Academic Journal of Applied Mathematical Sciences
ISSN(e): 2415-2188, ISSN(p): 2415-5225
Vol. 7, Issue. 3, pp: 167-178, 2021
URL: https://arpgweb.com/journal/journal/17
DOI: https://doi.org/10.32861/ajams.73.167.178
Academic Research Publishing Group
Original Research
Open Access

Auto-Transformations of the Probability Density Functions

Alexander Harin
Modern University for the Humanities, Moscow, Russia
Email: aaharin@gmail.com

Article History
Received: 11 June, 2021
Revised: 13 July, 2021
Accepted: 25 July, 2021
Published: 29 July, 2021

Abstract
The two goals of the present article are: 1) To define transformations (named here as auto-transformations) of the probability density functions (PDFs) of random variables into some similar functions having smaller sizes of their domains. 2) To research and outline basic features of these auto-transformations of PDFs. Particularly, auto-transformations from infinite to finite domains are analyzed. The goals are caused by the well-known problems of behavioral sciences.
Keywords: Variance; Noise; Bias; Measurement; Utility; Prospect theory.

1. Introduction
1.1. Preliminaries
The present article is devoted to modifications of the probability density functions (PDFs) of continuous random variables (r.v.s).

The next stage of this work will be a mathematical description of influence of noise (and other causes that can cause dispersion of data) on measurement data near the boundaries of the measurement intervals.

The article develops works [1-3], where biases of the expectations of random variables are revealed near the boundaries of finite intervals, and an applied mathematical method (approach) and model are created. The method and model have qualitatively explained some well-known generic problems of behavioral economics, decision theories, and the social sciences. The examples of such problems are the underweighting of high and the overweighting of low probabilities, risk aversion, the Allais paradox, risk premium, etc.

1.2. Main Goals of the Article
The main goals of the article are to determine transformations of probability density functions of continuous r.v.s from some larger (especially from infinite) domains into some smaller ones and to outline general properties of these transformations (named here as auto-transformations or ATs).

The transformations can be performed from infinite into half-infinite, from infinite and half-infinite into finite, and also from larger finite into smaller finite domains (intervals). These transformations of probability density functions are the first example of such transformations and can be extended in future to the general case of the real-valued random variables.

Suppose PDFs whose domains are infinite. The initial point, that has motivated this goal, is a question how could such or similar PDFs be modified if their domains were half-infinite or finite. The modifications near the boundaries of the intervals (domains) are of particular importance.

1.3. Review of the Literature
Bounds for moments and functions are considered in many works, see, e.g., Cacoullos [4]; Chernoff [5]; Madansky [6]; Moriguti [7]. Situations considered in Bhatia and Davis [8]; Dokov and Morton [9]; Pinelis [10]; Prekopa [11]; Sharma and Bhandari [12]; Sharma, *et al.* [13] are, in the mathematical aspects, the most similar to those analyzed here.

Mathematical aspects of behaviour sciences, utility and prospect theories are considered in a number of works, see, e.g., Biagini and Frittelli [14]; Choulli and Ma [15]; von Neumann and Morgenstern [16]; Steingrimsson and Luce [17]. Similar aspects are kept in mind in the present article as well. Works Aczel and Luce [18] and Steingrimsson and Luce [17] constitute one of initial points for the considerations of Harin [1]; [2], and of this article.

Qualitative influences of noise are considered in some works. For example, stabilization and synchronization by noise are considered in a number of works, see, e.g., Applebaum and Siakalli [19]; Arnold, *et al.* [20]; Barbu [21]; Cerrai [22]; Flandoli and Gess [23]; Hua, *et al.* [24]. A noise as a possible cause of some periodic behavior is considered in, e.g., Giacomin and Poquet [25]; Scheutzow [26]. The forbidden zones that will be considered in the next section can be treated as some qualitative influence of noise (and of other sources of dispersion) upon the expectations of data.

2. The Problematic That Has Motivated This Article

2.1. Problems of Behavioral Sciences

A man as an individual actor is a key subject of economics and some other sciences. There are basic problems concerned with the mathematical description of the behavior of a man. They are the most actual in behavioral economics, especially in utility and prospect theories, and also in decision theories, the social sciences and psychology, see, e.g., Kahneman and Thaler [27].

Examples of the problems are the underweighting of high and the overweighting of low probabilities, risk aversion, the Allais paradox, risk premium, the four-fold pattern paradox, etc.

The essence of the problems consists in biases of preferences and decisions of a man in comparison with predictions of the probability theory. These biases are maximal near the boundaries of the probability scale.

2.2. Existence Theorem. Applied Mathematical Method and Model

Bounds on the expectations of r.v.s (or bounding inequalities or, in other words, forbidden zones for the expectations of r.v.s) that take on values in a finite interval are considered in, e.g., Harin [1], and Harin [2].

