Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

On Functional Deviations About Functionals

STHITADHI DAS

sthdas999@gmail.com

Brainware University https://orcid.org/0009-0001-2484-9530

Research Article

Keywords: Statistical functional, functional deviation, mean absolute deviation, median absolute deviation, quantile deviation, dispersion inequalities, Monte Carlo simulation

Posted Date: October 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11471410/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

On Functional Deviations About Functionals

Sthitadhi Das*
Department of Mathematics
Brainware University, India

Abstract

A general framework for measuring dispersion through statistical functionals is developed. Given a random variable and a reference functional, the deviation is defined by applying a statistical functional to the absolute deviations from the reference value. The framework includes mean, median, quantile, and other functional-based measures of dispersion. A two-functional extension distinguishes the functional determining the reference location from the functional summarizing the deviations. Several inequalities are established, including the comparison of mean absolute deviations about the mean and median, an upper bound involving the standard deviation, and quantile bounds for nonnegative deviations. A simulation study based on normal, uniform, exponential, Student's t, log-normal, and Gaussian-mixture distributions illustrates these relationships. The numerical results demonstrate the effect of distributional shape and reference-functional selection on the resulting dispersion measures.

Keywords: Statistical functional; functional deviation; mean absolute deviation; median absolute deviation; quantile deviation; dispersion inequalities; Monte Carlo simulation.

2020 Mathematics Subject Classification: 62G35; 62G30; 62F10; 62F40.

1 Introduction

The description of variability is a fundamental component of statistical analysis, since the location of a distribution alone is insufficient to characterize the dispersion, asymmetry, or tail behaviour of the underlying observations. Classical measures of dispersion, including the variance, standard deviation, mean absolute deviation, and interquartile range, quantify different aspects of variation and may therefore provide complementary information about a distribution. Their interpretations depend, in part, on the reference location around which deviations are evaluated and on the manner in which those deviations are summarized. Consequently, the choice of a dispersion measure is not merely computational; it also determines which features of the distribution are emphasized. These considerations are particularly relevant when the distribution is skewed, heavy-tailed, or multimodal, since different location summaries need not coincide and may respond differently to extreme observations [2, 1].

The mean, median, mode, and quantiles represent distinct approaches to describing the location of a probability distribution. Under suitable moment conditions, the mean provides a natural reference for quadratic loss and is directly related to the variance through squared deviations. The median, in contrast, minimizes expected absolute loss under the usual integrability conditions and is consequently closely associated with absolute-deviation-based summaries. Quantiles characterize the distribution through specified probability levels and allow locations other than the centre to be investigated. The mode identifies a maximizer of the density, when a density exists, and may be particularly relevant when the concentration of probability near a

*Corresponding author: sthdas999@gmail.com

density peak is of interest. These location functionals describe different distributional features and need not yield the same numerical reference point. In symmetric distributions satisfying appropriate regularity conditions, some of these functionals may coincide; under skewness or multimodality, however, their differences may become substantial. The resulting choice of reference location can therefore affect the numerical description of dispersion.

Absolute deviations provide a natural means of constructing dispersion measures relative to such reference locations. For a random variable X with distribution function F, let $T(F)$ denote a selected location functional. The corresponding absolute-deviation random variable is

$$D_T = |X - T(F)|.$$

The distribution of D_T describes the magnitudes of departures from the chosen reference point without distinguishing positive from negative deviations. Different summaries of this random variable yield different measures of dispersion. For example, its expectation gives the mean absolute deviation about $T(F)$, its median gives the median absolute deviation about that reference, and its pth quantile describes the deviation magnitude associated with a specified probability level. Thus, even after a reference location has been selected, the dispersion description is not uniquely determined: it also depends on the functional applied to the absolute deviations.

The distinction between the reference location and the summary of the deviations is important in the study of robust statistics. The median absolute deviation about the median is a well-established robust scale estimator, valued for its bounded influence function and high breakdown point under the standard formulation [4]. More generally, robust statistical procedures have been developed to reduce sensitivity to contamination and departures from idealized distributional assumptions [1, 2]. Nevertheless, different measures involving absolute deviations should not automatically be regarded as equally robust. For example, the mean absolute deviation about the median still uses an expectation and does not possess the same breakdown properties as the median absolute deviation about the median. The robustness of a deviation measure therefore depends on both the reference functional and the functional used to summarize the resulting deviations. Separating these two choices makes it possible to examine their respective roles without conflating measures that share an absolute-deviation construction but have different statistical properties.

The classical relationships among these measures provide a useful starting point for a unified treatment. If m is a median of X and $\mu = \mathbb{E}(X)$, with $\mathbb{E}|X| < \infty$, the median minimizes expected absolute loss. Consequently,

$$\mathbb{E}|X - m| \le \mathbb{E}|X - \mu|.$$

If, in addition, $\mathbb{E}(X^2) < \infty$, the Cauchy–Schwarz inequality yields

$$\mathbb{E}|X - \mu| \le \sqrt{\mathbb{E}(X - \mu)^2} = \sigma,$$

where σ denotes the standard deviation. These inequalities establish an ordering among selected population-level measures: the mean absolute deviation about a median is bounded by the mean absolute deviation about the mean, which is itself bounded by the standard deviation. The inequalities do not, however, imply a universal ordering among all possible summaries of absolute deviations. In particular, the median, quantiles, and mode of D_T represent different features of its distribution and require separate treatment.

Quantile-based summaries offer a complementary description of dispersion because they distinguish the central portion of the absolute-deviation distribution from its upper tail. For a nonnegative random variable D, its pth quantile may be defined by

$$Q_D(p) = \inf\{d : F_D(d) \ge p\}, \qquad 0 < p < 1,$$

where F_D is the distribution function of D. The median of D corresponds to $Q_D(0.5)$ under this convention, whereas larger probability levels characterize increasingly extreme deviations. Such

summaries can be particularly informative when the distribution is skewed or heavy-tailed, since a single measure of average deviation may not adequately describe the upper portion of the deviation distribution. Quantiles are also central to the broader theory of statistical inference and distributional description [3]. However, the interpretation of a quantile-based deviation must remain distinct from that of an expected deviation: the former describes a specified position in the distribution of absolute deviations, whereas the latter aggregates their magnitudes through an expectation.

Mode-based deviations require additional care. Unlike the mean and median, a mode is defined through maximization of a density and need not be unique. In a multimodal distribution, the global mode may represent only one of several regions of high density, and its estimated location may depend on the method used to estimate the density. Accordingly, a mode-centred deviation can provide a meaningful description of departures from a density peak, but it does not generally minimize expected absolute loss and cannot be assumed to satisfy the inequalities that hold for mean- or median-centred expected absolute deviations. Furthermore, a mode-based summary of the absolute deviations, such as the mode of $|X - T(F)|$, is conceptually distinct from the mode used as the reference location itself. These distinctions motivate explicit treatment of the reference and summary functionals as separate components of the construction.

The present work develops a unified formulation in which the reference location and the functional used to summarize the absolute deviations may be selected independently. Let $T(F)$ denote a reference functional and let S denote a suitable functional acting on the distribution of the absolute deviations. The resulting deviation measure can be represented abstractly as

$$\mathcal{D}_{S,T}(F) = S\big(\mathcal{L}\big(|X - T(F)|\big)\big)\,,$$

where $\mathcal{L}(|X - T(F)|)$ denotes the probability distribution of the absolute-deviation random variable. The precise definition and admissibility of this expression depend on the functional S and the conditions under which the relevant distributional quantity exists. This formulation separates two statistical choices: T determines the point from which deviations are measured, while S determines how their magnitudes are summarized. When S is the expectation and T is the median, the construction gives the mean absolute deviation about the median. When both functionals are selected to be the mean, the usual mean absolute deviation about the mean is recovered. Other choices yield median-based, quantile-based, and, when well defined, mode-based summaries of the absolute deviations. The framework therefore places several familiar quantities within a common mathematical representation without presupposing that they share the same robustness, existence, or ordering properties.

Three related objectives guide the investigation. The first is to formulate the construction precisely and clarify the roles of the reference and deviation functionals. The second is to establish rigorous inequalities for selected members of the resulting family under explicit moment, integrability, and quantile assumptions. The third is to examine how distributional characteristics influence the numerical relationships among these measures. The theoretical analysis focuses on population-level relationships, whereas the computational investigation examines sample-based deviations across distributions with different degrees of symmetry, skewness, tail weight, and multimodality. This distinction is essential: a mathematical inequality established under specified assumptions is a population-level result, while agreement with that inequality in a finite sample or Monte Carlo experiment constitutes numerical evidence rather than a proof.

The computational component is designed to complement the theoretical results rather than replace them. By considering distributions with contrasting shapes, the investigation examines whether the relative behaviour of mean-centred, median-centred, and quantile-based deviations changes with distributional characteristics and sample size. Particular attention is given to the distinction between symmetric and skewed distributions, the possible influence of heavy tails on sampling variability, and the sensitivity of mode-based summaries to density estimation. For the latter, bandwidth sensitivity and the number of detected local density maxima provide numerical

diagnostics of the estimated density shape. Such diagnostics must be interpreted cautiously, since the estimated mode may vary with the bandwidth and the rule used to identify a global maximum. The analysis therefore treats mode-based behaviour separately from the inequalities established for expected absolute deviations.

The contribution of the study is thus threefold. First, a common functional representation is provided for deviations defined relative to potentially different statistical functionals. Second, selected distribution-free inequalities are established under clearly stated conditions, thereby identifying relationships that can be justified mathematically rather than inferred solely from numerical comparisons. Third, a simulation-based investigation is used to illustrate how the resulting measures behave across representative distributional settings and to identify aspects requiring further study, particularly the finite-sample variability of quantile-based summaries and the estimation sensitivity of mode-based quantities. The proposed formulation is not intended to impose a universal ranking on all functional deviations; rather, it provides a structured basis for determining which comparisons are mathematically valid and which depend on the underlying distribution or estimation procedure.

The remainder of the paper is organized as follows. The general construction and the relevant assumptions are introduced first. The principal inequalities for selected deviation measures are then established, followed by numerical investigations across representative distributions. The sensitivity of mode-based summaries is considered separately, and the final sections discuss limitations and possible extensions to broader classes of statistical functionals, multivariate observations, and functional data.

2 Notation and definitions

Let X be a real-valued random variable with distribution function F. For a measurable function g, write F_g for the distribution of $g(X)$.

A statistical functional T assigns a numerical characteristic to a distribution. Its value at F is written $T(F)$. For convenience, we also write $T(X)$ when the underlying distribution is understood.

Examples include

$$\mu = \mathbb{E}[X], \qquad m = \mathrm{Med}(X), \qquad \theta = \mathrm{Mode}(X), \qquad q_p = Q_F(p).$$

The existence and uniqueness of these quantities are assumed whenever required.

Definition 2.1 (Functional deviation about a functional)**.** Let T be a statistical functional such that $T(F)$ is well-defined. The functional deviation of X about T is defined as

$$\boxed{D_T(X) = T\big(F_{|X-T(F)|}\big)\,.} \tag{1}$$

Equivalently, using the abbreviated notation,

$$D_T(X) = T\big(|X - T(X)|\big).$$

In this definition, the functional is first applied to the distribution of X to obtain the reference value. The same functional is then applied to the distribution of the absolute deviations.

Definition 2.2 (Two-functional deviation)**.** Let T determine the reference location and let S summarize the absolute deviations. Define

$$\boxed{D_{S|T}(X) = S\big(F_{|X-T(F)|}\big)\,.} \tag{2}$$

In abbreviated notation,

$$D_{S|T}(X) = S\big(|X - T(X)|\big).$$

The original construction is recovered when $S = T$:

$$D_T(X) = D_{T|T}(X).$$

Definition 2.3 (Powered functional deviation)**.** For $r > 0$, define

$$\boxed{D^{(r)}_{S|T}(X) = S\big(F_{|X-T(F)|^r}\big)\,.} \tag{3}$$

The powered formulation allows the deviation to be summarized after transforming the absolute deviations through a positive power.

3 Examples of functional deviations

3.1 Mean deviation about the mean

Taking $S = T = \mathbb{E}$ gives

$$D_{\mathbb{E}|\mathbb{E}}(X) = \mathbb{E}\,|X - \mu|\,. \tag{4}$$

This is the mean absolute deviation about the mean.

3.2 Median deviation about the median

Taking $S = T = \text{Med}$ gives

$$D_{\text{Med}|\text{Med}}(X) = \text{Med}\,|X - m|\,. \tag{5}$$

This is the conventional median absolute deviation, provided the median of the absolute deviations is defined using the stated quantile convention.

3.3 Mode deviation about the mode

If a unique mode θ exists, the mode-based deviation is

$$D_{\text{Mode}|\text{Mode}}(X) = \text{Mode}\,|X - \theta|\,. \tag{6}$$

The mode of the absolute deviations need not be unique. Consequently, this quantity requires a specified mode selection rule whenever multiple modes are possible.

3.4 Quantile deviation about a quantile

For $0 < p < 1$, let $q_p = Q_F(p)$. Then

$$D_{Q_p|Q_p}(X) = Q_p(|X - q_p|)\,. \tag{7}$$

This is the p-th quantile of the absolute deviations from the p-th quantile of the original distribution.

3.5 Mixed-functional deviations

The two-functional definition also includes

$$D_{\mathbb{E}|\text{Med}}(X) = \mathbb{E}\,|X - m|\,, \tag{8}$$

$$D_{\text{Med}|\mathbb{E}}(X) = \text{Med}\,|X - \mu|\,, \tag{9}$$

$$D_{Q_p|Q_q}(X) = Q_p\,|X - Q_q(F)|\,. \tag{10}$$

The distinction between the two functionals is important: T determines the centre, while S summarizes the deviations from that centre.

4 Basic properties

Throughout this section, let X be a real-valued random variable with distribution function F, and let T and S denote statistical functionals. The functional $T(F)$ is assumed to be well-defined, and the distribution of $|X - T(F)|$ is assumed to belong to the domain of S. The functional deviation is defined by

$$D_{S|T}(X) = S\big(F_{|X-T(F)|}\big)\,,$$

where $F_{|X-T(F)|}$ denotes the distribution function of the nonnegative random variable $|X - T(F)|$. For $r > 0$, the corresponding powered functional deviation is

$$D^{(r)}_{S|T}(X) = S\big(F_{|X-T(F)|^r}\big)\,.$$

The following propositions establish nonnegativity, translation invariance, and positive scale equivariance under suitable conditions on the functionals.

Proposition 4.1 (Nonnegativity of functional deviations)**.** *Suppose that S is a nonnegative-valued statistical functional on its domain; that is,*

$$S(G) \geq 0$$

for every distribution function G in its domain. Then, for every random variable X for which $D_{S|T}(X)$ is well-defined,

$$\boxed{D_{S|T}(X) \geq 0.}$$

In particular, if $S = T$, then

$$D_{T|T}(X) \geq 0.$$

The same conclusion holds for the powered deviation $D^{(r)}_{S|T}(X)$ for every $r > 0$, provided that the corresponding functional is well-defined.

Proof. Let

$$Y_T = |X - T(F)|.$$

Since the absolute value of every real number is nonnegative, it follows that

$$Y_T \geq 0 \qquad \text{almost surely.}$$

Consequently, the distribution function F_{Y_T} is supported on $[0, \infty)$, in the sense that

$$F_{Y_T}(y) = 0, \qquad y < 0.$$

By the assumed nonnegative-valued property of S,

$$D_{S|T}(X) = S(F_{Y_T}) \geq 0.$$

Setting $S = T$ gives the same conclusion for the same-functional deviation.

For the powered deviation, observe that $r > 0$ implies

$$Y_T^r = |X - T(F)|^r \geq 0 \qquad \text{almost surely.}$$

Its distribution is therefore also supported on $[0, \infty)$. Applying the nonnegative-valued functional S to this distribution yields

$$D^{(r)}_{S|T}(X) = S(F_{Y_T^r}) \geq 0.$$

□

Remark 4.2. Nonnegativity does not imply that a functional deviation is strictly positive. Whether $D_{S|T}(X) = 0$ can occur depends on the functional S and on the distribution of the absolute deviations. For example, if S is expectation, then

$$\mathbb{E}|X - T(F)| = 0$$

if and only if $X = T(F)$ almost surely. In contrast, if S is a median, a zero median of the absolute deviations does not necessarily imply that $X = T(F)$ almost surely, because at least half of the probability mass may be concentrated at the reference point.

Proposition 4.3 (Translation invariance)**.** *Suppose that T is translation equivariant, meaning that for every $c \in \mathbb{R}$,*

$$T(F_{X+c}) = T(F_X) + c.$$

Suppose also that S depends only on the distribution of its argument. Then

$$\boxed{D_{S|T}(X + c) = D_{S|T}(X).}$$

The same result holds for $D^{(r)}_{S|T}(X)$ for every $r > 0$, whenever the relevant functionals exist.

Proof. Let F_X denote the distribution of X, and write

$$T_X = T(F_X), \qquad T_{X+c} = T(F_{X+c}).$$

By translation equivariance,

$$T_{X+c} = T_X + c.$$

The absolute deviation of $X + c$ from its reference functional is therefore

$$\begin{aligned} |(X + c) - T_{X+c}| &= |(X + c) - (T_X + c)| \\ &= |X - T_X|. \end{aligned}$$

This equality holds pointwise, and hence almost surely. In particular, the random variables $|(X + c) - T_{X+c}|$ and $|X - T_X|$ have identical distributions:

$$F_{|(X+c)-T_{X+c}|} = F_{|X-T_X|}.$$

Applying the same distributional functional S to both sides gives

$$\begin{aligned} D_{S|T}(X + c) &= S\big(F_{|(X+c)-T_{X+c}|}\big) \\ &= S\big(F_{|X-T_X|}\big) \\ &= D_{S|T}(X). \end{aligned}$$

Thus, the functional deviation is invariant under translation.

For the powered deviation, the preceding pointwise identity also gives

$$|(X + c) - T_{X+c}|^r = |X - T_X|^r.$$

The powered absolute-deviation random variables therefore have identical distributions. Applying S to their common distribution establishes

$$D^{(r)}_{S|T}(X + c) = D^{(r)}_{S|T}(X).$$

□

Remark 4.4. Translation equivariance is satisfied by common location functionals, including the mean, median, quantiles, and a suitably defined mode. Thus, whenever these functionals are well-defined and translation equivariant, their associated functional deviations are invariant under changes in the origin of measurement. For example, converting a temperature variable from Celsius to a shifted temperature scale does not alter an absolute deviation from a translation-equivariant reference functional.

