is easily computed:

$$I(z) = \int_0^Z xf(x)dx = z^2/2$$

We now compute MAD-based kurtosis as follows:

$$\text{compute } H : \mu = 1/2,\ F(\mu) = 1/2,\ I(\mu) = 1/8$$
$$\Rightarrow H = 2\mu F(\mu) - 2I(\mu) = 1/4$$
$$\text{compute } H_L : \mu_L = 1/4,\ F_L(\mu_L) = 1/2,\ I_L(\mu) = 1/16$$
$$\Rightarrow H_L = 2\mu LFL(\mu_L) - 2I_L(\mu_L) = 1/8$$
$$\text{compute } H_R : \mu_R = 3/4,\ F_R(\mu_R) = 1/2,\ I_R(\mu_R) = 5/16$$
$$\Rightarrow H_R = 2\mu_R F_R(\mu_R) - 2I_R(\mu_R) = 1/8$$
$$\text{compute } T_\mu = (F(\mu)H_L + (1-F(\mu))H_R)/H = 1/2$$

For this distribution, the left and right means μ_L and μ_R coincide with the quartiles Q_1 = 1/4 and Q_3 = 3/4. The MAD (around median) kurtosis is $T_M = T_\mu$, whereas the quantile kurtosis is T_Q = 1 [2,4].

Standard normal distribution

Suppose X is distributed according to the standard normal distribution with density f (x) and cumulative distribution function F (x):

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad and \quad F(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

where Φ (.) denotes the cumulative distribution of the standard normal. The classical kurtosis for normal distribution K = 3.

We compute the auxiliary integral

$$I(z) = \int_{x \leq z} xf(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} xe^{-x^2/2} dx = -\frac{e^{-x^2/2}}{\sqrt{2\pi}}\Bigg|_{-\infty}^{z} = -\frac{e^{-z^2/2}}{\sqrt{2\pi}}$$

and now compute MAD-based kurtosis as follows:

$$\text{1.compute } H : \mu = 0,\ F(\mu) = 1/2,\ I(\mu) = -1/\sqrt{2\pi}$$
$$\Rightarrow H(\mu) = 2\mu F(\mu) - 2I(\mu) = \sqrt{2/\pi}$$
$$\text{2. compute } H_L : \mu_L = -\sqrt{2/\pi},\ F_L(\mu_L) = 2(1 - \phi(\sqrt{2/\pi})),$$
$$I_L(\mu_L) = -(\sqrt{2/\pi})\ e^{-1/\pi}$$
$$\Rightarrow H_L = 2\mu_L F_L(\mu_L) - 2I_L(\mu_L) = 2\sqrt{2/\pi}(e^{-1/\pi} - 2 + 2\phi(\sqrt{2/\pi})),$$
$$\text{3.compute } H_R : \mu_R = \sqrt{2/\pi},\ F_R(\mu_R) = 2\phi(\sqrt{2/\pi}) - 1$$
$$I_R(\mu_R) = \sqrt{2/\pi}(1 - e^{-1/\pi})$$
$$\Rightarrow H_R = 2\mu_R F_R(\mu_R) - 2I_R(\mu_R) = 2\sqrt{2/\pi}(e^{-1/\pi} - 2 + 2\phi(\sqrt{2/\pi}))$$

$$\text{4. compute } T_\mu = \frac{F(\mu)H_L + (1-F(\mu))H_R}{H} = 2(e^{-1/\pi} - 2 + 2\phi(\sqrt{2/\pi})) \approx 0.60$$

The quantile kurtosis T_Q≈1.23 and the MAD (around median) TM≈0.59 [2,4]. For this distribution, the mad (about mean) and mad (about median) metrics for kurtosis are very close. For this distribution, the means μ_L, μ_R = ± 2π = ± 0.80 do not coincide with quartiles and are further away from μ = 0 than the quartiles. This results in larger values for mean absolute deviations for X_L and X_R, whereas the denominator is the same for X since μ = M. Note that for this distribution, we could compute T_μ differently: the random "right" variable X_{RL} follows half-normal distribution, and its mean absolute deviation H_R is known [6]. The "left" variable X_L ~−X_R and by symmetry has the same MAD H_R = H_L. The mean absolute deviation for normal distribution is well-known [6]. This gives us an alternative approach to compute T_μ.