Suppose a set $\{X_i\}$, $i = 1, \dots, n$, of random variables X_i whose values lie within an interval $[a, b]$. If $0 < (b-a) < \infty$ holds for $[a, b]$, and if the condition $\sigma^2{}_{\mathrm{i}} \geq \sigma^2{}_{\mathrm{min}} > 0$ of the minimal variance $\sigma^2{}_{\mathrm{min}}$ holds for their variances $\sigma^2{}_{\mathrm{i}}$, then their expectations μ_i are separated from the boundaries a and b of the interval $[a, b]$ by forbidden zones of the non-zero width

$$a < \left(a + \frac{\sigma^2{}_{\mathrm{min}}}{b-a} \right) \leq \mu_i \leq \left(b - \frac{\sigma^2{}_{\mathrm{min}}}{b-a} \right) < b. \tag{1}$$

That is, the theorem proves the existence of forbidden zones of a certain non-zero width for the expectations of measurement data under the condition that the variance of the data is not less than some minimal variance of a non-zero value.

Such non-zero minimal variances of measurement data can be caused, e.g., by noise, imprecision, errors, incompleteness, various types of uncertainty. Noise can be one of usual and important causes.

This theorem has led to an applied mathematical method (approach) of the biases of the expectations. The method supposes that the subjects (peoples) make their choices (at least to a considerable degree) as if there were some biases of the expectations for games.

The first stage of the approach is a concept of qualitative mathematical models. The models are concentrated on cases when the signs of the presupposed differences (for the choices of subjects), that are required to obtain the observed data, are not equal to the signs of the real differences between the expectations for the uncertain and sure games.

These general models enable formal solutions of the considered paradoxes, but some problems were aroused. In particular, the limits of their applicability need additional research (see the next subsection).

The first step of realization of the concept is a special qualitative mathematical model (see, e.g., Harin [3]) designed for the cases when the expectations for the sure and uncertain games are equal to each other. It had allowed working reliably within the above limits of the applicability, and the model was successfully and uniformly applied in more than one domain of experimental data.

Nevertheless, the above general problems of the limits of applicability of the models still stand over.

2.3. Need for Further Research

One of main questions concerning these theorem of existence of forbidden zones and method of the biases of the expectations is to analyze the widths of the forbidden zones for various PDFs. In particular, the width r_μ of the forbidden zones for the expectations can be determined from (1) as

$$r_\mu \equiv \frac{\sigma^2{}_{\mathrm{min}}}{b-a}. \tag{2}$$

We see that this width can evidently be neglected with respect to the minimal standard deviation (SD) σ_{min} when $\sigma_{min} \rightarrow 0$.

Suppose the level of the noise (or other source of the variance of the data) tends to zero. Hence, the variance of the data also tends to zero. Then the width of the forbidden zones tends to zero as well. One can estimate also the ratio of this width to the SD. This ratio represents, in particular, the decreasing of the width with respect to the decreasing of the SD (when the noise level decreases).

Therefore expression (2) leads to an important consequence: "When the minimum of the standard deviation tends to zero then the ratio of the width of the forbidden zones to the minimum of the standard deviation tends to zero as well." It means (approximately) that the widths of the forbidden zones decrease much faster than the SD does. This means that, at least in some cases, the widths of such forbidden zones can be neglected at low levels of the noise (and other sources of the forbidden zones).

So, the practical problem is whether the considered forbidden zones and biases for results of experiments can be neglected at low levels of their causes (such as noise). This leads to the future goals of the present research.

There is an additional problem as well. The above theorem has been proven for finite intervals. However, a number of important PDFs are defined on infinite and half-infinite intervals. So there is a need for a tool for estimations or at least for hypotheses and assumptions about these or similar important PDFs. This leads to the main goal of the present article.

3. Auto-Transformations. Main Definitions

3.1. Auto-Transformations as a Tool for Modifications and Hypotheses

The domains of many important probability density functions are infinite. The normal distribution can be mentioned as one of the most important examples. Questions can arise about how such or similar probability density functions could be modified if their domains were half-infinite or finite.

Generally, questions can arise about how PDFs can be modified when their domains are modified from larger to smaller sizes of the domains. These questions may be relevant in particular in connection with possible expansions and generalizations of the results of, e.g., Harin [1] and Harin [2] obtained for finite intervals.

Such questions can be too hard to be solved immediately and exactly. Therefore a tool can be proposed here to modify probability density functions, and put forward hypotheses and make assumptions about such modified functions. It will modify mainframe probability density functions into transformed ones those will be, depending on parameters of the transformations, similar to the mainframe functions to a greater or lesser degree.

This tool can be named as auto-transformations of probability density functions.

3.2. Main Definitions, Assumptions, and Abbreviations

3.2.1. Preliminary Notes

For the purposes of the present article and research, following definitions, assumptions, and abbreviations are used here.

Random variables are abbreviated to **r.v.s**.

For the sake of simplicity and to not increase excessively the volume of the present first article devoted to this item, only continuous and piecewise continuous probability density functions are considered here as a rule.

The standard deviation is abbreviated to **SD**. Probability density functions are abbreviated to **PDFs**. The Heaviside step function is referred to as $\boldsymbol{\theta(x)}$: $\theta(x|x < 0) = 0$, $\theta(x|x \geq 0) = 1$. The symbol Δ is used both without round brackets as, e.g., Δf and with them as, e.g., $\Delta_{gen}(E(X))$.

For the purposes of this article, a **finite boundary** of an interval is defined as the boundary which coordinate is finite and an **infinite boundary** is defined as the boundary which coordinate is infinite.