Proposition 4.5 (Positive scale equivariance)**.** *Let $a > 0$. Suppose that the reference functional T is positively scale equivariant:*

$$T(F_{aX}) = aT(F_X).$$

Suppose further that S is positively scale equivariant on nonnegative random variables, in the sense that

$$S(F_{aY}) = aS(F_Y), \qquad a > 0,$$

whenever $Y \geq 0$ and the relevant functionals exist. Then

$$\boxed{D_{S|T}(aX) = aD_{S|T}(X).}$$

Moreover, for every $r > 0$,

$$\boxed{D^{(r)}_{S|T}(aX) = a^r D^{(r)}_{S|T}(X).}$$

Proof. Write

$$T_X = T(F_X), \qquad T_{aX} = T(F_{aX}).$$

By the positive scale equivariance of T,

$$T_{aX} = aT_X.$$

It follows that

$$\begin{aligned} |aX - T_{aX}| &= |aX - aT_X| \\ &= a|X - T_X|, \end{aligned}$$

where the last equality uses $a > 0$. Hence, if

$$Y_T = |X - T_X|,$$

then the absolute-deviation random variable associated with aX is exactly aY_T. Therefore,

$$F_{|aX - T_{aX}|} = F_{aY_T}.$$

Using the definition of the functional deviation and the positive scale equivariance of S, we obtain

$$\begin{aligned} D_{S|T}(aX) &= S\big(F_{|aX - T_{aX}|}\big) \\ &= S(F_{aY_T}) \\ &= aS(F_{Y_T}) \\ &= aD_{S|T}(X). \end{aligned}$$

This proves the first assertion.

For the powered deviation, the same pointwise identity gives

$$\begin{aligned} |aX - T_{aX}|^r &= |aX - aT_X|^r \\ &= a^r |X - T_X|^r. \end{aligned}$$

Let $W_T = |X - T_X|^r$. The powered-deviation random variable associated with aX consequently has the same distribution as $a^r W_T$. Positive scale equivariance of S now yields

$$\begin{aligned} D^{(r)}_{S|T}(aX) &= S\big(F_{|aX - T_{aX}|^r}\big) \\ &= S(F_{a^r W_T}) \\ &= a^r S(F_{W_T}) \\ &= a^r D^{(r)}_{S|T}(X). \end{aligned}$$

The second assertion follows. □

Remark 4.6. The restriction $a > 0$ ensures that the scale transformation preserves the ordering of observations and that $|au| = a|u|$. For $a < 0$, a corresponding result may hold if the reference functional is equivariant under reflection and the deviation functional satisfies the required positive scaling property with factor $|a|$. The result is not asserted without these additional conditions.

The expectation, median, and quantile functionals are positively scale equivariant under their usual definitions. A mode functional is also positively scale equivariant for a density transformed under a positive scale change, provided the mode is uniquely defined or a scale-equivariant selection rule is used.

Corollary 4.7 (Location-scale transformation)**.** *Suppose the assumptions of the preceding translation and scale propositions hold. For $a > 0$ and $c \in \mathbb{R}$,*

$$\boxed{D_{S|T}(aX + c) = aD_{S|T}(X).}$$

For the powered deviation,

$$\boxed{D^{(r)}_{S|T}(aX + c) = a^r D^{(r)}_{S|T}(X).}$$

Proof. First apply translation equivariance of T:

$$T(aX + c) = T(aX) + c.$$

Positive scale equivariance then gives

$$T(aX + c) = aT(X) + c.$$

Consequently,

$$\begin{aligned} |(aX + c) - T(aX + c)| &= |(aX + c) - (aT(X) + c)| \\ &= a|X - T(X)|. \end{aligned}$$

Applying S to the distribution of this random variable and using its positive scale equivariance establishes

$$D_{S|T}(aX + c) = aD_{S|T}(X).$$

Raising the pointwise identity to the power r gives

$$|(aX + c) - T(aX + c)|^r = a^r |X - T(X)|^r.$$

A final application of the scale equivariance of S proves the powered result. □

5 Theoretical inequalities

The functional deviation framework permits comparisons between different measures of dispersion by selecting suitable reference functionals and deviation-summary functionals. The following results establish several fundamental inequalities involving the mean, median, standard deviation, and quantiles. Particular attention is given to the assumptions required for each inequality and to the conditions under which equality can occur.

5.1 Mean absolute deviation about the median and the mean

Lemma 5.1 (Median minimizes expected absolute loss)**.** *Suppose that* $\mathbb{E}|X| < \infty$*, and let* m *be any median of* X*. Then, for every* $a \in \mathbb{R}$*,*

$$\boxed{\mathbb{E}|X - m| \leq \mathbb{E}|X - a|.} \tag{11}$$

Proof. Define the expected absolute-loss function

$$L(a) = \mathbb{E}|X - a|, \qquad a \in \mathbb{R}.$$

The integrability condition $\mathbb{E}|X| < \infty$ ensures that $L(a) < \infty$ for every finite a, because

$$|X - a| \leq |X| + |a|.$$

Furthermore, since the absolute-value function is convex, L is convex as an expectation of convex functions.

For a fixed realization of X, the function $a \mapsto |X - a|$ has derivative -1 when $a < X$ and 1 when $a > X$. At $a = X$, its subdifferential is the interval $[-1, 1]$. Consequently, the subdifferential of L at any $a \in \mathbb{R}$ is

$$\partial L(a) = \left[\Pr(X < a) - \Pr(X \geq a),\, \Pr(X \leq a) - \Pr(X > a)\right].$$

For a convex function, a point minimizes the function if and only if zero belongs to its subdifferential. Since m is a median,

$$\Pr(X \leq m) \geq \frac{1}{2}, \qquad \Pr(X \geq m) \geq \frac{1}{2}.$$

The first inequality implies

$$\Pr(X \leq m) - \Pr(X > m) \geq 0,$$

whereas the second gives

$$\Pr(X < m) - \Pr(X \geq m) \leq 0.$$

It follows that

$$0 \in \partial L(m).$$

Thus m is a global minimizer of L, and hence

$$L(m) \leq L(a)$$

for every $a \in \mathbb{R}$. Substitution of the definition of L establishes the desired inequality. □

Theorem 5.2 (Comparison of mean absolute deviations)**.** *Suppose that* $\mathbb{E}|X| < \infty$*. Let* $\mu = \mathbb{E}[X]$ *and let* m *be any median of* X*. Then*

$$\boxed{D_{\mathbb{E}|\mathrm{Med}}(X) \leq D_{\mathbb{E}|\mathbb{E}}(X).} \tag{12}$$

Equivalently,

$$\mathbb{E}|X - m| \leq \mathbb{E}|X - \mu|.$$

Proof. By the definition of the functional deviation,

$$D_{\mathbb{E}|\mathrm{Med}}(X) = \mathbb{E}|X - m|, \qquad D_{\mathbb{E}|\mathbb{E}}(X) = \mathbb{E}|X - \mu|.$$

The median-minimization lemma establishes that the expected absolute loss is minimized at any median. In particular, the choice $a = \mu$ in Lemma 5.1 yields

$$\mathbb{E}|X - m| \leq \mathbb{E}|X - \mu|.$$

Therefore, the mean absolute deviation about the median cannot exceed the mean absolute deviation about the mean. □

Remark 5.3 (Equality and interpretation)**.** Equality holds if and only if the mean μ is also a minimizer of the expected absolute-loss function $a \mapsto \mathbb{E}|X - a|$; equivalently, μ is a median of X. Thus, symmetry about the mean provides a sufficient condition for equality whenever the mean is finite, but symmetry is not necessary. The result identifies the median as an optimal reference functional when dispersion is summarized by the expectation of absolute deviations.

5.2 Mean absolute deviation and standard deviation

Theorem 5.4 (Mean absolute deviation is bounded by standard deviation)**.** *Suppose that* $\mathbb{E}[X^2] < \infty$*, and write* $\mu = \mathbb{E}[X]$*. Then*

$$\boxed{D_{\mathbb{E}|\mathbb{E}}(X) \leq \mathrm{SD}(X).} \tag{13}$$

Equality holds if and only if $|X - \mu|$ *is constant almost surely.*

Proof. The finite second-moment assumption guarantees that μ and $\mathrm{Var}(X)$ are well-defined and finite. Applying the Cauchy–Schwarz inequality to the random variables $|X - \mu|$ and 1 gives

$$(\mathbb{E}|X - \mu|)^2 = \big(\mathbb{E}\big[|X - \mu| \cdot 1\big]\big)^2 \leq \mathbb{E}\left[|X - \mu|^2\right] \mathbb{E}[1^2].$$

Since $\mathbb{E}[1^2] = 1$ and $|X - \mu|^2 = (X - \mu)^2$, it follows that

$$(\mathbb{E}|X - \mu|)^2 \leq \mathbb{E}[(X - \mu)^2] = \mathrm{Var}(X).$$

Both sides are nonnegative. Taking square roots therefore yields

$$\mathbb{E}|X - \mu| \leq \sqrt{\mathrm{Var}(X)} = \mathrm{SD}(X),$$

which proves the inequality.

The equality condition follows from the equality criterion for Cauchy–Schwarz: the random variables $|X - \mu|$ and 1 must be linearly dependent almost surely. This is equivalent to $|X - \mu|$ being constant almost surely. □

Corollary 5.5 (Combined inequality for median-based and mean-based deviations)**.** *If* $\mathbb{E}[X^2] < \infty$*, with* $\mu = \mathbb{E}[X]$ *and* m *any median of* X*, then*

$$\boxed{D_{\mathbb{E}|\mathrm{Med}}(X) \leq D_{\mathbb{E}|\mathbb{E}}(X) \leq \mathrm{SD}(X).} \tag{14}$$

Proof. The finite second-moment assumption implies $\mathbb{E}|X| < \infty$, so both Theorem 5.2 and Theorem 5.4 apply. The former gives

$$D_{\mathbb{E}|\mathrm{Med}}(X) \leq D_{\mathbb{E}|\mathbb{E}}(X),$$

while the latter gives

$$D_{\mathbb{E}|\mathbb{E}}(X) \leq \mathrm{SD}(X).$$

Combining these inequalities establishes (14). □

Remark 5.6 (Equality in the combined inequality)**.** For a nondegenerate random variable, equality in $D_{\mathbb{E}|\mathbb{E}}(X) \leq \mathrm{SD}(X)$ requires

$$|X - \mu| = c \quad \text{almost surely}$$

for some $c > 0$. Hence X is concentrated on the two points $\mu - c$ and $\mu + c$. The condition $\mathbb{E}[X] = \mu$ then forces both points to have probability $1/2$. Consequently, for a nondegenerate distribution, equality in the mean-absolute-deviation–standard-deviation inequality occurs precisely for a two-point distribution with equal probabilities at symmetric locations about the mean. For a degenerate random variable, both quantities are zero.

5.3 Median of absolute deviations

Lemma 5.7 (Median bounded by twice the expectation)**.** *Let Y be a nonnegative random variable satisfying $\mathbb{E}[Y] < \infty$. If m_Y is any median of Y, then*

$$\boxed{m_Y \leq 2\mathbb{E}[Y].} \tag{15}$$

Proof. Since $Y \geq 0$ almost surely, every median m_Y is nonnegative. By the median property,

$$\Pr(Y \geq m_Y) \geq \frac{1}{2}.$$

Furthermore, nonnegativity of Y implies

$$Y \geq m_Y \mathbf{1}\{Y \geq m_Y\} \quad \text{almost surely.}$$

Taking expectations on both sides gives

$$\mathbb{E}[Y] \geq m_Y \mathbb{E}[\mathbf{1}\{Y \geq m_Y\}] = m_Y \Pr(Y \geq m_Y).$$

Using the median property, we obtain

$$\mathbb{E}[Y] \geq \frac{m_Y}{2}.$$

Rearrangement yields $m_Y \leq 2\mathbb{E}[Y]$, as required. The argument also covers $m_Y = 0$, for which the inequality is immediate. □

Corollary 5.8 (Median absolute deviation bound)**.** *Let $a \in \mathbb{R}$ be a fixed reference value and suppose that $\mathbb{E}|X - a| < \infty$. Then*

$$\boxed{D_{\mathrm{Med}|a}(X) \leq 2D_{\mathbb{E}|a}(X),} \tag{16}$$

where the notation denotes the median and expectation, respectively, of $|X - a|$. In particular, for $\mu = \mathbb{E}[X]$,

$$\mathrm{Med}\,|X - \mu| \leq 2\mathbb{E}|X - \mu|.$$

Proof. Set $Y = |X - a|$. This random variable is nonnegative and has finite expectation by assumption. Its median is precisely the median of the absolute deviations from a. Applying Lemma 5.7 gives

$$\mathrm{Med}\,|X - a| \leq 2\mathbb{E}|X - a|,$$

which establishes the result. □

Remark 5.9 (Nature of the bound)**.** The factor 2 is a general distribution-free bound obtained from nonnegativity and the median property. It does not assert that the median of absolute deviations is always close to its expectation in a relative sense. Sharper bounds may be possible under additional assumptions on the distribution of $|X - a|$.

5.4 Quantile bounds for nonnegative deviations

For a nonnegative random variable Y, define its p-th quantile, for $0 < p < 1$, by

$$Q_Y(p) = \inf\{y \in \mathbb{R} : F_Y(y) \geq p\},$$

where F_Y denotes the cumulative distribution function of Y.

Theorem 5.10 (Quantile bounded by the expectation)**.** *Let $Y \geq 0$ almost surely and suppose that $\mathbb{E}[Y] < \infty$. Then, for every $p \in (0,1)$,*

$$\boxed{Q_Y(p) \leq \frac{\mathbb{E}[Y]}{1-p}.} \tag{17}$$

Proof. Write $q = Q_Y(p)$. Since Y is nonnegative, $q \geq 0$. If $q = 0$, the desired result follows immediately from $\mathbb{E}[Y] \geq 0$.

Now suppose that $q > 0$. By the definition of the quantile as the lower generalized inverse of the distribution function, one has

$$F_Y(q-) \leq p,$$

where $F_Y(q-) = \Pr(Y < q)$ denotes the left limit of the distribution function at q. Therefore,

$$\Pr(Y \geq q) = 1 - \Pr(Y < q) \geq 1 - p.$$

The nonnegativity of Y further gives the almost-sure inequality

$$Y \geq q\mathbf{1}\{Y \geq q\}.$$

Taking expectations yields

$$\mathbb{E}[Y] \geq q \Pr(Y \geq q) \geq q(1-p).$$

Since $1 - p > 0$, division by $1 - p$ gives

$$q \leq \frac{\mathbb{E}[Y]}{1-p}.$$

Substituting $q = Q_Y(p)$ completes the proof. □

Corollary 5.11 (Quantile deviation inequality)**.** *Let $T(X)$ be a well-defined reference functional and suppose that*

$$\mathbb{E}|X - T(X)| < \infty.$$

Then, for every $p \in (0,1)$,

$$\boxed{D_{Q_p|T}(X) \leq \frac{D_{\mathbb{E}|T}(X)}{1-p}.} \tag{18}$$

Proof. Define the nonnegative deviation variable

$$Y_T = |X - T(X)|.$$

By assumption, $\mathbb{E}[Y_T] < \infty$. The definitions of the quantile-based and expectation-based functional deviations give

$$D_{Q_p|T}(X) = Q_{Y_T}(p), \qquad D_{\mathbb{E}|T}(X) = \mathbb{E}[Y_T].$$

Application of Theorem 5.10 therefore yields

$$D_{Q_p|T}(X) = Q_{Y_T}(p) \leq \frac{\mathbb{E}[Y_T]}{1-p} = \frac{D_{\mathbb{E}|T}(X)}{1-p},$$

which proves the claim. □

Remark 5.12 (Dependence on the quantile level)**.** The upper bound increases as p approaches one. This reflects the fact that upper quantiles of the absolute deviation can be substantially larger than its expectation. The inequality remains valid without assuming continuity of the distribution of Y_T; atoms and ties are permitted under the stated generalized-inverse convention.

5.5 Ordering of quantile deviations

Proposition 5.13 (Monotonicity in the quantile level)**.** *Let $T(X)$ be a fixed, well-defined reference functional, and define $Y_T = |X - T(X)|$. For any $0 < p_1 < p_2 < 1$,*

$$\boxed{D_{Q_{p_1}|T}(X) \leq D_{Q_{p_2}|T}(X).} \tag{19}$$

Proof. Let F_{Y_T} be the distribution function of Y_T. By definition,

$$Q_{Y_T}(p) = \inf\{y : F_{Y_T}(y) \geq p\}.$$

Because $p_1 < p_2$, every value y satisfying $F_{Y_T}(y) \geq p_2$ also satisfies $F_{Y_T}(y) \geq p_1$. Hence

$$\{y : F_{Y_T}(y) \geq p_2\} \subseteq \{y : F_{Y_T}(y) \geq p_1\}.$$

Taking the infimum over these sets yields

$$Q_{Y_T}(p_1) \leq Q_{Y_T}(p_2).$$

Finally, since $D_{Q_p|T}(X) = Q_{Y_T}(p)$, the desired inequality follows. □

Remark 5.14. The monotonicity result is distribution-free. It does not depend on the existence of moments or on continuity of the distribution of the absolute deviations. It states that increasing the quantile level cannot decrease the corresponding quantile-based functional deviation.

5.6 Why a universal mode inequality is not asserted

Remark 5.15 (Mode-based deviations)**.** A universal inequality analogous to Theorem 5.2 is not asserted for the mode. Unlike the median, which minimizes expected absolute loss, the mode generally does not minimize $a \mapsto \mathbb{E}|X - a|$. Moreover, the mode may be nonunique, particularly for discrete distributions or multimodal densities, and its value can depend on the convention used to select among competing modes.

Consequently, without additional distributional assumptions and a precise definition of the mode functional, neither

$$D_{\text{Mode}|\text{Mode}}(X) \leq D_{\mathbb{E}|\mathbb{E}}(X)$$

nor the reverse inequality can be concluded from the preceding results. Establishing mode-based inequalities requires a separate analysis of the distributional conditions and mode-selection rule.

6 Simulation study

6.1 Objectives

The simulation study is designed to investigate the empirical behaviour of the proposed functional deviations under a variety of distributional configurations and to examine whether the sample-based counterparts of the theoretical inequalities are satisfied in finite samples. The investigation focuses on the relationships between mean absolute deviation about the mean, mean absolute deviation about the median, standard deviation, median absolute deviation, and quantile-based deviations.

Six distributions are considered to represent different distributional characteristics, including symmetry, bounded support, positive skewness, heavy tails, and multimodality. This selection makes it possible to examine whether the comparative behaviour of the functional deviations depends on the underlying distribution and to identify differences between symmetric and asymmetric settings.

The distributions considered in the study are specified as follows:

(i) The standard normal distribution,

$$X \sim N(0,1),$$

representing a continuous, symmetric, unimodal distribution with finite moments.

(ii) The uniform distribution,

$$X \sim \text{Uniform}(-1,1),$$

representing a symmetric distribution with bounded support and constant density over its support.

(iii) The standard exponential distribution,

$$X \sim \text{Exponential}(1),$$

representing a nonnegative, positively skewed distribution with an exponentially decreasing tail.

(iv) Student's t distribution with three degrees of freedom,

$$X \sim t_3,$$

representing a symmetric, heavy-tailed distribution. Its variance is finite, whereas its fourth moment is infinite.