Exponential distribution

Suppose X is distributed according to an exponential distribution with rate λ > 0 [7]. Its density f (x) and its cumulative distribution function F (x) are given by

$$f(x) = \lambda e^{-\lambda x} \quad and\, F(x) = 1 - e^{-\lambda x}, \qquad x \in [0, \infty)$$

We compute the auxiliary integral (integral #2.322 in Gradshteyn & Ryzhik [8])

$$I(z) = \int_0^z x\lambda e^{-\lambda x} dx = -e^{-\lambda x}(x + 1/\lambda)|_o^z = 1/\lambda - e^{-\lambda z}(z + 1/\lambda)$$

We can now compute MAD-based kurtosis T_μ

$$\text{compute } H : H : \; \mu = 1/\lambda, F(\mu) = (e-1)/e,\; I(\mu) = (e-2)/\lambda e$$
$$\Rightarrow H = 2\mu F(\mu) - 2I(\mu) = 2/\lambda e$$
$$\text{compute HL: } H_L : \mu_L = (e-2)/\lambda(e-1), \quad F_L(\mu_L) = (e - e^{1/(e-1)})/(e-1),$$
$$I_L(\mu_L) = (e^2 - (1 + 2e^{1/(e-1)})e + 3e^{1/(e-1)})/\lambda(e-1)^2,$$
$$H_L = 2\mu_L F_L(\mu_L) - 2I_L(\mu_L) = 2(e^{1/(e-1)}(e-1) - e)/\lambda(e-1)^2$$
$$\text{compute } H_R : \mu_R = 2/\lambda, \quad F_R(\mu_R) = 1 - 1/e,\; I_R(\mu_R) = (2e-3)/\lambda e$$
$$\Rightarrow H_R = 2\mu_R F_R(\mu_R) - 2I_R(\mu_R) = 2/\lambda e$$
$$\text{compute } T_\mu = \frac{F(\mu)H_L + (1 - F(\mu))H_R}{H} = e^{1/(e-1)} + 1/e - e/(e-1) \approx 0.58$$

The means μ_L ≈ 0.42/λ and μ_R = 2/λ are not equal to quartiles Q_1 = log(4/3)/λ ≈ 0.29/λ and Q_3 = log(4)/λ ≈ 1.39 [1]. The quantile kurtosis T_Q ≈ 1.31 and MAD (about median) kurtosis T_M ≈ 0.62 [2,4].

$$f(x) = \frac{1}{2b} e^{-|x|/b} \qquad and \qquad F(x) = \left[\frac{1}{2} + \frac{1}{2}\text{sgn}(x)\right] e^{|x|/b}$$

Laplace Distribution

Suppose X is distributed according to Laplace (double exponential) distribution with location μ = 0 and scale b = 1/λ. Its density f (x) and cumulative distribution function F (x) are given by Feller, et al. [1,7]

For this distribution, we can easily compute T_μ as follows. The MAD H for Laplace is well-known and is H = b = 1/λ [6]. The "right" random variable X_R follows an exponential distribution with rate λ. Its MAD is H_R = 2/λe. By symmetry, for "left" variable X_L, we have the same value for H_L = 2/λe. Since F (μ) = 1 - F (μ) = 1/2, for this distribution, we immediately compute the MAD metric for kurtosis Tμ = 2/e ≈ 0.80. Note that this is larger than Tμ ≈ 0.60 for normal distribution, reflecting that tails are fatter for the Laplace than for normal distribution.

Conclusion

In this paper, we introduced a MAD (about mean) alternative metric to kurtosis. The new formula requires the computation of two special points and the construction of the corresponding distributions with these reference points as their means. The proposed alternative metric is computed by con- structing two derived distributions and computing their mean absolute deviations around their corresponding means. The new metric is simple to compute and requires the existence of the first-order moment. It is bounded and simple to interpret. Future work will focus on defining other MAD-based alternatives to other classical probability measures.

References

1. Johnson N, Kotz S (1970) Distributions in Statistics. J Wiley, New York, USA.
2. Moors J (1988) A quantile alternative to kurtosis. Journal of the Royal Statistical Society 37: 25-32.
3. Moors J (1986) The meaning of kurtosis: Darlington reexamined. The American Statistician 40: 283-284.
4. Pinsky E, Klawansky S (2023) Mad (about median) vs. quantile- based alternatives for classical standard deviation skew, and kurtosis. Frontiers in Applied Mathematics and Statistics 9: 1-14.
5. Habib E (2012) Mean absolute deviation about median as a tool of exploratory data analysis. International Journal of Research and Reviews in Applied Sciences 11: 517-523.
6. Weisstein E (2006) Mean deviation.
7. Feller J (1956) Probability theory and applications. J Wiley, New York, USA.
8. Gradshteyn IS, Ryzhik IM (1980) Table of integrals, series, and products. Academic Press, New York, USA.