Consider the domain of the probability density function of a random variable. Suppose this domain is the infinite or a half-infinite (or a finite) interval. Further, this interval is referred to as a mainframe interval (**MF-interval** or **MFI**). The corresponding probability density function of this variable is referred to as a mainframe PDF (**MF-PDF** or **MFF** or $\boldsymbol{f}$).

3.2.2. Auto-Transformations. General Definitions

Auto-transformations are abbreviated to **AT**s.

A half-infinite or finite part of the infinite MF-interval and a finite part of a half-infinite (or finite) MF-interval are defined as an **interval of auto-transformation** or an **auto-transformation interval** (**AT-interval** or **ATI**) under the following three determining conditions:

1. The AT-interval contains (at least at its boundary) at least one of the main features of the mainframe probability density function such as the expectation, median or mode.
2. The part of the MF-PDF that is situated in the ATI is identically mapped into the ATI (or, in other words, is unchanged).
3. The part or parts of the MF-PDF that lie outside the ATI are mapped into ATI. Types or modes of this mapping can be chosen in accordance with the goals and conditions of the mapping.

So the summarized transformation of the MF-function is fully enclosed in the AT-interval and consists of the following two parts:

1) the part that is identically mapped (or, in other words, is unchanged);
2) part(s) that is (are) mapped from outside of the ATI into inside of the ATI (usually they are denoted as **out-ATI part(s)**).

This resultant transformed MF-function is referred to as an auto-transformed probability density function (**AT-PDF** or **ATF**).

3.2.3. Auto-Transformations. Specific Definitions

The boundaries of mainframe intervals are usually denoted with a subscript "MF" as, e.g., a_{MF} or b_{MF}. They can be either infinite or finite.

The boundaries of AT-intervals are denoted either without any subscript as, e.g., "a" or "b", or with the subscript "AT" as e.g., "a_{AT}" or "b_{AT}".

Finite boundaries of AT-intervals those are the nearest to the expectations of the MF-PDFs are often marked by a subscript "boundary" and referred to as $a_{boundary}$ and/or $b_{boundary}$. Such left finite boundaries are often determined as $a_{boundary} = 0$. So the minimal distance between the expectation of the variable and the nearest left boundary $a_{boundary}$ of the interval is referred to as $min(|E(X)\text{-}a_{boundary}|)$ or simply $|E(X)\text{-}a_{boundary}|$ or simply $E(X)$.

Finite AT-intervals are usually referred to as $[a, b]$, where $-\infty < a < b < +\infty$ (while a_{MF} and b_{MF} usually can be supposed as infinite).

The medians of MF-functions are referred to as m_X or m.

Normal-like PDFs are defined as symmetric PDFs with non-increasing sides. In other words, for a normal-like PDF f we have:

$$f(E(X)+x_c)=f(E(X)-x_c)$$

and if $|x_c-E(X)|\le|x_d-E(X)|$, then $f(x_c)\ge f(x_d)$.

3.2.4. Types of Auto-Transformations and Situations

A **contiguous** situation is defined as the situation when a finite boundary of the MF-interval touches a finite boundary of the AT-interval.

A hypothetical **reflection** situation is defined as the situation when the MF-PDF f is modified to the AT-PDF f_{refl} such that a part of f is reflected (as an addition) with respect to some point (dot) $d_{refl} \equiv d_{reflection}$ of the reflection. This point can coincide with the median m_X of MF-PDF. In this case m_X and d_{refl} are often supposed to be $d_{refl} = m_X = 0$ for the left point d_{refl} and

$$f_{refl}(x)=\theta(x)[f(x)+f(-x)].$$

The hypothetical reflection situation is, in a sense, similar to the reflection of a wave of light from a mirror.

If an auto-transformation interval is finite, especially if the mainframe interval is infinite or at least semi-infinite, then the reflection can be multiple. This can occur also in similar cases but in the absence of reflection. In all these cases the AT can be referred to as the **repeated auto-transformation** or **multiple AT** or **many-fold AT** or **two-mirror AT**. Otherwise (and as a rule in this article), ATs can be referred to as **one-fold AT**s.

A hypothetical **adhesion** situation is modified from the reflection one. The reflected part of the MF-PDF is "adhered" (as an addition) to the boundary $d_{refl} = a_{boundary} = 0$. The MF-PDF is transformed to the function of the mixed type, such that its discrete part is equal to the integral of the reflected part of the MF-PDF. In other words, the reflected part of the mainframe PDF is adhered to the point $x = 0$. In Appendix 1 the necessary statement is proven that adhesion ATs provide the minimal distances from the boundaries of ATIs to the expectations of AT-PDFs.

ATs that transform the out-ATI parts of the MF-PDFs into linear additions to the AT-PDFs are referred to as **linear AT**s. ATs that transform the out-ATI parts of the MF-PDFs into stepwise additions to the AT-PDFs are referred to as **stepwise ATs** (the both linear and stepwise ATs can be many-fold).

Auto-transformations are referred to as **boundary-matched AT**s if they transform the out-ATI parts of the MF-PDFs into additions to auto-transformed PDFs such that those values are equal at the corresponding boundaries to the values of the MF PDFs. When those values are equal to the means of the MF PDFs, then these ATs are referred to as **mean-matched AT**s.