(v) The lognormal distribution,

$$X \sim \text{Lognormal}(0,1),$$

representing a positive and strongly right-skewed distribution whose mean and variance are finite.

(vi) A two-component Gaussian mixture,

$$X \sim 0.7N(0,1) + 0.3N(3, 0.5^2),$$

representing a mixture distribution with potentially distinct modal structure and an asymmetric overall distribution.

For each distribution, a random sample of size $N = 300000$ is generated using R. The relatively large sample size is intended to reduce Monte Carlo fluctuations in the empirical summaries and to provide a useful approximation to the corresponding population quantities. Nevertheless, the computed values remain sample estimates, and their agreement with the theoretical inequalities is interpreted as empirical verification rather than as a substitute for mathematical proof.

6.2 Step-by-step simulation procedure

The simulation is carried out through the following sequence of computational steps.

Step 1. Generation of random observations.

For each of the six distributions, $N = 300000$ independent and identically distributed observations are generated. The same sample size is used for all distributions to ensure that the empirical summaries can be compared under a common sample-size configuration. A fixed random seed is specified in the R implementation to make the numerical experiment reproducible.

Step 2. Estimation of location and dispersion.

For each generated sample $X_1, \ldots, X_N$, the sample mean, sample median, and standard deviation are calculated as

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \qquad \widehat{m} = \text{Median}(X_1, \ldots, X_N),$$

and

$$s = \left\{ \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2 \right\}^{1/2}.$$

Here, $\bar{X}$ and $\widehat{m}$ estimate the population mean and median, respectively, while s estimates the population standard deviation. The use of the conventional sample standard deviation with denominator $N-1$ is distinguished from the empirical root mean squared deviation with denominator N when numerical inequalities are checked.

Step 3. Calculation of mean absolute deviations.

Two expectation-based functional deviations are computed: the empirical mean absolute deviation about the sample mean,

$$\widehat{D}_{\mathbb{E}|\mathbb{E}} = \frac{1}{N}\sum_{i=1}^{N} |X_i - \bar{X}|,$$

and the empirical mean absolute deviation about the sample median,

$$\widehat{D}_{\mathbb{E}|\text{Med}} = \frac{1}{N}\sum_{i=1}^{N} |X_i - \widehat{m}|.$$

These quantities are compared to examine the empirical relationship between mean-centered and median-centered absolute deviations. In particular, the comparison assesses whether the median-centered quantity is no larger than the mean-centered quantity, as established by the corresponding population inequality.

Step 4. Calculation of median- and quantile-based deviations.

The median of the absolute deviations about the sample mean is calculated as

$$\widehat{D}_{\text{Med}|\mathbb{E}} = \text{Median}\big(|X_1 - \bar{X}|, \ldots, |X_N - \bar{X}|\big).$$

The distribution of absolute deviations about the sample median is then summarized through its empirical quantiles:

$$\widehat{Q}_{\text{abs},\widehat{m}}(p) = \widehat{Q}_p\big(|X_1 - \widehat{m}|, \ldots, |X_N - \widehat{m}|\big),$$

where $p \in \{0.25, 0.50, 0.75, 0.90\}$ and $\widehat{Q}_p$ denotes the empirical p-th quantile. These summaries describe the typical, central, and upper-tail magnitudes of the absolute deviations. In particular, the 90th percentile provides information about the upper portion of the deviation distribution that is not captured by the mean or median alone.

Step 5. Verification of sample analogues of theoretical inequalities.

The computed summaries are used to examine the following relationships:

(a) The median–mean inequality,

$$\widehat{D}_{\mathbb{E}|\text{Med}} \le \widehat{D}_{\mathbb{E}|\mathbb{E}}.$$

(b) The mean absolute deviation–standard deviation inequality, using the empirical root mean squared deviation

$$s_N^* = \left\{ \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})^2 \right\}^{1/2},$$

so that

$$\widehat{D}_{\mathbb{E}|\mathbb{E}} \leq s_N^*.$$

The denominator N is used here because the inequality follows directly from the empirical Cauchy–Schwarz inequality. The conventional sample standard deviation s may also be reported descriptively.

(c) The median–expectation bound for nonnegative deviations,

$$\widehat{D}_{\mathrm{Med}|\mathbb{E}} \leq 2\widehat{D}_{\mathbb{E}|\mathbb{E}}.$$

(d) The ordering of empirical quantiles,

$$\widehat{Q}_{\mathrm{abs},\widehat{m}}(0.25) \leq \widehat{Q}_{\mathrm{abs},\widehat{m}}(0.50) \leq \widehat{Q}_{\mathrm{abs},\widehat{m}}(0.75) \leq \widehat{Q}_{\mathrm{abs},\widehat{m}}(0.90).$$

(e) The quantile–expectation bound for the absolute deviations about the sample median,

$$\widehat{Q}_{\mathrm{abs},\widehat{m}}(p) \leq \frac{\widehat{D}_{\mathbb{E}|\mathrm{Med}}}{1-p}, \qquad p \in \{0.25, 0.50, 0.75, 0.90\}.$$

The empirical inequalities are evaluated using the same generated observations for each distribution. Their verification provides a direct computational check of the corresponding mathematical results. The distinction between the population quantities and their empirical counterparts is retained throughout the analysis.

Step 6. Comparison and graphical presentation.

The resulting summaries are compared across the six distributions to investigate how symmetry, skewness, tail behaviour, boundedness, and mixture structure influence the magnitude of the functional deviations. Particular attention is given to the difference between mean-centered and median-centered deviations and to the relative size of the mean absolute deviation compared with the standard deviation.

The numerical results are presented in tabular form, while selected comparisons are visualized using PGFPlots. The graphs are intended to facilitate comparison of the dispersion summaries across distributions and to illustrate the ordering of the quantile-based deviations. All plotted values are obtained from the simulation output, rather than from hypothetical or externally assigned numerical values.

The reproducible R code is available at Subsection (A.1) in Appendix.

6.3 Simulation results

The simulation results are examined to investigate the empirical behaviour of the proposed functional deviations under symmetric, bounded, skewed, heavy-tailed, and mixture distributions. The analysis is based on the numerical output obtained from the R implementation with $N = 300000$ observations generated from each distribution. The results are presented in three tables: the first compares the mean absolute deviations about the mean and median with the standard deviation; the second describes the quantile structure of the absolute deviations about the median; and the third reports the median of the absolute deviations about the mean.

6.3.1 Comparison of mean absolute deviations

Table 1 reports the mean absolute deviation about the sample median, $\widehat{D}_{\mathbb{E}|\mathrm{Med}}$, the mean absolute deviation about the sample mean, $\widehat{D}_{\mathbb{E}|\mathbb{E}}$, the estimated standard deviation, and the ratio of the two mean absolute deviations.

Table 1: Mean absolute deviations and standard deviations.

Distribution	$\widehat{D}_{\mathbb{E}\|\mathrm{Med}}$	$\widehat{D}_{\mathbb{E}\|\mathbb{E}}$	$\widehat{\sigma}$	Ratio
Normal	0.7973	0.7973	0.9993	1.0000
Uniform	0.4996	0.4996	0.5771	1.0000
Exponential	0.6922	0.7345	0.9983	0.9424
Student-$t(3)$	1.1044	1.1044	1.8004	1.0000
Lognormal	1.1228	1.2589	2.1418	0.8919
Normal mixture	1.3757	1.4007	1.6328	0.9821

The final column of Table 1 is calculated as

$$R_{\mathrm{MAD}} = \frac{\widehat{D}_{\mathbb{E}|\mathrm{Med}}}{\widehat{D}_{\mathbb{E}|\mathbb{E}}}.$$

This ratio measures the relative magnitude of the median-centered mean absolute deviation compared with the mean-centered counterpart. A value equal to one indicates that the two empirical deviations are equal, whereas a value below one indicates that the mean absolute deviation about the median is smaller. The latter behaviour is consistent with the theoretical inequality

$$D_{\mathbb{E}|\mathrm{Med}}(X) \leq D_{\mathbb{E}|\mathbb{E}}(X).$$

For the standard normal distribution, both mean absolute deviations are approximately 0.7973, and their ratio is 1.0000. This agreement is expected because the normal distribution is symmetric about its mean, and its population mean is also its median. In this setting, the mean and median provide the same population reference location. The empirical equality of the two deviations therefore illustrates the equality case of the median-minimization result. The estimated standard deviation is approximately 0.9993, which is close to the population value of one. Both mean absolute deviations are smaller than the standard deviation, consistent with the theoretical bound

$$D_{\mathbb{E}|\mathbb{E}}(X) \leq \mathrm{SD}(X).$$

The normal distribution thus provides a useful benchmark for comparing the two location-based measures of absolute dispersion.

A similar pattern is observed for the uniform distribution on $(-1, 1)$. The two mean absolute deviations are both approximately 0.4996, while the estimated standard deviation is 0.5771. The ratio is again 1.0000 to the reported precision. The uniform distribution is symmetric about zero, so its mean and median coincide. Consequently, the two absolute-deviation summaries have the same population value. The smaller standard deviation of the uniform distribution relative to that of the standard normal distribution also reflects the bounded support of the uniform model. All observations lie between -1 and 1, preventing the occurrence of the extreme observations possible under an unbounded normal distribution. The empirical results therefore illustrate how the support and shape of a distribution affect the scale of its functional deviations.

The exponential distribution exhibits a more noticeable difference between the two measures. The mean absolute deviation about the median is approximately 0.6922, whereas the mean absolute deviation about the mean is approximately 0.7345. Their ratio is 0.9424, meaning that the median-centered mean absolute deviation is approximately 5.76% smaller than the

mean-centered deviation. This difference is associated with the positive skewness of the exponential distribution. For an exponential random variable with rate one, the population mean is 1, whereas the population median is $\log 2 \approx 0.6931$. The median is therefore substantially smaller than the mean. The comparatively large observations in the right tail influence the sample mean more strongly than the sample median, whereas the median is determined by the central probability mass. Since the median minimizes expected absolute loss, the lower median-centered deviation is consistent with the theoretical result. The estimated standard deviation, approximately 0.9983, is close to the population standard deviation of one. The mean-centered absolute deviation remains smaller than this standard deviation, as required by the Cauchy–Schwarz inequality.

For Student's t distribution with three degrees of freedom, both mean absolute deviations are approximately 1.1044, and the reported ratio is 1.0000. This equality is consistent with the symmetry of the distribution about zero. Unlike the normal distribution, however, the t_3 distribution has substantially heavier tails. Its population variance exists, but its fourth moment is infinite. The estimated standard deviation of 1.8004 is consequently appreciably larger than the mean absolute deviation of 1.1044. This difference illustrates an important distinction between absolute-deviation measures and quadratic dispersion measures. Mean absolute deviation increases linearly with the magnitude of an observation's deviation, whereas variance is based on squared deviations and therefore gives greater weight to extreme observations. The heavier tails of the t_3 distribution can thus produce a comparatively large standard deviation relative to its mean absolute deviation. The results also show that the mean absolute deviation remains a meaningful finite dispersion measure even for distributions with heavier tails, provided the first absolute moment exists.

The lognormal distribution displays the largest difference between the two mean absolute deviations among the first five distributions considered. The mean absolute deviation about the median is approximately 1.1228, compared with 1.2589 about the mean. The ratio of 0.8919 indicates that the median-centered quantity is approximately 10.81% smaller than the mean-centered quantity. This difference is consistent with the strong right skewness of the lognormal distribution. Its population median is 1, while its population mean is $e^{1/2} \approx 1.6487$. The mean is pulled towards the upper tail by relatively large positive observations, whereas the median describes the central location of the distribution. Centering the absolute deviations at these two different locations therefore produces different average deviations. The relatively large estimated standard deviation of 2.1418 further reflects the influence of the right tail. The lognormal results provide particularly clear empirical evidence that the choice of reference functional can materially affect the measured dispersion when the underlying distribution is asymmetric.

For the Gaussian mixture, the mean absolute deviation about the median is approximately 1.3757, whereas the corresponding mean-centered deviation is approximately 1.4007. The ratio of 0.9821 indicates a difference of approximately 1.79% relative to the mean-centered deviation. The estimated standard deviation is 1.6328, which is larger than either absolute-deviation summary. The mixture combines a dominant component centered at zero with a secondary component centered at three. This separation introduces additional dispersion arising from differences between the component locations, in addition to the variation within each component. The resulting distribution is asymmetric overall and may exhibit more than one mode, depending on the component weights and separation. The difference between the two mean absolute deviations is smaller than that observed for the exponential and lognormal distributions, suggesting that the effect of the location choice is relatively modest in this particular simulated sample. Nevertheless, the median-centered deviation remains smaller, as predicted by the theoretical inequality. The ratio should be interpreted as a property of the simulated distribution and its component configuration, rather than as a general feature of Gaussian mixtures.

Taken together, the results in Table 1 demonstrate two important patterns. First, the

median-centered mean absolute deviation does not exceed its mean-centered counterpart in any of the six simulated samples. Second, the difference between these quantities depends on the distributional shape. The two quantities agree to the reported precision for the symmetric normal, uniform, and Student-$t(3)$ distributions, while noticeable differences occur for the skewed exponential and lognormal distributions. The Gaussian mixture shows a smaller but still visible difference. These findings are consistent with the theoretical role of the median as a minimizer of expected absolute loss. The degree of agreement between the two deviations, however, is not determined by symmetry alone in every possible distribution; it also depends on whether the mean is a median.

The results also support the distinction between absolute-deviation measures and standard deviation. For every distribution in the table, the reported mean absolute deviation about the mean is smaller than the estimated standard deviation. This is consistent with the population inequality, although the exact empirical comparison depends on the denominator used in calculating the standard deviation. The empirical root mean squared deviation with denominator N satisfies the corresponding sample inequality directly, while the conventional sample standard deviation uses denominator $N - 1$. The values reported here are close enough for the same qualitative pattern to be observed.

6.3.2 Quantile structure of absolute deviations

Table 2 reports the 25th, 50th, 75th, and 90th percentiles of the absolute deviations about the sample median. These quantiles provide a more detailed description of dispersion than a single average because they distinguish the central concentration of the deviations from their upper-tail behaviour.

Table 2: Quantiles of absolute deviations about the median.

Distribution	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
Normal	0.3186	0.6739	1.1495	1.642
Uniform	0.2494	0.4996	0.7494	0.900
Exponential	0.2474	0.4801	0.6926	1.609
Student-$t(3)$	0.3492	0.7662	1.4234	2.361
Lognormal	0.3095	0.5977	0.9677	2.595
Normal mixture	0.5413	1.2287	2.1748	2.707

For every distribution, the reported quantiles are nondecreasing as the probability level increases:

$$\widehat{Q}_{0.25} \leq \widehat{Q}_{0.50} \leq \widehat{Q}_{0.75} \leq \widehat{Q}_{0.90}.$$

This ordering is an intrinsic property of quantiles and is not specific to the six distributions selected for the simulation. The empirical results illustrate the proposition that higher quantile levels cannot produce smaller quantile-based deviations. The magnitude of the increments between quantiles, however, differs substantially across distributions and therefore provides information about their dispersion structures.

For the normal distribution, the absolute-deviation quantiles increase from 0.3186 at the 25th percentile to 1.642 at the 90th percentile. The median absolute deviation is approximately 0.6739, and the 75th percentile is 1.1495. The increase towards the upper quantiles reflects the contribution of observations farther from the central location. Because the normal distribution is symmetric and has relatively regular tails, its absolute deviations are distributed symmetrically around zero only in the original variable's location structure; the absolute deviations themselves are nonnegative and follow a folded normal-type distribution. The increasing quantiles therefore describe how the distance from the median accumulates as progressively larger proportions of observations are included.

The uniform distribution has the smallest 90th-percentile absolute deviation among the six distributions, at 0.900. Its quantiles increase from 0.2494 to 0.4996, 0.7494, and 0.900. The comparatively regular increments are consistent with the bounded support and constant density of the original distribution. For a population uniform distribution on $(-1, 1)$, the absolute deviation from its median zero is uniformly distributed on $(0, 1)$, so its population quantile function is $Q(p) = p$. The empirical quantiles in the table are close to these theoretical values. Small differences are expected from random sampling and the use of the estimated sample median. This result demonstrates how a bounded distribution restricts the upper range of absolute deviations, in contrast to distributions with unbounded support.

For the exponential distribution, the first three quantiles are 0.2474, 0.4801, and 0.6926, whereas the 90th percentile rises to 1.609. The relatively pronounced increase at the upper quantile reflects the right-skewed structure of the distribution. A substantial fraction of observations lies relatively close to the median, while a smaller fraction extends considerably farther into the right tail. Consequently, central quantiles alone would not fully describe the distribution of absolute deviations. The difference between the 75th and 90th percentiles is especially informative because it highlights the contribution of larger deviations. This pattern complements the mean-based results, which also show a difference between deviations centered at the mean and those centered at the median.

The Student-$t(3)$ distribution exhibits a 25th-percentile absolute deviation of 0.3492, a median of 0.7662, a 75th percentile of 1.4234, and a 90th percentile of 2.361. These upper quantiles are larger than the corresponding values for the normal distribution. The difference becomes more pronounced at higher probability levels, consistent with the heavier tails of the t_3 distribution. Observations far from the center occur more frequently than under the standard normal distribution, increasing the upper quantiles of absolute deviations. This illustrates why quantile summaries at several probability levels can be useful when comparing distributions with different tail behaviour. In particular, the 90th percentile provides a direct description of a relatively high deviation without depending as strongly on the magnitude of the most extreme observations as a squared-deviation measure would.

For the lognormal distribution, the 25th, 50th, and 75th percentiles are 0.3095, 0.5977, and 0.9677, respectively, while the 90th percentile reaches 2.595. The substantial increase between the 75th and 90th percentiles reflects the strong right-skewness of the original distribution. Although many observations remain relatively close to the median, a smaller fraction of observations can lie much farther away. Since the deviations are measured in absolute distance from the median, large observations in the right tail translate directly into large absolute deviations. The high 90th percentile is therefore consistent with the relatively large mean absolute deviation about the mean and the large standard deviation reported in Table 1. It also demonstrates that the median and lower quantiles, considered in isolation, may not adequately describe the upper tail of the deviation distribution.

The Gaussian mixture yields the largest 25th-, 50th-, and 75th-percentile absolute deviations among the six distributions, with values of 0.5413, 1.2287, and 2.1748, respectively. Its 90th percentile is 2.707. These values indicate a relatively broad distribution of distances from the overall sample median. The separation between the two Gaussian components contributes to this behaviour: observations from the component centered near three may be substantially displaced from the overall median, while observations from the component centered near zero may lie closer to it. The mixture structure therefore produces variation both within and between the components. The resulting quantiles reflect the combined distribution rather than the spread of either component considered separately. The comparatively large central quantiles indicate that the increased dispersion is not confined solely to a small number of extreme observations.