For the purposes of the present research, one can denote a **vanishing auto-transformation** as the auto-transformation such that the additional part Δf of the AT-PDF starts at the corresponding boundary of the ATI with a non-zero value and vanishes to the zero value.

Vanishing ATs can be of convex, triangle, concave, and mixed types. Stepwise ATs can be considered in a sense as the upper (unattainable) limit cases for convex vanishing ATs.

In all situations the non-modified standard deviations of the non-modified mainframe probability density functions are used to analyze and estimate the auto-transformed probability density functions.

Usually, $\hbar$ denotes the value (**h**eight) of the probability density function (***h*** will be used in future for the probability mass function) and ***l*** denotes the length. Usually, the index "1" denotes the center or top of a function, that is $\hbar_1 \equiv \hbar_{centre}$ and $l_1 \equiv l_{centre}$. The index "2" denotes the side or tail or bottom of a function, that is $\hbar_2 \equiv \hbar_{side} \equiv \hbar_{tail}$ and $l_2 \equiv l_{side} \equiv l_{tail}$.

Let us denote the **means** of the mainframe PDFs within the AT-intervals as $f_{mean} = f_{MF.ATI.mean}$. They are equal to

$$f_{mean} \equiv \frac{1}{b_{AT}-a_{AT}}\int_{a_{AT}}^{b_{AT}} f_{MF}(x)dx.$$

These means can be used as objective reference points for various purposes.

Auto-transformations can be used as a basis for approximation, modelling etc. Approaches and first approximation models that use auto-transformations can be referred to as AT-approaches and AT-models.

3.3. Possible Generalizations

Similar auto-transformations can be evidently used for functions beyond the scope of the theory of probability.

4. Auto-Transformations. General Considerations

4.1. Modifications of Out-ATI Parts

Let us denote the norms of f on the out-ATI parts as

$$\delta_{out.left} \equiv \int_{a_{MF}}^{a_{AT}} f(x)dx \equiv \int_{a_{MF}}^{a} f(x)dx$$

and

$$\delta_{out.right} \equiv \int_{b_{AT}}^{b_{MF}} f(x)dx \equiv \int_{b}^{b_{MF}} f(x)dx$$

and denote

$$\delta_{out} \equiv \delta_{out.left} + \delta_{out.right}.$$

Let us consider some possible cases of modifications of the out-ATI parts of the mainframe probability density functions into the auto-transformed probability density functions.

4.1.1. Non-Increasing Auto-Transformations

Auto-transformations that transform the out-ATI parts of the MF-PDFs into the additions to the AT-PDFs (that are added to the unchanged parts of the MF-PDFs in the AT-intervals to constitute the resulting AT-PDFs) that do not increase in the directions from the corresponding boundaries to the middles of the AT-intervals (et seq.) are referred to as **non-increasing ATs**.

Non-increasing auto-transformations correspond to intuitive condition that the auto-transformed out-ATI part should contribute near the boundary that is the nearest to this out-ATI part, at least, not less than near the opposite boundary. Stepwise ATs belong evidently to non-increasing ATs.

Non-increasing auto-transformations are the main type of ATs considered in the present article.

4.1.2. Full-ATI Auto-Transformations. Uniform Auto-Transformations

ATs that transform the out-ATI parts of the MF-PDFs into the total AT-intervals are referred to as **full-ATI ATs** or **filling ATs**.

Suppose that the out-ATI parts are uniformly mapped into the ATI, e.g., by means of some uniform coefficient or addition.

Uniform multiplication (including increasing ATs). The MF PDF that is situated in the AT-interval is multiplied by the uniform rising coefficient

$$f_{AT}(x) = \frac{f_{MF}(x \mid x \in [a,b])}{1-\delta_{out}}.$$

Uniform addition. The out-ATI parts are added uniformly to the unchanged part of the MF-PDF and the transformed PDF is

$$f_{AT}(x) = f_{MF}(x \mid x \in [a,b]) + \frac{\delta_{out}}{b-a}.$$

Uniform addition auto-transformations evidently belong to stepwise (and non-increasing) full-ATI ATs. In Appendix 2, statements are proven that uniform addition ATs provide the maximal distances from the expectations of the AT-PDFs to the sides of reflection of the ATIs (and the minimal distance to their opposite sides) among the non-increasing auto-transformations.

4.1.3. Boundary-Matched and Mean-Matched Auto-Transformations

Various types of boundary-matched (BM) and mean-matched (MM) ATs can be considered. Here are two examples of linear non-increasing ATs.

The length of, e.g., left stepwise part of a stepwise MM AT equals

$$l_{stepwist.mean.left} = \frac{\delta_{out.left}}{f_{mean}}.$$

The length of, e.g., left triangle part of a triangle BM AT equals

$$l_{triangle.bound.left} = \frac{2\delta_{out.left}}{f(a)}.$$

Trapezium auto-transformations are evidently an intermediate case between the stepwise and triangle ones. Stepwise, trapezium, and triangle linear ATs constitute a family of linear and piecewise linear auto-transformations.

Necessary (see subsection 4.3) auto-transformations ensure the necessary condition (see proofs in Appendix 4) to provide mean-matched one-fold non-increasing (linear and nonlinear) auto-transformations.

Suppose non-increasing auto-transformations, e.g., mean-matched ones. Let us consider out-ATI additions to the AT-PDFs. If the AT-intervals are sufficiently short, then the non-increasing mean-matched ATs cannot be one-fold and are many-fold. Let us increase the lengths of the ATIs. Then the transformed out-ATI parts will be stepwise mean-matched one-fold. Further they will be consequently stepwise, trapezium, and triangle (and also convex, linear, and concave).

4.2. About the Distances from the Expectations to the Boundaries

One of the main goals of the auto-transformations is to estimate possible distances from the expectations of the PDFs to the boundaries of the intervals. In turn, the estimations of these distances in real economic measurements are the main ultimate goal of the present research. At that the auto-transformations are the tool for such estimations.

Appendix 1 provides a simple but necessary proof that the distances from the expectations of the AT-PDFs to the nearest boundaries of the AT-intervals are the minimal for the adhesion auto-transformations.

For the reflected auto-transformations one can prove (see Appendix 3) that the distances from the expectations of the AT-PDFs to the nearest boundaries of the AT-intervals are the minimal when the points of reflection coincide with the medians of the MF-PDFs.

4.3. Necessary Auto-Transformations

An auto-transformation of a PDF is referred to as a **necessary AT** or **Norm-necessary AT** or **N-necessary AT** if

$$\int_{a}^{b} f(x)dx \equiv \int_{a_{AT}}^{b_{AT}} f(x)dx \geq \frac{1}{2} \int_{a_{MF}}^{b_{MF}} f(x)dx = \frac{1}{2},$$

or, equivalently, if

$$\int_{a_{MF}}^{b_{MF}} f(x)dx - \int_{a_{AT}}^{b_{AT}} f(x)dx \equiv \int_{a_{MF}}^{a_{AT}} f(x)dx + \int_{b_{AT}}^{b_{MF}} f(x)dx \leq \frac{1}{2}.$$

That is the norm of the unchanged part of the MF-PDF is not less than $1/2$. That is the difference between the norms calculated for the whole MF-PDF and its unchanged part is not more than $1/2$.

Necessary auto-transformations ensure the condition that provides some properties of transformed PDFs (see the proofs of the corresponding statements in Appendix 4) including the following.

Necessary stepwise one-step auto-transformations ensure that if the values of the additions are $\Delta f \geq f_{mean}$, then the lengths of the stepwise one-step additions to the AT-PDFs are not more than the lengths of the AT-intervals. That is such ATs belong to the one-fold auto-transformations.

Necessary uniform addition ATs ensure that the values of the uniform additions are $\Delta f \geq f_{mean}$.

Necessary symmetric triangle ATs ensure for symmetric PDFs that:

a) if the boundary value (the height) of the AT addition is $\Delta f_{bound} \geq f_{mean}$ then the length of the triangle is not more than the length of the AT-interval;

b) vice versa, if the length of the triangle is not less than the length of the AT-interval (that is if the auto-transformation is at least filling), then its boundary value (its height) is $\Delta f_{bound} \leq f_{mean}$.

Necessary symmetric vanishing filling ATs ensure for symmetric PDFs that if $\Delta f_{bound} \geq f_{mean}$ then such ATs are concave (at least partially), except the triangle ATs in the case of $\Delta f_{bound} = f_{mean}$ and $\Delta_{out} = 1/2$.

4.4. Sufficient Auto-Transformations

An auto-transformation of a PDF f is referred to as a **sufficient auto-transformation** or **Norm-sufficient AT** or **N-sufficient AT** if

$$\int_{a_{MF}}^{a} f(x)dx + \int_{b}^{b_{MF}} f(x)dx = \delta_{out} << \int_{a_{MF}}^{b_{MF}} f(x)dx = 1,$$

that is the difference between the norms calculated for the MF-PDF and unchanged part of the AT-PDF is negligibly small in comparison with the norm calculated for the MF-PDF. That is the difference is much less than unit.

Sufficient ATs evidently belong to necessary ATs.

For the normal distribution, the auto-transformation interval that corresponds to the "three-sigma rule" can be used as a sufficient AT-interval.

Generally, a sufficient auto-transformation (and auto-transformation interval) can be referred to as the sufficient one with respect to some parameter, if the difference between the values of this parameter calculated for the MF-PDF and AT-PDF is negligibly small in comparison with the value of this parameter calculated for the whole MF-PDF. For example:

An auto-transformation will be referred to as the V-sufficient one (where "V" denotes Variation) if the part(s) of the variation of the MF-PDF that is (are) calculated outside the AT-interval is (are) much less than the whole variation of the mainframe PDF.

An auto-transformation will be referred to as the SD-sufficient one (where "SD" denotes the Standard Deviation) if the part(s) of the standard deviation that is (are) calculated for the mainframe PDF outside the AT-interval is (are) much less than the standard deviation calculated for the whole MF-PDF.

Sufficient auto-transformations are evidently the most prospective auto-transformations for hypotheses and estimations.