Across all six distributions, the quantile results show that the magnitude of a functional deviation depends on the probability level used to summarize the absolute deviations. The 25th percentile describes the distance within which a relatively concentrated quarter of the

observations falls, whereas the median provides a central summary of the deviation distribution. The 75th and 90th percentiles describe progressively more distant observations and are therefore useful for distinguishing bounded, light-tailed, heavy-tailed, skewed, and mixture settings. The simultaneous consideration of these quantiles offers a richer characterization of dispersion than a single mean absolute deviation.

6.3.3 Median of absolute deviations about the mean

Table 3 reports the median of the absolute deviations measured about the mean. Unlike the mean absolute deviation about the mean, which averages all absolute deviations, this quantity summarizes their central value. It therefore provides a robust complementary measure of dispersion.

Table 3: Median of absolute deviations about the mean.

Distribution	$\mathrm{Med}(\lvert X-\mu\rvert)$
Normal	0.6740
Uniform	0.4996
Exponential	0.6340
Student-$t(3)$	0.7661
Lognormal	0.9951
Normal mixture	1.3958

For the normal distribution, the median absolute deviation about the mean is 0.6740. This value is close to the population median of the absolute value of a standard normal random variable, approximately 0.6745. Since the normal distribution is symmetric about its mean, the mean is also the median. Hence, the median absolute deviation about the mean and the median absolute deviation about the median should agree at the population level. The small difference between the reported value 0.6740 and the median absolute deviation about the sample median, 0.6739, is attributable to sampling variation and rounding.

For the uniform distribution, the median absolute deviation about the mean is approximately 0.4996. This is consistent with the population result for a uniform distribution on $(-1, 1)$, for which the absolute deviation from zero is uniformly distributed on $(0, 1)$ and has median 0.5. The agreement between the median absolute deviation and the mean absolute deviation about the median in the table reflects the simple bounded and symmetric structure of this distribution. The result also illustrates that median-based and expectation-based summaries need not differ substantially when the absolute deviations have a regular distribution.

For the exponential distribution, the median absolute deviation about the mean is 0.6340, which is lower than the mean absolute deviation about the median, 0.6922. This is not a contradiction: the former quantity is the median of absolute deviations centered at the mean, whereas the latter is the expectation of absolute deviations centered at the median. The two summaries differ both in their reference location and in the functional used to summarize the resulting deviations. Their numerical values should therefore not be interpreted as a direct inequality comparison. Rather, the result illustrates how different combinations of reference location and deviation-summary functional can produce distinct measures of dispersion under asymmetry.

The corresponding value for Student's $t(3)$ distribution is 0.7661. This is close to the median of absolute deviations about the sample median, 0.7662, as expected from the symmetry of the population distribution. The agreement supports the observation that changing the reference from the mean to the median has little effect on the population deviation distribution when the two location functionals coincide. The numerical value is larger than the corresponding normal value, consistent with the greater spread and heavier tails of the t_3 distribution.

For the lognormal distribution, the median absolute deviation about the mean is 0.9951, substantially larger than the corresponding value for the exponential distribution. This reflects the larger overall scale of the lognormal distribution and its strong right skewness. The result should be distinguished from the mean absolute deviation about the median, which is 1.1228: the two summaries use different reference locations and different functionals. The median of absolute deviations about the mean is less influenced by the largest absolute deviations than their average, making it useful for describing the central magnitude of deviations in a distribution with a pronounced right tail.

For the Gaussian mixture, the median absolute deviation about the mean is 1.3958, the largest value in Table 3. This result is consistent with the broad spread produced by the separation between the two component means. The central absolute deviation from the overall mean is relatively large because the two components occupy different regions of the real line. The value also exceeds the median absolute deviation about the sample median, approximately 1.2287, indicating that the change in reference location affects the central summary of absolute deviations. The mixture example therefore illustrates that the effect of choosing the mean or median as the reference need not be limited to the expectation of absolute deviations; it can also influence median-based deviation summaries.

Overall, Table 3 complements the preceding tables by showing that the median of absolute deviations and the mean of absolute deviations capture different aspects of dispersion. The median emphasizes the central magnitude of the deviations, whereas the mean incorporates contributions from the entire deviation distribution, including its upper tail. This distinction is particularly relevant for the exponential, lognormal, and mixture distributions, for which asymmetry or component separation leads to a less uniform spread of absolute deviations.

6.3.4 Overall interpretation

The three tables jointly illustrate the usefulness of the functional deviation framework for comparing different distributional structures. The first table shows that the empirical mean absolute deviation about the median does not exceed the corresponding deviation about the mean for any of the six simulated distributions. The differences are negligible to the reported precision for the symmetric distributions, but are more pronounced for the exponential and lognormal distributions. The Gaussian mixture falls between these cases, with a relatively small difference between the two mean absolute deviations.

The second table demonstrates that quantile-based deviations provide information about the distribution of absolute distances that cannot be summarized by the mean alone. The bounded uniform distribution has the smallest upper quantile, whereas the heavy-tailed Student-$t(3)$ and the right-skewed lognormal distribution exhibit relatively large upper quantiles. The Gaussian mixture displays large deviations across several quantile levels, consistent with the separation between its components. These results emphasize the importance of examining several quantiles when the objective is to distinguish central dispersion from upper-tail behaviour.

The third table further shows that the median of absolute deviations about the mean provides a complementary summary to the expectation of absolute deviations. Because it summarizes the central deviation rather than the average deviation, it responds differently to skewness, heavy tails, and mixture structure. Together, the expectation-based, median-based, and quantile-based summaries provide a more comprehensive description of dispersion than any one measure considered in isolation.

Finally, the simulation should be interpreted as an empirical illustration of the theoretical properties rather than as an independent proof of them. The large sample size reduces random fluctuations, but the results remain sample-dependent and are reported to a limited number of decimal places. The observed inequalities agree with the theoretical relationships for the distributions considered, while the differences between the numerical summaries demonstrate

how the choice of reference functional and deviation-summary functional influences the empirical assessment of dispersion.

7 Graphical presentation using PGFPlots

7.1 Mean absolute deviations across distributions

Figure 1 compares the mean absolute deviations about the median and the mean. The plotted values are those reported by the simulation.

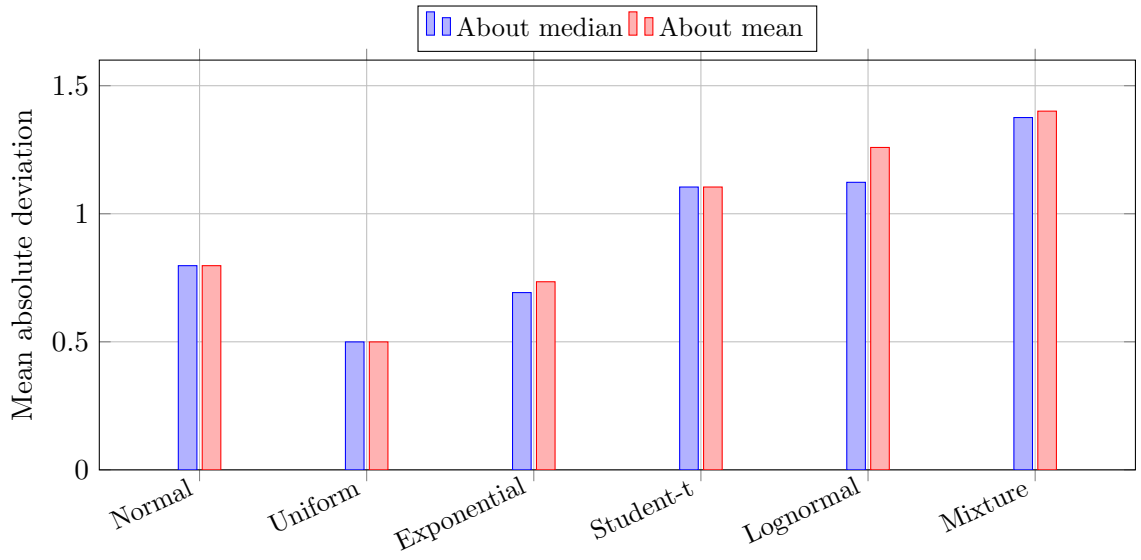

Figure 1: Comparison of mean absolute deviations about the median and the mean.

7.2 Mean absolute deviation versus standard deviation

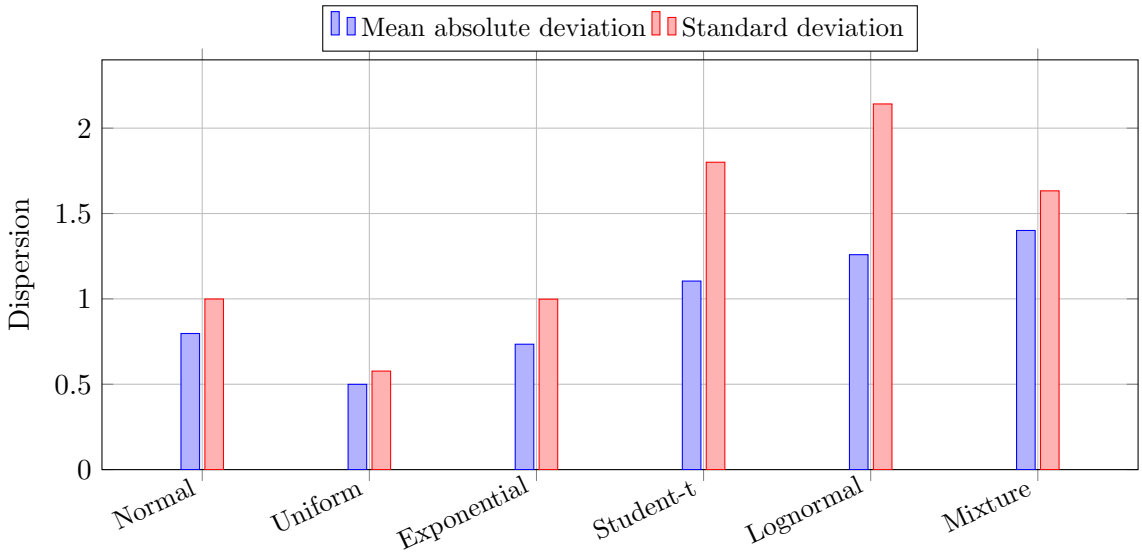

Figure 2: Mean absolute deviation about the mean compared with the standard deviation.

7.3 Quantiles of absolute deviations

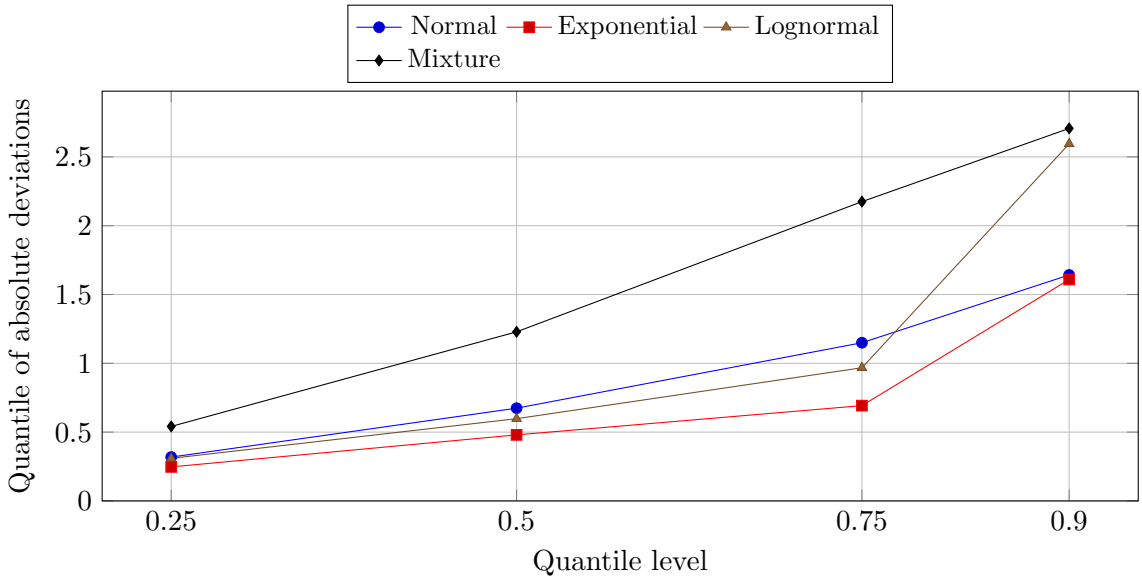

Figure 3: Quantiles of absolute deviations about the median for selected distributions.

8 Interpretation and discussion

The simulation results provide an empirical illustration of how functional deviations behave under different distributional configurations. By comparing expectation-based, median-based, and quantile-based summaries, the analysis demonstrates that the numerical assessment of dispersion depends not only on the underlying distribution but also on the choice of reference functional and the functional used to summarize the resulting absolute deviations. The findings are discussed in relation to the theoretical inequalities established in the preceding section.

8.1 Comparison of the two mean absolute deviations

The first important finding concerns the comparison between the mean absolute deviation about the median and the mean absolute deviation about the mean. The empirical results satisfy

$$\widehat{D}_{\mathbb{E}|\mathrm{Med}} \leq \widehat{D}_{\mathbb{E}|\mathbb{E}}$$

for all six simulated distributions. The numerical values show that the magnitude of the difference depends on the distributional shape and on the relative positions of the mean and median.

For the normal, uniform, and Student's $t(3)$ distributions, the two mean absolute deviations agree to four decimal places. This behaviour is consistent with the symmetry of these distributions about their population locations. For a symmetric distribution with a finite mean and a well-defined central median, the mean and median coincide. Consequently, both absolute-deviation summaries use the same population reference location, and their population values are equal. The observed agreement therefore illustrates the equality case of Theorem 5.2.

The normal distribution provides a particularly useful benchmark. Its mean and median are both zero, and its standard deviation is one. The two empirical mean absolute deviations are approximately 0.7973, which is close to the population value $\sqrt{2/\pi} \approx 0.7979$. Thus, the simulation reproduces the expected behaviour of a symmetric light-tailed distribution. The agreement between the two absolute-deviation measures should not, however, be interpreted as a general property of all distributions. It follows from the coincidence of the population reference locations in this setting.

A similar pattern is observed for the uniform distribution on $(-1, 1)$. Both mean absolute deviations are approximately 0.4996, close to their population value of 0.5. The uniform distribution has bounded support and is symmetric about zero, so its mean and median coincide. In addition, the bounded support restricts the maximum possible distance from the central location. The agreement of the two deviations is therefore accompanied by a relatively small overall dispersion. This example illustrates that the equality of mean-centered and median-centered deviations can occur in a distribution whose shape differs substantially from that of the normal distribution.

For Student's $t(3)$ distribution, the two mean absolute deviations are both approximately 1.1044. This agreement is again consistent with symmetry. However, the absolute scale of the deviations is substantially larger than that observed for the standard normal distribution. The difference illustrates that equality between two functional deviations does not imply equality of their magnitudes across distributions. The shape and tail behaviour of the underlying distribution continue to determine the size of the deviations, even when the mean and median coincide.

A more pronounced difference is observed for the exponential distribution. The mean absolute deviation about the median is 0.6922, while the corresponding deviation about the mean is 0.7345. The absolute difference is

$$0.7345 - 0.6922 = 0.0423.$$

Relative to the mean-centered deviation, this represents a reduction of approximately 5.76% when the median is used as the reference location. The ratio of the two quantities is 0.9424, indicating that the median-centered deviation is approximately 94.24% of the mean-centered deviation.

This difference is associated with the positive skewness of the exponential distribution. For an exponential distribution with rate one, the population mean is 1, whereas the population median is $\log 2 \approx 0.6931$. The mean is consequently located farther into the right-hand side of the distribution than the median. Relatively large observations in the right tail influence the mean more strongly than the median, producing a difference between the two reference locations. The median, however, minimizes the expected absolute loss over all possible reference values. Hence, the smaller median-centered mean absolute deviation is not merely an empirical regularity: it is consistent with the minimization property established theoretically.

The lognormal distribution exhibits the largest difference between the two mean absolute deviations among the six simulated distributions. The median-centered deviation is 1.1228, whereas the mean-centered deviation is 1.2589. Their absolute difference is

$$1.2589 - 1.1228 = 0.1361.$$

The reduction relative to the mean-centered deviation is approximately 10.81%, and the ratio of the two deviations is 0.8919. Thus, the median-centered deviation is approximately 89.19% of its mean-centered counterpart.

The larger difference for the lognormal distribution is consistent with its pronounced right skewness. For a lognormal random variable with parameters $(0, 1)$, the population median is 1, whereas the population mean is $e^{1/2} \approx 1.6487$. The separation between these two locations is greater than that for the exponential distribution considered in this simulation. Large positive observations also contribute substantially to the mean and to the average absolute deviation. The median provides a more central reference location in terms of probability mass, and its minimizing property ensures that the expected absolute deviation about it is no larger than the expected absolute deviation about the mean.

The Gaussian mixture exhibits a smaller but still noticeable difference. The mean absolute deviation about the median is 1.3757, compared with 1.4007 about the mean, giving an absolute difference of 0.0250. This corresponds to a reduction of approximately 1.79% relative to the

mean-centered deviation. The relatively small difference does not indicate that the mixture has low dispersion; both deviations are larger than those for the normal, uniform, and exponential distributions. Rather, it indicates that the two reference choices produce similar average absolute deviations for this particular mixture configuration.

The mixture combines a dominant normal component centered at zero with a secondary component centered at three. This separation increases the overall dispersion and may produce a multimodal density. Nevertheless, the relative difference between the two mean absolute deviations depends on the overall distribution of observations around the mean and median, not simply on the number of modes. Therefore, the observed ratio of 0.9821 should be interpreted as a feature of this particular mixture rather than as a universal characteristic of multimodal distributions.

Taken together, these findings support two conclusions. First, the median-centered mean absolute deviation is no greater than the mean-centered counterpart in all six samples, in agreement with Theorem 5.2. Second, the magnitude of the difference depends on the distributional configuration. The difference is negligible to the reported precision for the three symmetric distributions, more pronounced for the exponential and lognormal distributions, and comparatively small for the Gaussian mixture. The results thus illustrate the importance of distinguishing the choice of reference location from the overall scale of dispersion.

8.2 Comparison with the standard deviation

The second comparison concerns the mean absolute deviation about the mean and the standard deviation. The reported simulation output indicates that the mean absolute deviation about the mean is smaller than the estimated standard deviation in every distribution. The theoretical basis for this relationship is the inequality

$$D_{\mathbb{E}|\mathbb{E}}(X) = \mathbb{E}|X - \mu| \leq \sqrt{\mathbb{E}[(X-\mu)^2]} = \mathrm{SD}(X),$$

which follows directly from the Cauchy–Schwarz inequality whenever the second moment is finite.