5. Conclusions

The present article proposes a new tool to modify and transform the probability density functions of random variables into similar functions having smaller sizes of their domains.

The tool can be used, e.g., for hypotheses and assumptions. It is named as the auto-transformations (ATs) of PDFs. It can be treated as a first step to general ATs of real-valued random variables.

Evidently, the considerations of the article can be developed also to transformations of other functions beyond the scope of the probability theory.

The ultimate goal of this research is a mathematical description of influence of noise (and other causes that can cause dispersion of data) on measurement data near the boundaries of the measurement intervals.

The two main results of the present article are:

First result. The concept of the auto-transformations is proposed.

Second result. Some general properties of the auto-transformations are considered and the corresponding statements are proven.

Appendix 1. Adhesion

Let us consider a piecewise continuous mainframe probability density function $f_{MF} \equiv f$ that is defined on a mainframe interval whose left boundary is anyway finite or infinite. Let us consider an auto-transformation such that the coordinate of the left boundary of the AT-interval (that is the left dot of reflection) is $d_{left} \equiv a_{AT} = a$. For the left out-ATI part we have

$$\delta_{out.left} \equiv \int_{-\infty}^{d_{left}} f_{MF}(x)dx \quad \text{or} \quad \delta_{out.left} \equiv \int_{a_{MF}}^{a_{AT}} f_{MF}(x)dx .$$

Let us compare two cases:

1) The first is the case when the out-ATI part is adhered to the left boundary a_{AT} of the AT-interval. The contribution of this adhered part $\Delta_{adh}(E(X_{AT}))$ to the expectation of the AT-PDF can be written as

$$\Delta_{adh}(E(X_{AT})) = \delta_{out.left} \times d_{left} .$$

2) The second is the general case when the out-ATI part is mapped into the ATI as a general addition $\Delta_{gen}(f(x))$ to the MF-PDF, at least not only at the boundary a_{AT} of the ATI. Its contribution to $E(X_{AT})$ is

$$\Delta_{gen}(E(X_{AT})) = \int_a^b x\Delta_{gen}(f(x))dx,$$

where

$$\int_a^b \Delta_{gen}(f(x))dx = \delta_{out.left} .$$

The multiplier x (under the integral sign) is not less than d_{left}. Therefore, the contribution of this general part to the expectation of the auto-transformed PDF is not less than

$$\Delta_{gen}(E(X_{AT})) = \int_a^b x\Delta_{gen}(f(x))dx \geq \int_a^b d_{left}\Delta_{gen}(f(x))dx = d_{left}\int_a^b \Delta_{gen}(f(x))dx$$

$$= d_{left} \times \delta_{out.left} = \Delta_{adh}(E(X_{AT})).$$

So, the left boundary adhesion auto-transformation provides the minimal expectation of the auto-transformed PDF. Evidently, the right boundary adhesion AT provides the maximal expectation of the AT-PDF.

Appendix 2. Uniformity

A2.1. Uniformity for Contiguous Situations

Suppose a continuous or piecewise continuous PDF f such that: the domain of f is an interval $[0; 2l]$, $E(X) = l$, $f(l + y) = f(l - y)$ and if $|x_2 - l| \geq |x_1 - l|$ than $f(x_1) \geq f(x_2)$ (that is f is normal-like).

Due to $E(X)$ is predetermined, the ratio $E(X)/SD$ depends only on SD.

The maximal value of $f(x)$ can be denoted as $max(f(x)) = f(l) \equiv \hbar$ and the expression for the variance can be rewritten as

$$Var(X) = \int_0^{2l} (x-l)^2 f(x)dx = \int_0^{2l} (x-l)^2\{\hbar - [\hbar - f(x)]\}dx$$

$$= l^2 \frac{2\hbar l}{3} - \int_0^{2l} (x-l)^2[\hbar - f(x)]dx .$$

The members $(x\text{-}l)^2$ and $[\hbar\text{-}f(x)] \equiv [max(f(x))\text{-}f(x)]$ are nonnegative. Hence the variance is maximal when $[max(f(x))\text{-}f(x)] = 0$. That is it is maximal when $f(x) = max(f(x)) = \hbar$. This condition implies the normalization equality $2\hbar l = 1$. Under this condition the standard deviation is not more than the well-known value

$$SD = \sqrt{l^2 \frac{2\hbar l}{3}} = \frac{l}{\sqrt{3}}.$$

So the ratio $|E(X)\text{-}a_{boundary}|/SD$ is equal to

$$\frac{E(X) - a_{boundary}}{SD} \geq l \frac{\sqrt{3}}{l} = \sqrt{3}$$

and we have

$$E(X) - a_{boundary} > SD.$$

So, in the general case, for the contiguous situation, for any continuous or piecewise continuous normal-like PDF with finite domain, the minimal distance from its expectation to the nearest boundary cannot be much less than its standard deviation.

Moreover, this minimal distance is more than the standard deviation.

A2.2. Non-Increasing and Uniform Addition Auto-Transformations

Let us consider non-increasing, including uniform addition, one-fold ATs. Consider a piecewise continuous MF-PDF f_{MF} and the AT-PDF $f \equiv f_{AT}$.