To examine the relative magnitudes of the two measures, the simulation output also reports the ratio

$$R_{\mathrm{SD}} = \frac{\widehat{D}_{\mathbb{E}|\mathbb{E}}}{\widehat{\sigma}}.$$

The ratios range from approximately 0.5878 for the lognormal distribution to 0.8658 for the uniform distribution. A ratio closer to one indicates that the mean absolute deviation is relatively close to the standard deviation, whereas a smaller ratio indicates a larger difference between the two summaries relative to the standard deviation.

For the uniform distribution, the reported ratio is approximately 0.8658. This relatively high value is consistent with its bounded support and the comparatively regular distribution of absolute deviations. Because observations cannot fall outside $(-1, 1)$, the distribution cannot generate arbitrarily large deviations. The standard deviation is therefore not disproportionately enlarged by a small number of extreme observations. The mean absolute deviation and standard deviation remain distinct because they summarize deviations differently, but their relative magnitudes are comparatively close in this example.

The standard normal distribution has a ratio of approximately 0.7979, consistent with the population identity

$$\frac{\mathbb{E}|X|}{\mathrm{SD}(X)} = \sqrt{\frac{2}{\pi}}, \qquad X \sim N(0,1).$$

The mean absolute deviation is smaller than the standard deviation because the latter is based on squared deviations. Under a normal distribution, the relationship between the two measures

is stable and determined by the distribution's shape. The simulation value is close to the corresponding population ratio, as expected from the large sample size.

For the exponential distribution, the ratio is approximately 0.7358. This lower value reflects the asymmetric shape of the distribution and the influence of its right tail on quadratic dispersion. Although both the mean absolute deviation and standard deviation are finite, the standard deviation places relatively greater weight on large deviations because the deviations are squared before being averaged. The resulting ratio therefore illustrates how a skewed distribution can produce a greater relative separation between first-order and second-order dispersion summaries.

For Student's $t(3)$ distribution, the ratio is approximately 0.6135. This value is appreciably smaller than the normal ratio, reflecting the heavy-tailed nature of the t_3 distribution. The population variance of this distribution is finite, but its fourth moment is infinite. Its heavy tails increase the contribution of large observations to the sample variance and standard deviation, whereas the mean absolute deviation increases linearly with their magnitude. The observed ratio is therefore consistent with the greater sensitivity of quadratic dispersion measures to extreme deviations.

The lognormal distribution has the smallest reported ratio, approximately 0.5878. Its mean absolute deviation about the mean is 1.2589, whereas its estimated standard deviation is 2.1418. The relatively large difference between these values is consistent with the distribution's pronounced right skewness and extended upper tail. Large positive observations contribute strongly to the second moment and can substantially increase the standard deviation. The mean absolute deviation also reflects these observations, but its linear dependence on absolute deviations makes it less sensitive to very large values than the quadratic dispersion measure. The low ratio therefore illustrates the different responses of the two measures to asymmetric distributions with substantial upper-tail variation.

For the Gaussian mixture, the ratio is approximately 0.8578. The mixture has a relatively large standard deviation because the two component means are separated, creating between-component variation in addition to within-component variation. At the same time, the mean absolute deviation remains relatively close to the standard deviation compared with the exponential, Student-$t(3)$, and lognormal examples. This demonstrates that a distribution can have a large absolute scale of dispersion while retaining a relatively high ratio between mean absolute deviation and standard deviation. The ratio depends on the complete distributional shape and cannot be inferred solely from the number of modes or from the magnitude of the standard deviation.

Overall, the comparison demonstrates that the standard deviation and mean absolute deviation provide related but nonidentical descriptions of dispersion. The standard deviation emphasizes large deviations through quadratic weighting, whereas the mean absolute deviation summarizes their average absolute magnitude. The observed ratios vary across the six distributions, showing that no single ratio should be regarded as universal. In particular, skewness and heavy tails can lead to a substantial difference between the two measures. The findings are consistent with the theoretical inequality, while also illustrating why both summaries can be informative in applied analyses.

8.3 Quantile-based deviations

Quantile-based functional deviations provide a distribution-sensitive description of dispersion at different probability levels. Unlike the mean absolute deviation, which summarizes all absolute deviations through their average, a quantile-based deviation identifies the distance below which a specified proportion of the observations lies. For example, the 25th percentile describes the lower quarter of the absolute-deviation distribution, the median describes its central position, and the 75th and 90th percentiles characterize progressively larger deviations.

For each of the six distributions, the reported quantiles satisfy

$$\widehat{Q}_{0.25} \leq \widehat{Q}_{0.50} \leq \widehat{Q}_{0.75} \leq \widehat{Q}_{0.90}.$$

This ordering is a mathematical consequence of the nondecreasing quantile function and does not depend on normality, symmetry, or the existence of moments. The simulation therefore provides a computational check of the quantile calculations, while the validity of the ordering itself follows from the definition of quantiles.

The uniform distribution provides the clearest bounded-support example. Its absolute deviations about the population median zero are bounded by one, and the corresponding population quantile function is $Q(p) = p$ for $0 < p < 1$. The empirical quantiles in the simulation are close to the values 0.25, 0.50, 0.75, and 0.90. The small differences arise from finite-sample variation and the use of the sample median rather than the known population median. The uniformly spaced quantiles reflect the constant density of the absolute-deviation distribution over its support. This behaviour contrasts with distributions whose absolute-deviation quantiles increase more rapidly in the upper part of the distribution.

For the standard normal distribution, the absolute deviation quantiles increase from 0.3186 at the 25th percentile to 0.6739 at the median, 1.1495 at the 75th percentile, and 1.642 at the 90th percentile. The increasing values indicate that a progressively larger distance from the median is needed to encompass higher proportions of the observations. The upper quantiles are substantially larger than the lower quantiles, but the pattern remains consistent with a light-tailed continuous distribution. The normal results also provide a benchmark against which the effects of heavier tails and asymmetry can be assessed.

The exponential distribution has absolute-deviation quantiles of 0.2474, 0.4801, 0.6926, and 1.609 at the four reported probability levels. The relatively large increase from the 75th to the 90th percentile is noteworthy. It indicates that the upper portion of the absolute-deviation distribution is more spread out than would be suggested by the lower and central quantiles alone. This behaviour is consistent with the right-skewed structure of the exponential distribution. A substantial proportion of observations is relatively close to the median, but the right tail allows some observations to lie much farther away. Thus, the upper quantiles reveal information about asymmetry that is not fully captured by the median absolute deviation.

Student's $t(3)$ distribution exhibits a 90th-percentile absolute deviation of 2.361, compared with 1.642 for the normal distribution. Its 75th percentile is also larger, at 1.4234, compared with 1.1495 for the normal distribution. These differences are consistent with the heavier tails of the t_3 distribution. The increased upper quantiles indicate that relatively large absolute deviations occur more frequently than under the standard normal model. Consequently, quantile-based summaries can help distinguish heavy-tailed distributions even when the two distributions are both symmetric about their central locations.

The lognormal distribution has a 90th-percentile absolute deviation of 2.595, which is larger than the corresponding normal, uniform, and exponential values in the reported table. Its 75th percentile is 0.9677, while its median absolute deviation is 0.5977. The substantial increase at the 90th percentile indicates that the upper tail contributes disproportionately large absolute deviations. This is consistent with the right-skewed lognormal distribution, where a relatively small proportion of observations can be much larger than the central values. The result reinforces the importance of reporting upper quantiles alongside central measures of dispersion, especially when the distribution is asymmetric.

The Gaussian mixture produces the largest 25th-, 50th-, and 75th-percentile absolute deviations among the distributions considered, with values of 0.5413, 1.2287, and 2.1748, respectively. Its 90th percentile is 2.707. The large central quantiles indicate that the dispersion is not confined to a small number of extreme observations: a substantial proportion of the sample lies relatively far from the overall median. This pattern is consistent with the separation between the two Gaussian components. Observations from the component centered near three can be

far from the overall median, while observations from the component centered near zero occupy a different region of the support. The combined distribution therefore produces a broad distribution of absolute distances. Nevertheless, the numerical quantiles alone do not establish bimodality; the modal structure of a mixture must be assessed from its density or another suitable diagnostic.

The comparison of quantiles also highlights why a single measure of dispersion may be insufficient. Two distributions can have similar median absolute deviations but substantially different upper quantiles. Such a difference would indicate that their central deviations are similar while their upper-tail behaviour differs. Conversely, large differences in central quantiles, as observed for the Gaussian mixture, indicate broader dispersion throughout a substantial portion of the distribution. The quantile-based functional deviations therefore provide a flexible way of distinguishing central spread from upper-tail spread.

8.4 Verification of the theoretical inequalities

The R program reports that all five implemented sample checks are satisfied for each of the six simulated distributions. These checks correspond to the median–expectation bound, the comparison between the two mean absolute deviations, the mean absolute deviation–standard deviation inequality, monotonicity of quantile-based deviations, and the quantile–expectation bound.

The first check concerns the median of the absolute deviations about the mean:

$$\mathrm{Med}(|X-\mu|) \le 2\mathbb{E}|X-\mu|.$$

In the simulation, the population mean μ is replaced by the sample mean $\bar{X}$, and the population expectation and median are replaced by their empirical counterparts. The resulting check is

$$\widehat{D}_{\mathrm{Med}|\mathbb{E}} \le 2\widehat{D}_{\mathbb{E}|\mathbb{E}}.$$

The inequality follows from the general fact that the median of a nonnegative random variable cannot exceed twice its expectation. The simulation shows that the empirical median of the absolute deviations is within this bound for all six distributions. This check is useful for verifying the consistency of the computed median and mean of the absolute-deviation sample.

The second check concerns the median-minimization property:

$$\mathbb{E}|X-m| \le \mathbb{E}|X-\mu|.$$

Its empirical counterpart is

$$\widehat{D}_{\mathbb{E}|\mathrm{Med}} \le \widehat{D}_{\mathbb{E}|\mathbb{E}}.$$

All six samples satisfy this comparison. This is consistent with the fact that the sample median minimizes the sum of absolute deviations from the observed data, while the sample mean is the reference value associated with the conventional mean-centered absolute deviation. The result holds for the empirical distribution, just as the population inequality holds for the underlying distribution when the first absolute moment is finite. The simulation therefore illustrates both the theoretical minimizing property and its sample counterpart.

The third check compares the mean absolute deviation about the mean with the standard deviation:

$$\mathbb{E}|X-\mu| \le \mathrm{SD}(X).$$

For the empirical verification, the root mean squared deviation with denominator N is the most direct analogue:

$$\widehat{D}_{\mathbb{E}|\mathbb{E}} \le \left\{ \frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2 \right\}^{1/2}.$$

This follows directly from the finite-sample Cauchy–Schwarz inequality. The conventional sample standard deviation, which uses denominator $N-1$, is also reported in the results table, but it should be distinguished from this denominator-N quantity when describing the exact empirical inequality. The observed numerical pattern is consistent with the theoretical relationship between absolute and quadratic dispersion.

The fourth check verifies the monotonicity of quantile-based deviations:

$$Q_{Y_T}(p_1) \leq Q_{Y_T}(p_2), \qquad 0 < p_1 < p_2 < 1,$$

where $Y_T = |X - T(X)|$. The empirical quantiles are ordered from the 25th percentile through the 90th percentile in every simulated sample. This is expected because the quantile function is nondecreasing by definition. The result does not depend on any distributional assumptions and remains valid for discrete distributions, continuous distributions, and mixtures. In the present setting, the check confirms that the quantile calculations have been implemented consistently.

The fifth check concerns the quantile–expectation bound:

$$Q_{Y_T}(p) \leq \frac{\mathbb{E}[Y_T]}{1-p}.$$

The corresponding empirical comparison replaces the population quantile and expectation by their sample estimates. This bound becomes less restrictive as p approaches one because the denominator $1-p$ decreases. For example, at $p = 0.90$, the bound is ten times the expectation of the nonnegative deviation variable, whereas at $p = 0.25$, it is only four-thirds of that expectation. The bound therefore provides a distribution-free upper limit for a quantile in terms of the mean absolute deviation. The reported checks show that the computed empirical quantiles satisfy the corresponding inequalities for the samples considered.

The satisfaction of these checks across all six distributions is consistent with the mathematical results established earlier. However, the checks have different logical roles. The median–mean, median–expectation, and quantile–expectation comparisons assess sample analogues of inequalities proved for population functionals. Quantile monotonicity, by contrast, is a defining property of quantiles, and the mean-absolute-deviation comparison follows from the median's empirical minimization of absolute loss. Thus, the simulation provides a numerical verification of the calculations and an illustration of the theoretical relationships, rather than an independent proof of the population inequalities.

8.5 Overall implications

The simulation results demonstrate that the proposed functional deviation framework can describe dispersion from several complementary perspectives. The choice of reference functional determines the point about which deviations are measured, while the choice of summary functional determines how those deviations are aggregated. These two choices should be treated separately because they answer different questions about the distribution.

The mean-centered and median-centered mean absolute deviations are equal, to the reported precision, for the symmetric normal, uniform, and Student-$t(3)$ distributions. Their difference is larger for the exponential and lognormal distributions, for which the mean and median are separated by positive skewness. The Gaussian mixture exhibits a smaller difference between the two mean absolute deviations despite its larger overall dispersion. These findings show that the relative effect of changing the reference location depends on the complete distributional structure rather than on dispersion magnitude alone.

The comparison with standard deviation reveals a different aspect of distributional behaviour. The standard deviation gives greater weight to large deviations through squaring, while the mean absolute deviation measures their average absolute magnitude. This distinction

becomes especially important for the lognormal and Student-$t(3)$ distributions, where skewness and heavy tails respectively contribute to larger quadratic dispersion relative to absolute dispersion.

Finally, the quantile-based deviations demonstrate that dispersion is not adequately described by a single central summary in every setting. The upper quantiles distinguish distributions with bounded support from those with heavy tails, pronounced skewness, or separated mixture components. Considering several quantile levels therefore provides additional information about the distribution of distances from a reference location.

The overall findings support the use of multiple functional deviations as complementary descriptive measures. The expectation-based deviation quantifies average absolute dispersion, the median-based deviation describes its central magnitude, and quantile-based deviations characterize dispersion at selected probability levels. The theoretical inequalities establish useful relationships among these summaries, while the simulation illustrates how their numerical values respond to different distributional shapes. The conclusions are limited to the six distributions and the simulation configuration considered; further experiments with varying sample sizes, mixture weights, component separation, and tail parameters would be required to assess the finite-sample behaviour more broadly.

9 Limitations and Further Investigation

The numerical investigation provides an initial assessment of the relationships among functional deviations defined relative to the mean, median, and other reference functionals. However, the results should be interpreted in light of the sampling design, the tail behaviour of the underlying distributions, the choice of reference functional, and the numerical procedures used to estimate the corresponding deviations. The following discussion identifies the principal limitations of the present investigation and develops a reproducible Monte Carlo framework for examining the stability, finite-sample behaviour, and distributional sensitivity of the proposed measures.

9.1 Limitations of the single-sample numerical investigation

The initial simulation considered six distributions: the standard normal distribution, the uniform distribution on $(-1, 1)$, the exponential distribution with unit rate, the Student's t distribution with three degrees of freedom, the standard lognormal distribution, and a two-component Gaussian mixture. A single sample of size $N = 300000$ was generated from each distribution. The resulting empirical deviations provide a useful numerical illustration of the proposed relationships. Nevertheless, a single generated sample does not provide a direct estimate of the sampling variability of the deviation estimators.

For example, the estimated mean absolute deviations about the sample median and the sample mean were, respectively, 0.7973 and 0.7973 for the standard normal distribution, 0.4996 and 0.4996 for the uniform distribution, 0.6922 and 0.7345 for the exponential distribution, 1.1044 and 1.1044 for the Student's t_3 distribution, 1.1228 and 1.2589 for the lognormal distribution, and 1.3757 and 1.4007 for the Gaussian mixture. These values illustrate the expected equality or near-equality for approximately symmetric distributions and the advantage of median-centred absolute deviations for asymmetric distributions.

The ratio of the estimated mean absolute deviation about the median to that about the mean was approximately 0.9424 for the exponential distribution, 0.8919 for the lognormal distribution, and 0.9821 for the Gaussian mixture. Thus, in the generated samples, the corresponding median-centred deviations were approximately 5.76%, 10.81%, and 1.79% smaller, respectively, than the mean-centred deviations. These percentages describe the observed samples; they should not be interpreted as estimates of Monte Carlo bias or as universal improvements associated with median centering.

The estimated standard deviations were approximately 0.9993, 0.5771, 0.9983, 1.8004, 2.1418, and 1.6328, in the order listed above. The corresponding ratios of mean absolute deviation about the mean to standard deviation were 0.7979, 0.8658, 0.7358, 0.6135, 0.5878, and 0.8578. The substantial reduction in these ratios for the Student's t_3 and lognormal distributions illustrates that a deviation based on absolute loss and a deviation based on squared loss can respond differently to distributional tails.

These numerical differences motivate repeated sampling. In particular, it remains necessary to determine whether the observed differences persist across independent samples, how rapidly the corresponding estimators stabilize as the sample size increases, and whether the sampling variability depends strongly on skewness, tail weight, or multimodality.

9.2 Repeated Monte Carlo experiments and finite-sample behaviour

A natural extension is to generate independent samples at $n \in \{50, 100, 250, 500, 1000\}$ and repeat the experiment B times for each combination of distribution and sample size. A value such as $B = 2000$ provides a practical starting point, subject to computational resources. Let $X_{1b}, \ldots, X_{nb}$ denote the observations in replication b, and let $\widehat{\mu}_b$ and $\widehat{m}_b$ denote the sample mean and sample median, respectively.

The principal quantities of interest may be defined as

$$\widehat{D}_{\mu,b} = \frac{1}{n}\sum_{i=1}^{n} |X_{ib} - \widehat{\mu}_b|, \qquad \widehat{D}_{m,b} = \frac{1}{n}\sum_{i=1}^{n} |X_{ib} - \widehat{m}_b|.$$

The first quantity measures the average absolute deviation from the sample mean, whereas the second measures the average absolute deviation from the sample median. Their difference and ratio are

$$\widehat{\Delta}_b = \widehat{D}_{\mu,b} - \widehat{D}_{m,b}, \qquad \widehat{R}_b = \frac{\widehat{D}_{m,b}}{\widehat{D}_{\mu,b}},$$

provided that $\widehat{D}_{\mu,b} > 0$.

Since the sample median minimizes the sum of absolute deviations, $\widehat{\Delta}_b$ is nonnegative, apart from possible floating-point error. The ratio $\widehat{R}_b$ therefore provides a normalized measure of the reduction in absolute deviation achieved by median centering. A ratio close to one indicates that the two reference functionals produce similar mean absolute deviations, whereas a smaller ratio indicates a more substantial reduction from median centering.