Left auto-transformations. Suppose that f_{MF} is defined on the mainframe interval whose left boundary a_{MF} is anyway finite or infinite and whose right boundary b_{MF} is finite. Suppose the auto-transformation interval is $[a_{AT}, b_{AT}]$ where $a \equiv a_{AT} > a_{MF}$ and $b \equiv b_{AT} = b_{MF}$.

The left out-ATI part $\delta_{out.left}$ can be defined similarly to Appendix 1.

The left out-ATI contribution to the expectation of the AT-PDF is

$$\Delta_{left}(E(X_{AT})) = \int_{a_{AT}}^{a_{MF}} x\Delta_{left}(f_{AT}(x))dx$$

under the condition

$$\int_{a_{AT}}^{a_{MF}} \Delta_{left}(f_{AT}(x))dx = \delta_{out.left}.$$

The boundary left value of the addition $\Delta_{left}(f(x))$ to the MF-PDF is $\Delta_{left}(f(a_{AT}))$. Due to the non-increasing character of the AT,

$$\Delta_{left}(f(x)) \leq \Delta_{left}(f(a_{AT})).$$

That is in any point of the AT-interval the addition $\Delta_{left}(f(x))$ is not more than its boundary value.

The minimal possible left value $f(a_{AT})$ for a non-increasing one-fold AT can be obtained from $\Delta_{left}(f_{AT}(a_{AT}))(b_{AT}-a_{AT}) \geq \delta_{out.left}$. It is

$$\Delta_{left}(f_{AT}(a_{AT})) \geq \frac{\delta_{out.left}}{b_{AT} - a_{AT}}.$$

Hence, left uniform addition ATs provide the maximal expectations

$$E(X_{AT}) = \int_{a_{AT}}^{b} xf_{AT}(x)dx = E(X_{MF} \mid x \in [a_{AT}, b_{AT}]) + \Delta_{left}(E(X_{AT}))$$

$$= E(X_{MF} \mid x \in [a_{AT}, b_{AT}]) + \int_{a_{AT}}^{b} x\Delta_{left}(f_{AT}(x))dx$$

$$\leq E(X_{MF} \mid x \in [a_{AT}, b_{AT}]) + \int_{a_{AT}}^{b} x\Delta_{left}(f_{AT}(a_{AT}))dx \qquad .$$

$$= E(X_{MF} \mid x \in [a_{AT}, b_{AT}]) + \Delta_{left}(f_{AT}(a_{AT}))(b_{AT} - a_{AT})$$

for the auto-transformed PDFs.

Right auto-transformations. Considerations for the right ATs lead to the same conclusions as for the left ATs.

So, stepwise, including uniform addition, auto-transformations provide the maximal distances from the expectations for AT-PDFs to the nearest boundaries and the minimal distances from the expectations to the opposite boundaries for non-increasing ATs.

Appendix 3. Reflection. The Minimal Distance

Let us consider reflection auto-transformations and determine conditions for the minimal distances from the expectations of auto-transformed PDFs to the boundaries of the AT-intervals.

Let us consider a piecewise continuous probability density function f. For the sake of generality of consideration, assume that, if f is defined on a finite interval $[a; b]$ or on a semi-infinite interval, then it will be expanded on the infinite interval by supposing that $f(x) \equiv 0$ for $x \in (-\infty, a_{MF})$ and/or $x \in (b_{MF}, \infty)$ respectively.

Let us perform a reflection auto-transformation with respect to an arbitrary finite reflection point (dot) $d \equiv d_{refl}$.

The auto-transformed PDF is

$$f_{AT}(x \,|\, x \geq d) = f_{MF}(x) + f_{MF}(2d - x) \equiv f(x) + f(2d - x).$$

The expectation of the AT-PDF (as a function of d) is

$$E(X_{AT}) \equiv \mu_{AT} = \int_d^{\infty} x f(x)dx + \int_d^{\infty} x f(2d - x)dx.$$

Let us consider the integral on the far right. Let us introduce temporarily a new and formal argument $y = 2d - x$. Then

$$\int_d^{\infty} x f(2d - x)dx = -\int_d^{-\infty} (2d - y) f(y)dy = \int_{-\infty}^{d} (2d - x) f(x)dx$$

and

$$\mu_{AT} = \int_d^{\infty} x f(x)dx + \int_{-\infty}^{d} (2d - x) f(x)dx.$$

The distance from this expectation to the reflection point is

$$\begin{aligned}\mu_{AT} - d &= \int_d^{\infty} x f(x)dx + \int_{-\infty}^{d} (2d - x) f(x)dx - d \\ &= \int_d^{\infty} x f(x)dx - \int_{-\infty}^{d} x f(x)dx + 2d \int_{-\infty}^{d} f(x)dx - d\end{aligned}.$$

Using the median m, the two right terms can be identically transformed to

$$2d \int_{-\infty}^{m} f(x)dx - d + 2d \int_m^d f(x)dx = 2d \int_m^d f(x)dx.$$

So the distance from this expectation to the reflection point is

$$\mu_{AT} - d = \int_d^{\infty} x f(x)dx - \int_{-\infty}^{d} x f(x)dx + 2d \int_m^d f(x)dx.$$