For each distribution and sample size, the Monte Carlo summaries should include the mean and standard deviation of $\widehat{D}_{\mu,b}$ and $\widehat{D}_{m,b}$, the empirical distribution of $\widehat{\Delta}_b$, and the mean, standard deviation, and selected quantiles of $\widehat{R}_b$. The Monte Carlo standard error of an estimated expectation $B^{-1}\sum_{b=1}^{B} A_b$ is

$$\mathrm{MCSE}(\overline{A}) = \frac{s_A}{\sqrt{B}},$$

where s_A is the sample standard deviation of $A_1, \ldots, A_B$. This quantity measures the simulation uncertainty in the estimated Monte Carlo average; it is distinct from the sampling standard deviation of an estimator within a single data sample.

A useful distinction should be maintained between bias and variability. If D denotes the population value of a particular deviation functional and $\widehat{D}_b$ its estimate in replication b, the Monte Carlo bias estimate is

$$\widehat{\mathrm{Bias}}(\widehat{D}) = \frac{1}{B}\sum_{b=1}^{B} \widehat{D}_b - D,$$

whereas the Monte Carlo standard deviation is

$$\widehat{\mathrm{SD}}_{\mathrm{MC}}(\widehat{D}) = \left\{ \frac{1}{B-1}\sum_{b=1}^{B} \left(\widehat{D}_b - \overline{\widehat{D}}\right)^2 \right\}^{1/2}.$$

Population benchmarks should be obtained analytically where possible or approximated independently using a much larger sample or numerical integration. The same simulated observations should not be used uncritically both to define the benchmark and to assess the estimator against that benchmark.

The expected patterns differ across the six distributions. For the standard normal distribution, the population mean and median coincide, and the population mean absolute deviation about the common centre equals $\sqrt{2/\pi} \approx 0.7979$. For the uniform distribution on $(-1, 1)$, the corresponding value is $1/2$. These analytical benchmarks allow direct assessment of finite-sample bias. For the exponential distribution, the population mean is 1 and the population median is $\log 2$. Its mean absolute deviation about the mean is $2/e \approx 0.7358$, while the mean absolute deviation about the median is approximately 0.6931. Consequently, the population ratio is approximately 0.9425.

For the Student's t_3 distribution, the population mean and median are both zero, but the distribution has heavier tails than the normal distribution. For the lognormal distribution, substantial right-skewness can produce a larger difference between mean-centred and median-centred absolute deviations. The Gaussian mixture introduces a different challenge: the mean and median need not represent the same central region when the component means and mixing weights differ. These characteristics make the six distributions complementary rather than interchangeable simulation settings.

9.3 Numerical verification of empirical inequalities

The theoretical inequalities should be distinguished from their empirical counterparts. In particular, the inequality

$$\mathbb{E}|X - \mu| \leq \sqrt{\mathrm{Var}(X)}$$

is a population statement under a finite second-moment assumption. Its empirical analogue is most directly verified by comparing the empirical mean absolute deviation about the sample mean with the empirical root mean square deviation using the same denominator n:

$$\frac{1}{n}\sum_{i=1}^{n} |X_i - \overline{X}| \leq \left\{ \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2 \right\}^{1/2}.$$

The inequality follows from the Cauchy–Schwarz inequality and holds for every finite sample. If the conventional sample standard deviation with denominator $n - 1$ is used instead, it is at least as large as the root mean square deviation with denominator n; the comparison therefore remains valid.

A repeated-sampling experiment can record the frequency of numerical violations of the empirical inequalities. These frequencies must, however, be interpreted correctly. The empirical minimization property of the sample median and the empirical Cauchy–Schwarz inequality hold by construction. Accordingly, a nonzero violation frequency for these two checks should prompt an examination of numerical tolerances, data handling, or implementation errors rather than a substantive statistical interpretation.

Quantile inequalities require additional attention to the quantile convention. Let $Y_i = |X_i - \widehat{m}|$ and let $\widehat{Q}_Y(p)$ be the empirical quantile of order p. The empirical mean of Y_i is $\widehat{D}_m$, and the Markov-type bound

$$\widehat{Q}_Y(p) \leq \frac{\widehat{D}_m}{1-p}, \qquad 0 < p < 1,$$

holds for the empirical distribution when the lower generalized-inverse empirical quantile is used. Since $\widehat{D}_m \leq \widehat{D}_\mu$, the further bound

$$\widehat{Q}_Y(p) \leq \frac{\widehat{D}_\mu}{1-p}$$

also follows. For example, at $p = 0.75$, the upper bound is $4\widehat{D}_{\mu}$. Although this bound is valid, it may be conservative; the simulation should report the ratio of the observed quantile to its upper bound rather than merely recording whether the inequality holds.

The empirical median of Y_i is similarly bounded by $2\widehat{D}_m$ under the lower generalized-inverse convention, and therefore by $2\widehat{D}_{\mu}$. It is important to distinguish these deterministic sample inequalities from statements about estimator bias, consistency, or the probability of a theoretical population inequality being true. The repeated-sampling experiment evaluates the behaviour of estimated quantities; it does not establish a population theorem by simulation.

9.4 Quantile-based deviations and tail sensitivity

Mean absolute deviation compresses the distribution of absolute deviations into a single expectation. Quantile-based deviations provide complementary information about the distribution of $|X - T(F)|$. In the proposed investigation, the empirical quantiles at levels 0.25, 0.50, 0.75, and 0.90 may be computed for absolute deviations about the sample median. The resulting vector

$$\left(\widehat{Q}_{0.25}, \widehat{Q}_{0.50}, \widehat{Q}_{0.75}, \widehat{Q}_{0.90}\right)$$

summarizes the central spread and upper tail of the absolute-deviation distribution.

The earlier numerical experiment produced the following quantiles, rounded to four decimal places. For the standard normal distribution, the values were 0.3186, 0.6739, 1.1495, and 1.6420. For the uniform distribution, they were 0.2494, 0.4996, 0.7494, and 0.9000. The corresponding values for the exponential distribution were 0.2474, 0.4801, 0.6926, and 1.6090; for the Student's t_3 distribution, they were 0.3492, 0.7662, 1.4234, and 2.3610; for the lognormal distribution, they were 0.3095, 0.5977, 0.9677, and 2.5950; and for the Gaussian mixture, they were 0.5413, 1.2287, 2.1748, and 2.7070.

The upper quantiles are particularly informative. The 0.90 quantile for the lognormal sample was 2.5950, compared with 1.6090 for the exponential sample and 2.3610 for the Student's t_3 sample. This pattern illustrates that a distribution may have a moderate central deviation while retaining a relatively extended upper tail. The Gaussian mixture also exhibited a large central quantile, consistent with substantial dispersion across mixture components. These observations are descriptive of the simulated samples and should not be taken as universal rankings of the corresponding population quantiles.

In the repeated experiment, the variability of each quantile estimator should be examined separately. In particular, the 0.90 quantile will generally be estimated less precisely than a central quantile because it is determined by fewer upper-tail observations. For the Student's t_3 and lognormal distributions, this distinction is especially relevant. A complete analysis should therefore report Monte Carlo standard deviations or interquartile ranges for the quantile estimators, in addition to their average values.

9.5 Mode-based deviations and density-estimation uncertainty

Mode-based deviations require a separate methodological treatment because the mode is not generally characterized by a minimization property analogous to that of the median under absolute loss. For a continuous distribution with density f, a mode is a point at which f attains its global maximum. A density may have several global maxima, several local maxima, or a broad region of nearly equal density. Consequently, the mode may be nonunique or sensitive to small perturbations in the data.

This difficulty is particularly relevant for Gaussian mixtures. Depending on the component means, component variances, and mixing weights, a Gaussian mixture may be unimodal or multimodal. The presence of two mixture components does not, by itself, guarantee two modes. Furthermore, the density of the absolute deviation $|X - \widehat{m}|$ may exhibit a shape different from

that of the original observations. Its mode should therefore be estimated separately rather than inferred from the mode of the original distribution.

A reproducible mode-based simulation should specify the density estimator, bandwidth-selection method, evaluation grid, and mode-selection convention. One possible approach is to use a Gaussian kernel density estimator with bandwidth selected by the normal-reference rule implemented in `bw.nrd0` in R. The estimated global mode can then be approximated by the grid point at which the estimated density is maximized. To investigate multimodality, local maxima can also be recorded, with a prespecified rule for distinguishing meaningful modes from minor numerical fluctuations.

Bandwidth sensitivity should be assessed by repeating the analysis with, for example, $0.75h$, h, and $1.25h$, where h is the default bandwidth. These multipliers constitute a sensitivity design rather than an assertion that any one bandwidth is optimal. A mode estimate that changes substantially across these bandwidths should be reported as unstable. For a multimodal density, the analysis should state whether the global mode, a specified local mode, or all detected modes are being used as the reference functional.

Unlike the mean and median, a mode-based reference does not automatically inherit the inequality

$$\mathbb{E}|X - \mathrm{Med}(F)| \leq \mathbb{E}|X - \mathrm{Mode}(F)|.$$

No universal inequality of this form should be asserted without additional assumptions. The mode-based investigation should consequently focus on estimation stability, sensitivity to bandwidth, the frequency of nonunique estimated modes, and the effect of the mode-selection convention on the resulting deviations.

9.6 Heavy tails and moment assumptions

The Student's t distribution with three degrees of freedom provides an important example of a heavy-tailed distribution with finite variance but infinite fourth moment. More generally, for a Student's t_ν distribution, the absolute moment of order r is finite only when $r < \nu$. Thus, for $\nu = 3$, the first and second moments exist, whereas the third and fourth absolute moments do not.

This distinction matters for interpreting simulation variability. The inequality

$$\mathbb{E}|X - \mu| \leq \sqrt{\mathrm{Var}(X)}$$

requires only a finite second moment. It therefore applies to the Student's t_3 distribution. However, arguments that rely on a finite fourth moment, such as certain higher-order variance approximations or standard error approximations for statistics involving squared observations, may fail or require different justification.

For $X \sim t_3$, the theoretical variance is 3 and the standard deviation is $\sqrt{3} \approx 1.7321$. The earlier single-sample estimate of the standard deviation was 1.8004. This discrepancy is a descriptive sample outcome, not evidence that the population variance is infinite. The finite variance assumption is satisfied, but large observations can produce appreciable sample-to-sample fluctuation. In particular, sample variance and standard deviation estimates may be much more sensitive to extreme observations than estimates based on absolute deviations.

A further study could compare t_3 with t_5 and the standard normal distribution to examine how the finite-sample behaviour changes as additional moments become available. Such an extension would help distinguish the consequences of heavy tails from those of skewness and multimodality. The lognormal and Gaussian mixture distributions would remain useful complementary examples because they introduce right-skewness and heterogeneous mixture structure, respectively.

9.7 Reproducible Monte Carlo implementation in R

The corresponding R implementation, as available in Subsection (A.2) in Appendix, provides a reproducible framework for the repeated-sampling experiment. It uses $B = 2000$ replications, the sample sizes $n = 50, 100, 250, 500, 1000$, and the six distributions used in the initial investigation. The code computes the mean absolute deviations about the sample mean and median, the standard deviation, selected quantiles of the absolute deviations about the median, and empirical inequality checks. It also reports Monte Carlo means, standard deviations, Monte Carlo standard errors, and empirical violation frequencies. The empirical root mean square deviation is calculated using denominator n, thereby matching the empirical Cauchy–Schwarz inequality exactly. The R quantile convention `type=1` is used for the quantile-based checks, corresponding to the lower generalized-inverse empirical quantile. A fixed random seed makes the experiment reproducible.

The code is designed to generate the repeated-sampling results; the resulting numerical summaries must be obtained by running it. In interpreting the output, a violation frequency of zero for the median-minimization and Cauchy–Schwarz checks is expected because these inequalities hold for every finite sample, up to numerical tolerance. The quantile checks should also be satisfied under the stated empirical-quantile convention. Nonzero violation frequencies should be investigated as potential implementation or numerical issues.

The most informative comparisons are likely to be the Monte Carlo standard deviations and quantile intervals for the estimated deviations and their ratios. The expected decrease in variability with increasing sample size should be assessed empirically rather than imposed as a conclusion, particularly for heavy-tailed distributions and high-order quantiles.

9.7.1 Output analysis & interpretation

The repeated Monte Carlo experiment was designed to investigate the finite-sample behaviour of mean absolute deviations defined relative to the sample mean and sample median, together with the stability of their ratio, the distribution of absolute deviations, and the empirical validity of the associated inequalities. Six distributions were considered: the standard normal distribution, the uniform distribution on $(-1, 1)$, the exponential distribution with unit rate, the Student's t distribution with three degrees of freedom, the standard lognormal distribution, and a two-component Gaussian mixture with mixing probability 0.30 for the component $N(3, 0.5^2)$ and probability 0.70 for the component $N(0, 1)$.

The sample sizes were $n \in \{50, 100, 250, 500, 1000\}$, with $B = 2000$ independent replications for each combination of distribution and sample size. Consequently, the experiment comprised $6 \times 5 \times 2000 = 60000$ simulated samples. For each replication, the mean absolute deviations about the sample mean and sample median, their difference and ratio, the root mean square deviation, the conventional sample standard deviation, and four quantiles of the absolute deviations about the sample median were calculated. All numerical summaries reported below were recalculated directly from the raw replication-level output, grouped by distribution and sample size.

9.7.2 Mean absolute deviations about the mean and median

The first comparison concerns the empirical quantities

$$\widehat{D}_{\text{mean}} = \frac{1}{n}\sum_{i=1}^{n} |X_i - \overline{X}|, \qquad \widehat{D}_{\text{median}} = \frac{1}{n}\sum_{i=1}^{n} |X_i - \widetilde{X}|,$$

where $\overline{X}$ and $\widetilde{X}$ denote the sample mean and sample median, respectively. Since the sample median minimizes the sum of absolute deviations, the difference

$$\widehat{\Delta} = \widehat{D}_{\text{mean}} - \widehat{D}_{\text{median}}$$

is nonnegative, apart from possible numerical rounding error. The ratio

$$\widehat{R} = \frac{\widehat{D}_{\text{median}}}{\widehat{D}_{\text{mean}}}$$

provides a normalized comparison of the two measures.

Table 4 reports the Monte Carlo means of these quantities at $n = 1000$, together with the Monte Carlo standard deviations of the two deviation estimators. The Monte Carlo means are averages over the 2000 replications; the reported standard deviations describe variation across replications rather than the dispersion of observations within a single sample.

Table 4: Monte Carlo results for the two mean absolute deviations at $n = 1000$ ($B = 2000$).

Distribution	$\overline{D}_{\text{mean}}$	$\overline{D}_{\text{median}}$	$\overline{\Delta}$	$\overline{R}$	$\text{MCSD}_{\text{mean}}$	$\text{MCSD}_{\text{median}}$
Normal$(0, 1)$	0.7969	0.7966	0.0002	0.9997	0.0189	0.0189
Uniform$(-1, 1)$	0.4995	0.4993	0.0002	0.9997	0.0090	0.0090
Exponential(1)	0.7340	0.6915	0.0425	0.9422	0.0285	0.0253
Student-t_3	1.1026	1.1019	0.0007	0.9994	0.0411	0.0409
Lognormal$(0, 1)$	1.2588	1.1223	0.1365	0.8920	0.0789	0.0616
Normal mixture	1.4001	1.3751	0.0250	0.9821	0.0279	0.0286

The results reveal three distinct patterns. First, the mean-centred and median-centred deviations are nearly identical for the standard normal and uniform distributions. Their Monte Carlo mean ratios are both approximately 0.9997, indicating that the two reference functionals yield essentially the same average absolute deviation at this sample size. The Student's t_3 distribution displays a similar pattern, with a ratio of 0.9994. These findings are consistent with the symmetry of the corresponding population distributions: for the normal, uniform, and Student's t_3 distributions, the population mean and median coincide.

Second, the distinction is more pronounced for asymmetric distributions. For the exponential distribution, the Monte Carlo mean of the median-centred deviation is 0.6915, compared with 0.7340 for the mean-centred deviation. The average ratio is 0.9422, corresponding to a reduction of approximately 5.78% relative to the mean-centred deviation. For the lognormal distribution, the corresponding estimates are 1.1223 and 1.2588, with an average ratio of 0.8920. The median-centred deviation is therefore approximately 10.80% smaller. This is the largest relative reduction among the six distributions.

Third, the Gaussian mixture produces an intermediate difference. Its mean-centred deviation is 1.4001, while its median-centred deviation is 1.3751. The average ratio of 0.9821 corresponds to a reduction of approximately 1.79%. Although the mixture is asymmetric for the specified mixing proportions and component locations, the difference between the two measures is smaller than that observed for the exponential and lognormal distributions.

These results illustrate that the difference between mean-centred and median-centred deviations is not determined solely by overall dispersion. It also depends on the distributional shape and the location of the chosen reference functional. In particular, a mean-centred absolute deviation can be appreciably larger than its median-centred counterpart when right-skewness is substantial.

9.7.3 Effect of sample size on the estimated deviations

The stability of the deviation estimators was assessed by examining their Monte Carlo means and standard deviations over the five sample sizes. Table 5 summarizes the Monte Carlo means of the two deviations and their ratio for each sample size.

For the standard normal distribution, the Monte Carlo mean of the mean-centred deviation increases from 0.7901 at $n = 50$ to 0.7969 at $n = 1000$, approaching the theoretical population

Table 5: Monte Carlo means of the mean-centred and median-centred absolute deviations and their ratio across sample sizes ($B = 2000$).

Distribution	n	$\mathbb{E}_{MC}(\widehat{D}_{\text{mean}})$	$\mathbb{E}_{MC}(\widehat{D}_{\text{median}})$	$\mathbb{E}_{MC}(\widehat{R})$
Normal$(0, 1)$	50	0.7901	0.7856	0.9943
	100	0.7928	0.7905	0.9971
	250	0.7962	0.7953	0.9988
	500	0.7971	0.7967	0.9994
	1000	0.7969	0.7966	0.9997
Uniform$(-1, 1)$	50	0.4939	0.4907	0.9936
	100	0.4955	0.4939	0.9966
	250	0.4980	0.4973	0.9987
	500	0.4993	0.4990	0.9993
	1000	0.4995	0.4993	0.9997
Exponential(1)	50	0.7309	0.6862	0.9408
	100	0.7302	0.6872	0.9421
	250	0.7320	0.6893	0.9421
	500	0.7354	0.6926	0.9420
	1000	0.7340	0.6915	0.9422
Student-t_3	50	1.1015	1.0883	0.9894
	100	1.1099	1.1000	0.9939
	250	1.1062	1.1035	0.9976
	500	1.1033	1.1020	0.9988
	1000	1.1026	1.1019	0.9994
Lognormal$(0, 1)$	50	1.2545	1.1105	0.8945
	100	1.2602	1.1200	0.8933
	250	1.2498	1.1135	0.8928
	500	1.2634	1.1254	0.8916
	1000	1.2588	1.1223	0.8920
Normal mixture	50	1.3847	1.3563	0.9793
	100	1.3951	1.3689	0.9811
	250	1.3969	1.3715	0.9817
	500	1.3989	1.3740	0.9822
	1000	1.4001	1.3751	0.9821

value

$$\mathbb{E}|X| = \sqrt{\frac{2}{\pi}} \approx 0.7979.$$

The corresponding median-centred estimate increases from 0.7856 to 0.7966. A comparable pattern occurs for the uniform distribution, for which the population absolute deviation about the common mean and median is 0.5. At $n = 1000$, the estimated Monte Carlo means are 0.4995 and 0.4993, respectively.

The Monte Carlo means for the exponential distribution are comparatively stable across sample sizes. The mean-centred deviation ranges from 0.7302 to 0.7354 over the five sample sizes, while the median-centred deviation ranges from 0.6862 to 0.6926. The ratio remains close to 0.942 from $n = 100$ onwards. This persistence suggests that the difference between the two reference functionals reflects the population asymmetry rather than a small-sample phenomenon alone.

For the lognormal distribution, the ratio remains close to 0.89 throughout the design. Its Monte Carlo mean is 0.8945 at $n = 50$ and 0.8920 at $n = 1000$. Thus, the relative reduction associated with median-centering persists even as the sample size increases. The Gaussian mixture also exhibits a stable difference: the ratio changes from 0.9793 at $n = 50$ to 0.9821 at $n = 1000$.

For the Student's t_3 distribution, the ratio moves towards one as the sample size increases, from 0.9894 at $n = 50$ to 0.9994 at $n = 1000$. This is consistent with the common population centre of the symmetric distribution. However, the rate at which the estimators stabilize should be assessed separately from their limiting population values, because heavy tails can produce substantial finite-sample fluctuations.

Overall, the sample-size results show that the Monte Carlo means become more stable as n increases, while the differences between the two deviation measures remain distribution-dependent. The ratios for symmetric distributions approach one, whereas the lower ratios for the exponential and lognormal distributions persist across the sample sizes considered.

9.7.4 Monte Carlo variability and distributional sensitivity

The Monte Carlo standard deviations provide an additional measure of the stability of the deviation estimators. At $n = 1000$, the Monte Carlo standard deviations of $\widehat{D}_{\text{mean}}$ are 0.0189 for the normal distribution, 0.0090 for the uniform distribution, 0.0285 for the exponential distribution, 0.0411 for the Student's t_3 distribution, 0.0789 for the lognormal distribution, and 0.0279 for the Gaussian mixture. The corresponding Monte Carlo standard deviations of $\widehat{D}_{\text{median}}$ are 0.0189, 0.0090, 0.0253, 0.0409, 0.0616, and 0.0286.

These results indicate that the lognormal deviation estimators exhibit the largest absolute Monte Carlo variability among the six distributions considered. The substantial variability is consistent with the long right tail of the lognormal distribution, which allows a relatively small number of large observations to influence the average absolute deviation. Median-centering reduces the Monte Carlo standard deviation from 0.0789 to 0.0616 at $n = 1000$. The corresponding reduction for the exponential distribution is from 0.0285 to 0.0253.

For the Student's t_3 distribution, the Monte Carlo standard deviation of the mean-centred absolute deviation is 0.2033 at $n = 50$ and 0.0411 at $n = 1000$. The corresponding median-centred standard deviation decreases from 0.1875 to 0.0409. The relatively high variability at small sample sizes reflects the heavy-tailed nature of the distribution. Although the variance of a t_3 random variable is finite, its fourth moment is infinite, so statistics that depend strongly on extreme squared observations can exhibit substantial variability.

The behaviour of the sample standard deviation reinforces this distinction. At $n = 1000$, the Monte Carlo standard deviation of the conventional sample standard deviation is approximately 0.2438 for the Student's t_3 distribution, compared with 0.0411 for the mean-centred absolute deviation. For the lognormal distribution, the corresponding values are approximately 0.2959

and 0.0789. Thus, in these simulations, the standard deviation estimator exhibits appreciably greater sampling variability than the mean absolute deviation estimator under the heavy-tailed and strongly right-skewed distributions.

These comparisons should not be interpreted as establishing that mean absolute deviation is uniformly superior to standard deviation. The two measures quantify different aspects of dispersion, and their relative performance depends on the distribution, the target parameter, and the criterion used to evaluate estimation. The results instead illustrate that squared-deviation-based estimators can be more sensitive to extreme observations than absolute-deviation-based estimators.

9.7.5 Quantile-based deviations

The mean absolute deviation provides a single summary of the distribution of absolute deviations. Quantile-based deviations provide additional information about the central spread and upper tail. Let

$$Y_i = |X_i - \widetilde{X}|,$$

and denote by $\widehat{Q}_Y(p)$ the empirical quantile of order p. Table 6 presents the Monte Carlo means of the 0.90 quantile of the absolute deviations about the sample median at $n = 1000$.

Table 6: Monte Carlo mean and standard deviation of the 0.90 quantile of absolute deviations about the sample median at $n = 1000$.

Distribution	MC mean	MC standard deviation
Normal$(0, 1)$	1.6422	0.0454
Uniform$(-1, 1)$	0.8980	0.0095
Exponential(1)	1.6021	0.0892
Student-t_3	2.3447	0.1003
Lognormal$(0, 1)$	2.5911	0.1875
Normal mixture	2.7036	0.0559

The largest Monte Carlo mean of the 0.90 absolute-deviation quantile is observed for the Gaussian mixture, at 2.7036, followed by the lognormal distribution, at 2.5911, and the Student's t_3 distribution, at 2.3447. The normal and exponential distributions produce similar values, 1.6422 and 1.6021, respectively, whereas the uniform distribution has the smallest value, 0.8980. These differences indicate that the upper part of the absolute-deviation distribution can vary substantially even when the mean absolute deviation alone does not fully capture the distributional shape.

The lognormal distribution has a relatively large Monte Carlo standard deviation for this quantile, 0.1875, compared with 0.0559 for the Gaussian mixture and 0.0454 for the normal distribution. This finding is consistent with the sensitivity of upper-tail quantiles to extreme observations in strongly right-skewed distributions. The Student's t_3 quantile also exhibits appreciable variability, with a Monte Carlo standard deviation of 0.1003.

The sample-size pattern is informative as well. For the normal distribution, the Monte Carlo mean of the 0.90 quantile increases from 1.5967 at $n = 50$ to 1.6422 at $n = 1000$. For the Student's t_3 distribution, it increases from 2.2600 to 2.3447, while for the lognormal distribution it increases from 2.4579 to 2.5911. The corresponding Monte Carlo standard deviations decrease as the sample size increases, although the rate of decrease differs across distributions.

The quantile-based results therefore complement the mean-deviation comparisons. In particular, a distribution may exhibit a relatively moderate mean absolute deviation but still possess a large upper quantile of absolute deviations. Reporting several quantile levels can consequently provide a more detailed description of dispersion than reporting a single deviation functional.

9.7.6 Numerical verification of the inequalities

Five empirical checks were recorded in each replication: the sample-median minimization inequality, the empirical Cauchy–Schwarz inequality, two quantile bounds based on the empirical mean absolute deviations, and the bound on the median of the absolute deviations. The raw results showed zero violations for all five checks across all six distributions and all five sample sizes. Thus, each of the 30 distribution–sample-size combinations had zero recorded violations in its 2000 replications.

The first check was

$$\widehat{D}_{\text{median}} \leq \widehat{D}_{\text{mean}}.$$

This inequality follows directly from the minimization property of the sample median under absolute loss. The second check was

$$\widehat{D}_{\text{mean}} \leq \left\{ \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right\}^{1/2},$$

which follows from the Cauchy–Schwarz inequality. Both are deterministic finite-sample properties and are expected to hold for every sample, apart from numerical rounding error.

The empirical quantile checks used

$$\widehat{Q}_Y(p) \leq \frac{\widehat{D}_{\text{median}}}{1-p} \quad \text{and} \quad \widehat{Q}_Y(p) \leq \frac{\widehat{D}_{\text{mean}}}{1-p}, \qquad p \in \{0.25, 0.50, 0.75, 0.90\}.$$

The final check was

$$\widehat{Q}_Y(0.50) \leq 2\widehat{D}_{\text{mean}}.$$

No violation of these inequalities was recorded under the empirical quantile convention and numerical tolerance used in the simulation.

The zero violation frequencies are consistent with the mathematical inequalities. They should not, however, be interpreted as evidence that the inequalities hold with an estimated probability of one in an arbitrary population. The checks are based on inequalities that hold by construction for the empirical quantities under the specified definitions. Their principal role is to validate the numerical implementation rather than to provide an independent statistical test of the population results.

9.7.7 Overall interpretation

The repeated Monte Carlo results support several conclusions. First, the difference between mean-centred and median-centred absolute deviations is strongly distribution-dependent. For the symmetric normal, uniform, and Student's t_3 distributions, the two deviations are close, and their ratio approaches one as the sample size increases. For the exponential and lognormal distributions, the median-centred deviation is consistently smaller, with relative reductions of approximately 5.8% and 10.8%, respectively, at $n = 1000$. The Gaussian mixture exhibits a smaller but persistent reduction of approximately 1.8%.

Second, the Monte Carlo standard deviations generally decrease as the sample size increases, indicating greater sampling stability. Nevertheless, the magnitude of this variability differs across distributions. The lognormal and Student's t_3 settings illustrate that heavy tails or strong skewness can substantially affect the variability of deviation estimators and upper quantiles.

Third, quantile-based deviations reveal distributional features that are not fully represented by the mean absolute deviation. The comparatively large 0.90 quantiles for the lognormal, Student's t_3, and Gaussian mixture distributions indicate substantial upper-tail or between-component dispersion.

Finally, all five empirical inequalities were satisfied in the recorded replications. This outcome is consistent with their deterministic empirical formulations and supports the correctness of the numerical checks, subject to the specified tolerance and quantile convention.

Taken together, the results demonstrate the practical value of considering several deviation functionals rather than relying on a single measure of dispersion. Mean absolute deviation, median-centred absolute deviation, and quantile-based deviations emphasize different aspects of distributional spread. Their comparative behaviour depends on symmetry, skewness, tail weight, and mixture structure, and these differences remain visible across the sample sizes investigated.

9.7.8 Supplementary R code for mode sensitivity

A separate sensitivity analysis of the estimated mode of the absolute deviations may be conducted using kernel density estimation (KDE), with the global mode and the number of detected interior local maxima recorded for each bandwidth choice. The number of local maxima serves as a numerical diagnostic of the estimated density's shape and may vary with the bandwidth, grid resolution, and peak-detection procedure; therefore, it should not be interpreted as a formal test of multimodality. The corresponding R implementation is provided in Subsection (A.3) of the Appendix. For reproducible and publication-quality reporting, the estimated global modes should be presented alongside the selected bandwidths and the numbers of detected local maxima. Where multiple maxima are identified, the estimated density plots should be examined to distinguish potentially meaningful modes from minor numerical fluctuations. Furthermore, the analysis should be repeated across independent Monte Carlo samples to assess the sampling variability of the estimated modes and the stability of mode selection under alternative bandwidth choices.

9.8 Bandwidth sensitivity of mode-based deviations

The sensitivity of mode-based deviations to the density estimation procedure was investigated using a Gaussian kernel density estimator. The analysis was conducted on the absolute deviations

$$Y_i = |X_i - \widetilde{X}|,$$

where $\widetilde{X}$ denotes the sample median of the Gaussian-mixture observations. The bandwidth selected by the normal-reference rule was denoted by h_0, and the sensitivity analysis was performed using the bandwidths $0.75h_0$, h_0, and $1.25h_0$. The estimated global mode was defined as the grid point at which the estimated kernel density attained its maximum. The number of interior local maxima was also recorded as a diagnostic of the estimated density's modal structure.

Table 7: Bandwidth sensitivity of the estimated mode of the absolute deviations for the Gaussian-mixture sample.

Bandwidth multiplier	Bandwidth	Estimated global mode	Number of local maxima
0.75	0.05093051	0.1636450	5
1.00	0.06790734	0.1912091	3
1.25	0.08488418	0.2189289	3

The estimated global mode increased from 0.1636450 at bandwidth $0.75h_0$ to 0.2189289 at bandwidth $1.25h_0$. The absolute difference between these estimates was 0.0552839, representing an increase of approximately 33.8% relative to the estimate obtained with the smallest bandwidth. The estimate obtained at the default bandwidth was 0.1912091, lying between

the estimates obtained under the narrower and wider bandwidth choices. This pattern indicates that the estimated mode is sensitive to the amount of smoothing applied to the empirical distribution of absolute deviations.

The number of detected interior local maxima decreased from five under the bandwidth multiplier 0.75 to three under the default and enlarged bandwidths. This behaviour is consistent with the suppression of small-scale density fluctuations under increased smoothing. Nevertheless, the number of detected local maxima should not be interpreted as a direct estimate of the true number of population modes. The count may depend on the sample, the bandwidth, the resolution of the evaluation grid, and the numerical rule used to identify local maxima.

The results illustrate an important distinction between mode-based deviations and deviations defined through the mean or median. The mean and median are determined by expectation and quantile properties, respectively, whereas the mode is obtained from the maximum of an estimated density. Consequently, the latter may exhibit additional sensitivity to smoothing and to the local shape of the distribution. This issue is particularly relevant when absolute deviations are used as the variable whose density is estimated, because the distribution of $|X - \widetilde{X}|$ need not preserve the modal structure of the original observations.

Accordingly, the estimated mode of the absolute deviations should be reported together with the bandwidth-selection rule and a sensitivity analysis. The present results provide numerical evidence of bandwidth sensitivity for the particular simulated sample; they do not establish the population mode, the true number of modes, or the sampling variability of the mode estimator. A more comprehensive investigation would repeat the bandwidth analysis across independent Monte Carlo samples and examine the stability of both the global mode and the number of detected local maxima.

9.9 Further directions

The present framework opens several directions for further theoretical investigation and computational assessment. Although the established inequalities provide distribution-free relationships among selected deviation measures, the numerical investigations remain dependent on the distributions considered, the sample sizes employed, and the manner in which the underlying statistical functionals are estimated. Further research may therefore be directed towards examining the sensitivity of functional deviations to tail behaviour, asymmetry, multimodality, quantile selection, and the dimensional structure of the observations.

First, the repeated-sampling investigation may be extended to a broader family of Student's t distributions, including t_5 and t_{10}. Such an extension would facilitate a systematic assessment of the influence of tail weight on the finite-sample behaviour of mean-centred, median-centred, and quantile-based deviations. In particular, the comparison between t_3, t_5, t_{10}, and the standard normal distribution would help distinguish the consequences of heavy tails from those attributable to other distributional characteristics. The t_3 distribution has a finite variance but an infinite fourth moment, whereas the t_5 and t_{10} distributions possess finite fourth moments. Accordingly, the comparison would also provide an opportunity to investigate whether the variability of sample-based deviation measures changes systematically as additional moments become finite. Nevertheless, the effects of tail weight should not be interpreted exclusively through moment existence, since the empirical behaviour of these measures also depends on the influence of extreme observations and the robustness of the corresponding location estimators. The investigation could therefore incorporate Monte Carlo bias, empirical standard deviation, root mean squared error, and coverage probabilities, where appropriate, to distinguish systematic estimation error from sampling variability.

Second, the Gaussian-mixture setting may be developed into a structured sensitivity analysis by varying the mixing probability, the separation between component means, and the component

variances. For example, a two-component mixture may be expressed as

$$X \sim \pi N(\mu_1, \sigma_1^2) + (1-\pi)N(\mu_2, \sigma_2^2), \qquad 0 < \pi < 1.$$

The mixing probability π determines the relative contribution of the two components, while the difference $|\mu_2 - \mu_1|$ controls their separation. Changes in these parameters may produce substantial differences between the mean, median, and mode, and may also alter the shape of the distribution of absolute deviations. When the component means are sufficiently separated, the density may exhibit multiple local maxima; however, the existence and location of these maxima depend jointly on the mixing probability, component variances, and separation. Consequently, a mode-centred deviation may be particularly sensitive to changes in mixture structure. A systematic investigation could examine the difference between the mean-centred and median-centred mean absolute deviations, the ratio of these two measures, and the corresponding mode-centred deviation across a grid of mixture parameters. The number and locations of local density maxima could also be recorded to assess whether changes in multimodality are accompanied by changes in the estimated deviation measures. Since a mode need not be unique and does not generally minimize expected absolute loss, mode-based comparisons should be treated separately from inequalities established for the mean and median.

Third, the quantile-based component of the framework may be extended to more extreme probability levels, including 0.95 and 0.99. For a nonnegative deviation variable D, the quantile $Q_D(p)$ describes the level below which a proportion p of the absolute deviations lies. Higher quantile levels consequently provide information about the upper tail of the deviation distribution that may not be adequately represented by the median or lower quantiles. This extension would be especially informative for heavy-tailed and strongly skewed distributions, for which the upper portion of the deviation distribution may behave differently from its central portion. However, extreme sample quantiles are determined by relatively few upper-order observations and may therefore exhibit appreciable finite-sample variability. Their assessment should be accompanied by Monte Carlo standard errors, empirical bias, and, where feasible, confidence intervals. Larger sample sizes and sufficiently many Monte Carlo replications would be required to obtain stable comparisons. In addition, the distinction between a population quantile and its sample estimator should be maintained explicitly, since the theoretical quantile inequalities do not by themselves establish the finite-sample accuracy of a particular quantile estimator.

Fourth, the effects of quantile conventions and finite-sample interpolation rules warrant separate consideration. In a finite sample, an empirical quantile may be defined through an order statistic or obtained by interpolation between adjacent order statistics. Different conventions can produce different numerical values, particularly when the sample size is small or the target probability lies near the upper tail. A sensitivity analysis could therefore compare several standard empirical quantile definitions while holding the underlying data and probability levels fixed. The analysis would help identify whether observed differences between quantile-based deviation measures arise primarily from the distribution, sample size, or quantile-estimation convention. It would also be useful to distinguish the empirical quantile of the observed absolute deviations from the population quantile of the corresponding absolute-deviation random variable. These distinctions are important when interpreting numerical evidence in support of theoretical relationships.

Fifth, the framework may be extended to investigate the robustness and stability of functional deviations under contamination. A contaminated distribution may be represented by

$$F_\epsilon = (1-\epsilon)F + \epsilon G, \qquad 0 \le \epsilon < 1,$$

where F denotes the reference distribution and G represents a contaminating distribution. The sensitivity of each deviation measure could then be examined as the contamination proportion ϵ and the contaminating distribution are varied. Mean-centred deviations may be affected

by substantial changes in the sample mean under extreme contamination, whereas median-centred measures may exhibit greater resistance when the contamination remains sufficiently limited. The behaviour of mode-centred deviations requires particular care because a small contaminating component may generate an additional local density maximum or alter which maximum is identified as the global mode. Such an investigation would provide a more direct connection between the proposed framework and robust statistical analysis. It would also clarify that robustness depends not only on the choice of reference functional but on the functional applied to the absolute deviations themselves.

Sixth, additional theoretical results may be sought for deviations defined relative to general location and quantile functionals. Let $T(F)$ denote a location functional associated with a distribution F, and let

$$D_T = |X - T(F)|.$$

A general deviation functional may then be defined as

$$\mathcal{D}_{S,T}(F) = S\big(\mathcal{L}(D_T)\big),$$

where $\mathcal{L}(D_T)$ denotes the distribution of D_T and S is a functional acting on that distribution. This formulation separates the selection of the reference location from the selection of the summary used to quantify the deviations. The mean absolute deviation about the median, for example, corresponds to choosing the median as the reference location and the expectation as the deviation summary. A quantile-based deviation corresponds to selecting a quantile of the absolute-deviation distribution, while a mode-based deviation uses a density mode of that distribution when such a mode is well defined. Further work could identify conditions under which comparisons between two choices of S or T are possible. In particular, it would be valuable to determine which inequalities depend only on general properties of the functionals and which require additional distributional assumptions, such as symmetry, unimodality, or moment existence. The nonuniqueness and potential instability of modes would need to be incorporated explicitly into any general theory involving mode-based deviations.

Finally, the framework may be generalized to multivariate and functional observations. For a random vector $X \in \mathbb{R}^d$, a vector-valued location functional $T(F)$ could serve as the reference point, and deviations could be defined through a norm,

$$D_T = \|X - T(F)\|,$$

where $\|\cdot\|$ is a suitable norm on $\mathbb{R}^d$. Different choices of norm may produce different measures of multivariate dispersion and may affect sensitivity to individual coordinates and extreme observations. Alternatively, depth-based distances could be used to define deviations relative to central regions of a multivariate distribution. For functional observations, where X is a random function on a domain $\mathcal{T}$, the distance from a mean function or a suitably defined median function could be measured using the L^2 norm,

$$D_T = \left\{ \int_{\mathcal{T}} \big[X(t) - T(F)(t)\big]^2 \, dt \right\}^{1/2}.$$

The resulting scalar deviation could subsequently be summarized through its expectation, median, or selected quantiles. Such extensions would require attention to the existence of the relevant functionals, the integrability of the observations, the choice of norm or depth, and the computational complexity of estimation. They would nevertheless permit the central idea of separating the reference functional from the deviation-summary functional to be applied beyond univariate scalar data.

In summary, further development should distinguish three complementary objectives: establishing additional population-level inequalities, evaluating the finite-sample behaviour of estimated deviation measures, and investigating the stability of these measures under changes

in distributional structure and estimation procedures. The inequalities already established are mathematical statements under their specified assumptions, whereas numerical agreement in a finite simulation constitutes empirical evidence rather than a proof. Similarly, the mode-based analysis provides a useful illustration of sensitivity to density estimation, but it does not establish a universal ordering between mode-centred and other deviations. Maintaining these distinctions would support a more comprehensive and theoretically coherent treatment of deviations defined relative to general statistical functionals.

10 Conclusion

A unified framework for functional deviations about statistical functionals has been developed to distinguish the location used as a reference from the functional used to summarize the resulting absolute deviations. This formulation encompasses conventional measures, including the mean absolute deviation about the mean and the mean absolute deviation about the median, while also accommodating quantile-based and mode-based summaries. The framework therefore provides a common perspective for examining how the reference location and the selected deviation summary contribute separately to the numerical description of dispersion.

Several fundamental inequalities have been established under appropriate assumptions. In particular, the expected absolute deviation about a median is no greater than the expected absolute deviation about the mean, and the latter is bounded above by the standard deviation when the second moment is finite. Additional bounds for medians and quantiles of nonnegative absolute-deviation variables provide complementary information about the distribution of deviations. These relationships follow from general properties of the relevant statistical functionals and do not depend on a particular parametric distribution, provided that the stated integrability and quantile conditions are satisfied. Their interpretation should nevertheless remain tied to the exact quantities being compared: inequalities for expected absolute deviations do not automatically extend to medians, quantiles, or modes of the deviation distribution.

The numerical investigations illustrate how the choice of reference functional can influence the resulting deviation measures across distributions with different shapes. For symmetric distributions, mean-centred and median-centred deviations may be close or coincide at the population level under suitable conditions. For skewed distributions, their differences may be more pronounced because the mean and median need not identify the same location. Heavy-tailed distributions introduce additional considerations concerning the stability of sample-based estimates, while Gaussian mixtures illustrate how component separation and multimodality can complicate the interpretation of a single reference location. Quantile-based deviations provide a complementary description by distinguishing the central and upper portions of the absolute-deviation distribution. These numerical patterns are illustrative and should not be interpreted as universal rankings beyond the assumptions and distributions examined.

The mode-based analysis further emphasizes the need to distinguish a mathematically defined population functional from its numerical estimate. A mode is determined by the maximum of a density rather than by minimization of expected absolute loss, and its estimation may depend on bandwidth selection, the presence of multiple local maxima, and the rule used to identify a global maximum. Consequently, mode-centred deviations can be informative in settings where density peaks have substantive meaning, but their properties require separate investigation and cannot be inferred directly from the inequalities established for means and medians.

Several limitations remain. The range of distributions and simulation settings may be expanded, the finite-sample performance of extreme quantiles may be studied more extensively, and the sensitivity of mode-based measures may be examined under alternative density-estimation procedures. Further theoretical work may also establish conditions for comparing general location and deviation functionals, while extensions to multivariate and functional data

may broaden the scope of the framework. Such developments would require explicit treatment of existence, integrability, uniqueness, and estimation uncertainty.

Overall, functional deviations provide a flexible way to describe dispersion relative to different statistical reference points. The principal contribution of the framework is the separation of the reference location from the functional used to summarize deviations, together with rigorous inequalities for selected cases and a basis for systematic numerical investigation. By keeping theoretical results distinct from simulation evidence and estimation-sensitive quantities, the framework supports a more precise understanding of the relationships among conventional, robust, quantile-based, and mode-based measures of deviation.

Acknowledgements

The author is grateful to the anonymous reviewers for their constructive comments and suggestions, which may help improve the clarity and presentation of the manuscript.

Funding

No external funding was received for this research.

Data Availability Statement

The data supporting the findings of this study are generated through numerical simulations using the probability distributions and simulation settings described in the manuscript. The simulation procedures and corresponding R code are provided in the manuscript and its Appendix to facilitate reproducibility.

Declaration of Conflicting Interests

The author declares no competing interests that could have influenced the research, authorship, or publication of this manuscript.

Ethics Approval

Ethics approval was not required for this study because it involved mathematical analysis and computer-based simulations and did not involve human participants, animals, or identifiable personal data.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Author Contributions

The author conceptualized the study, developed the mathematical framework, conducted the numerical investigations, interpreted the results, and prepared the manuscript.

Code Availability

The R code used for the numerical investigations is provided in the Appendix. It may be used to reproduce the reported simulation procedures, subject to the distributional specifications and computational settings described in the manuscript.

References

[1] Frank R. Hampel, Elvezio M. Ronchetti, Peter J. Rousseeuw, and Werner A. Stahel. Robust statistics: The approach based on influence functions. *Wiley Series in Probability and Statistics*, 1986.

[2] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. John Wiley & Sons, 2 edition, 2009.

[3] Roger Koenker. *Quantile Regression*. Cambridge University Press, 2005.

[4] Peter J. Rousseeuw and Christophe Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.

A Appendix: Codes

A.1 R implementation 1

```r
set.seed(20261022)
N <- 300000

distributions <- list(
  "Normal(0,1)" = function(n) rnorm(n),
  "Uniform(-1,1)" = function(n) runif(n, -1, 1),
  "Exponential(1)" = function(n) rexp(n),
  "Student-t(3)" = function(n) rt(n, df = 3),
  "Lognormal(0,1)" = function(n) rlnorm(n),
  "Normal mixture" = function(n) {
    z <- rbinom(n, 1, 0.3)
    ifelse(z == 1,
           rnorm(n, 3, 0.5),
           rnorm(n))
  }
)

calculate_deviations <- function(x) {
  mu <- mean(x)
  med <- median(x)

  # Population-form empirical standard deviation
  sdv <- sd(x) * sqrt((length(x)-1)/length(x))

  d_mean <- abs(x-mu)
  d_med <- abs(x-med)

  q25 <- quantile(d_med, .25, names=FALSE)
  q50 <- quantile(d_med, .50, names=FALSE)
  q75 <- quantile(d_med, .75, names=FALSE)
  q90 <- quantile(d_med, .90, names=FALSE)

  checks <- c(
    median_abs_dev_bound =
      median(d_mean) <= 2*mean(d_mean),
    median_reference_bound =
      mean(d_med) <= mean(d_mean),
    sd_bound =
      mean(d_mean) <= sdv,
    quantile_order =
      q25 <= q50 && q50 <= q75 && q75 <= q90,
    quantile_mean_bound =
      q90 <= mean(d_med)/(1-.90)
  )

  data.frame(
    Mean=mu,
    Median=med,
    MeanAD_Median=mean(d_med),
    MeanAD_Mean=mean(d_mean),
    SD=sdv,
    MedianAD_Mean=median(d_mean),
    Q25_AbsDevMedian=q25,
    Q50_AbsDevMedian=q50,
    Q75_AbsDevMedian=q75,
    Q90_AbsDevMedian=q90,
    Ratio_MedianMeanAD=mean(d_med)/mean(d_mean),
    Ratio_MeanAD_SD=mean(d_mean)/sdv,
    All_Inequalities_Hold=all(checks)
```

```r
  )
}

results <- do.call(
  rbind,
  lapply(names(distributions), function(nm) {
    ans <- calculate_deviations(
      distributions[[nm]](N)
    )
    cbind(Distribution=nm, ans)
  })
)

rownames(results) <- NULL
print(results, digits=4)
print(all(results$All_Inequalities_Hold))
```

A.2 R implementation 2

```r
# ============================================================
# REPEATED MONTE CARLO STUDY OF FUNCTIONAL DEVIATIONS
# ============================================================

set.seed(20261022)

B <- 2000
n_grid <- c(50, 100, 250, 500, 1000)
alpha_grid <- c(0.25, 0.50, 0.75, 0.90)
tol <- 1e-10

# Six distributions used in the initial experiment
generate_sample <- function(n, distribution) {
  switch(
    distribution,
    "Normal(0,1)" = rnorm(n, mean = 0, sd = 1),
    "Uniform(-1,1)" = runif(n, min = -1, max = 1),
    "Exponential(1)" = rexp(n, rate = 1),
    "Student-t(3)" = rt(n, df = 3),
    "Lognormal(0,1)" = rlnorm(n, meanlog = 0, sdlog = 1),
    "Normal mixture" = {
      component <- rbinom(n, size = 1, prob = 0.30)
      ifelse(
        component == 1,
        rnorm(n, mean = 3, sd = 0.5),
        rnorm(n, mean = 0, sd = 1)
      )
    },
    stop("Unknown distribution")
  )
}

distribution_names <- c(
  "Normal(0,1)",
  "Uniform(-1,1)",
```

```
  "Exponential(1)",
  "Student-t(3)",
  "Lognormal(0,1)",
  "Normal mixture"
)

# ------------------------------------------------------------
# Statistics from one sample
# ------------------------------------------------------------

calculate_statistics <- function(x) {
  n <- length(x)

  xbar <- mean(x)
  med <- median(x)

  # Absolute deviations about the sample mean and median
  dev_mean <- abs(x - xbar)
  dev_median <- abs(x - med)

  mad_mean <- mean(dev_mean)
  mad_median <- mean(dev_median)

  # RMS deviation with denominator n
  rms_mean <- sqrt(mean((x - xbar)^2))

  # Conventional sample standard deviation, denominator n - 1
  sd_sample <- sd(x)

  # Quantiles of absolute deviations about the sample median
  qdev <- quantile(
    dev_median,
    probs = alpha_grid,
    type = 1,
    names = FALSE
  )

  # Empirical median of absolute deviations about the mean
  med_dev_mean <- quantile(
    dev_mean,
    probs = 0.50,
    type = 1,
    names = FALSE
  )

  # Numerical checks
  violation_median <- mad_median > mad_mean + tol
  violation_cauchy <- mad_mean > rms_mean + tol

  # Empirical Markov-type inequalities
  violation_quantile_median <- any(
```

```
    qdev > mad_median / (1 - alpha_grid) + tol
  )

  violation_quantile_mean <- any(
    qdev > mad_mean / (1 - alpha_grid) + tol
  )

  # Median of nonnegative deviations <= 2 times their mean
  violation_median_bound <- med_dev_mean >
    2 * mad_mean + tol

  c(
    mean = xbar,
    median = med,
    MAD_mean = mad_mean,
    MAD_median = mad_median,
    difference = mad_mean - mad_median,
    ratio = if (mad_mean > 0) mad_median / mad_mean else NA,
    RMS_mean = rms_mean,
    SD_sample = sd_sample,
    median_absdev_mean = med_dev_mean,
    Q25_absdev_median = qdev[1],
    Q50_absdev_median = qdev[2],
    Q75_absdev_median = qdev[3],
    Q90_absdev_median = qdev[4],
    violation_median = as.numeric(violation_median),
    violation_cauchy = as.numeric(violation_cauchy),
    violation_q_median = as.numeric(violation_quantile_median),
    violation_q_mean = as.numeric(violation_quantile_mean),
    violation_median_bound = as.numeric(violation_median_bound)
  )
}

# ------------------------------------------------------------
# Monte Carlo experiment
# ------------------------------------------------------------

result_list <- vector(
  "list",
  length(n_grid) * length(distribution_names) * B
)

counter <- 0L

for (dist in distribution_names) {
  for (n in n_grid) {
    for (b in seq_len(B)) {
      x <- generate_sample(n, dist)
      stats_b <- calculate_statistics(x)

      counter <- counter + 1L
```

```
      result_list[[counter]] <- data.frame(
        distribution = dist,
        n = n,
        replication = b,
        as.list(stats_b),
        check.names = FALSE
      )
    }
  }
}

mc_results <- do.call(rbind, result_list)
rownames(mc_results) <- NULL

# ------------------------------------------------------------
# Summaries: Monte Carlo means, SDs, and MC standard errors
# ------------------------------------------------------------

measure_names <- c(
  "MAD_mean",
  "MAD_median",
  "difference",
  "ratio",
  "RMS_mean",
  "SD_sample",
  "Q25_absdev_median",
  "Q50_absdev_median",
  "Q75_absdev_median",
  "Q90_absdev_median"
)

summary_list <- list()
counter <- 0L

for (dist in distribution_names) {
  for (n in n_grid) {
    z <- subset(
      mc_results,
      distribution == dist & mc_results$n == n
    )

    for (measure in measure_names) {
      v <- z[[measure]]
      counter <- counter + 1L

      summary_list[[counter]] <- data.frame(
        distribution = dist,
        n = n,
        measure = measure,
        MC_mean = mean(v, na.rm = TRUE),
```

```
        MC_SD = sd(v, na.rm = TRUE),
        MCSE = sd(v, na.rm = TRUE) / sqrt(sum(!is.na(v))),
        MC_median = median(v, na.rm = TRUE),
        MC_Q025 = quantile(v, 0.025, na.rm = TRUE),
        MC_Q975 = quantile(v, 0.975, na.rm = TRUE)
      )
    }
  }
}

mc_summary <- do.call(rbind, summary_list)
rownames(mc_summary) <- NULL

# ------------------------------------------------------------
# Empirical inequality-violation frequencies
# ------------------------------------------------------------

check_names <- c(
  "violation_median",
  "violation_cauchy",
  "violation_q_median",
  "violation_q_mean",
  "violation_median_bound"
)

violation_list <- list()
counter <- 0L

for (dist in distribution_names) {
  for (n in n_grid) {
    z <- subset(
      mc_results,
      distribution == dist & mc_results$n == n
    )

    for (check in check_names) {
      counter <- counter + 1L

      violation_list[[counter]] <- data.frame(
        distribution = dist,
        n = n,
        inequality = check,
        violations = sum(z[[check]] > 0),
        replications = nrow(z),
        violation_rate = mean(z[[check]] > 0)
      )
    }
  }
}

violation_summary <- do.call(rbind, violation_list)
```

```
rownames(violation_summary) <- NULL

# ------------------------------------------------------------
# Display and save results
# ------------------------------------------------------------

cat("\n--- MONTE CARLO SUMMARY ---\n")
print(mc_summary, row.names = FALSE)

cat("\n--- INEQUALITY CHECK SUMMARY ---\n")
print(violation_summary, row.names = FALSE)

write.csv(
  mc_results,
  "functional_deviation_mc_raw.csv",
  row.names = FALSE
)

write.csv(
  mc_summary,
  "functional_deviation_mc_summary.csv",
  row.names = FALSE
)

write.csv(
  violation_summary,
  "functional_deviation_inequality_checks.csv",
  row.names = FALSE
)
```

A.3 R implementation 3

```
# ============================================================
# KDE MODE AND BANDWIDTH SENSITIVITY
# ============================================================

estimate_modes <- function(x, bandwidth_multiplier = 1,
                           n_grid = 2048) {
  h0 <- bw.nrd0(x)
  h <- bandwidth_multiplier * h0

  if (!is.finite(h) || h <= 0) {
    return(data.frame(
      bandwidth_multiplier = bandwidth_multiplier,
      bandwidth = h,
      global_mode = NA_real_,
      number_of_local_maxima = NA_integer_
    ))
  }

  den <- density(
    x,
```

```
    bw = h,
    kernel = "gaussian",
    n = n_grid
  )

  y <- den$y
  xx <- den$x

  # Interior local maxima on the evaluation grid
  local_index <- which(
    y[2:(length(y) - 1)] > y[1:(length(y) - 2)] &
    y[2:(length(y) - 1)] >= y[3:length(y)]
  ) + 1L

  # Grid-based global mode
  global_mode <- xx[which.max(y)]

  data.frame(
    bandwidth_multiplier = bandwidth_multiplier,
    bandwidth = h,
    global_mode = global_mode,
    number_of_local_maxima = length(local_index)
  )
}

# Example: investigate the Gaussian mixture
set.seed(20261022)
x_mix <- ifelse(
  rbinom(300000, size = 1, prob = 0.30) == 1,
  rnorm(300000, mean = 3, sd = 0.5),
  rnorm(300000, mean = 0, sd = 1)
)

mode_sensitivity <- do.call(
  rbind,
  lapply(
    c(0.75, 1, 1.25),
    function(multiplier) {
      estimate_modes(
        x_mix,
        bandwidth_multiplier = multiplier
      )
    }
  )
)

print(mode_sensitivity)

# The same analysis can be applied to absolute deviations
# about the sample median.
dev_mix <- abs(x_mix - median(x_mix))
```

```
mode_sensitivity_deviation <- do.call(
  rbind,
  lapply(
    c(0.75, 1, 1.25),
    function(multiplier) {
      estimate_modes(
        dev_mix,
        bandwidth_multiplier = multiplier
      )
    }
  )
)

print(mode_sensitivity_deviation)
```