Let us use another reflection point, e.g., $d_{refl.2} = d + \varepsilon$. Then

$$\mu_{AT} - d - \varepsilon = \int_{d+\varepsilon}^{\infty} x f(x)dx - \int_{-\infty}^{d+\varepsilon} x f(x)dx + 2d \int_m^{d+\varepsilon} f(x)dx.$$

The difference between the distances is equal to

$$\begin{aligned}&[\mu_{AT}(d+\varepsilon) - d - \varepsilon] - [\mu_{AT}(d) - d] \\ &= \int_{d+\varepsilon}^{\infty} x f(x)dx - \int_{-\infty}^{d+\varepsilon} x f(x)dx + 2(d+\varepsilon) \int_m^{d+\varepsilon} f(x)dx \\ &- \int_d^{\infty} x f(x)dx + \int_{-\infty}^{d} x f(x)dx - 2d \int_m^d f(x)dx \\ &= -\int_d^{d+\varepsilon} x f(x)dx - \int_d^{d+\varepsilon} x f(x)dx + 2d \int_d^{d+\varepsilon} f(x)dx + 2\varepsilon \int_m^{d+\varepsilon} f(x)dx \\ &= 2\int_d^{d+\varepsilon} (d - x) f(x)dx + 2\varepsilon \int_m^{d+\varepsilon} f(x)dx\end{aligned}.$$

Let us consider two cases: 1) $d > m$ and 2) $d+\varepsilon < m$, under the conditions $f(x) \geq 0$ and $\varepsilon \geq 0$.

1) $d > m$. The result is

$$\begin{aligned}&[\mu_{AT}(d+\varepsilon) - d - \varepsilon] - [\mu_{AT}(d) - d] \\ &= 2\int_d^{d+\varepsilon} (d + \varepsilon - x) f(x)dx + 2\varepsilon \int_m^d f(x)dx\end{aligned}.$$

The both integrals are non-negative. So the difference between the distances is non-negative as well.

2) $d + \varepsilon < m$. The result is

$$[\mu_{AT}(d+\varepsilon)-d-\varepsilon]-[\mu_{AT}(d)-d]$$
$$=2\int_{d}^{d+\varepsilon}(d-x)f(x)dx-2\varepsilon\int_{d+\varepsilon}^{m}f(x)dx\,.$$

The both integrals (taking into account the signs before them) are non-positive. So the difference between the distances is non-positive as well.

So in these two cases, when d is shifted, e.g., to the right, the distance from the expectation to the boundary is changed in two different ways. In the first case, when $d > m$, that is when d moves from m, the change is non-negative. In the second case, when $d + \varepsilon < m$, that is when d moves toward m, the change is non-positive.

The minimal distance from the expectation to the boundary is attained when the median belongs to $[d; d + \varepsilon]$. If $\varepsilon \rightarrow 0$ then we obtain

$$d_{refl.\min}=m\,. \tag{3}$$

So the reflection point $d_{refl.min}$ of the reflection auto-transformation that provides the minimal distance from the expectation of the auto-transformed PDF to the boundary of the AT-interval is equal to m.

Appendix 4. Necessary Auto-Transformations

Let us consider necessary auto-transformations (that is auto-transformations such that $\delta_{out} \le 1/2$) of piecewise continuous mainframe probability density functions $f_{MF} \equiv f$ into finite AT-intervals $[b_{AT}, a_{AT}] \equiv [b, a]$.

Let us refine that a **full-ATI AT** or **filling AT** is an AT such that the AT-addition to the MF-PDF is non-zero in the entire ATI (at least except of no more than one point). That is the AT-addition fills the entire AT-interval.

First of all, necessary auto-transformations ensure by their definition (for arbitrary mainframe PDFs f_{MF}) that

$$f_{mean}\equiv\frac{1}{b_{AT}-a_{AT}}\int_{a_{AT}}^{b_{AT}}f_{AMF}(x)dx\ge\frac{1}{2(b_{AT}-a_{AT})}\,.$$

A4.1. One-Step Stepwise Auto-Transformations

Let us consider, e.g. left, stepwise one-step auto-transformations and denote the coordinates of the right ends of the steps as $b_{step} \equiv b_s$. Without taking into account possible additions from right auto-transformations, we obtain the auto-transformed PDFs $f_{AT}(x) = f_{MF}(x) + \Delta f_{left}$ for $x\in[a,b_s]$, where $\Delta f_{left} = \delta_{out.left}/(b_s-a)$.

Consider the length b_s-a of the step of a necessary one-step AT.

Statement A.4.1. If, for a necessary, e.g. left, one-step auto-transformation, the value of the addition Δf_{left} is not less than f_{mean}, then the length b_s-a of the step is not more than the length $b-a$ of the AT-interval. In other words this auto-transformation is one-fold.

Proof. The length of the left one-step out-ATI addition is

$$b_s-a=\frac{\delta_{out.left}}{\Delta f_{left}}\le\frac{\delta_{out.left}}{f_{mean}}\le\frac{1}{2}2(b-a)=b-a\,. \qquad \square$$

So, this auto-transformation is one-fold.

A4.2. Uniform Addition Auto-Transformations

Consider the value $\Delta f_{un.left}$ of, e.g., the left addition of a necessary uniform addition auto-transformation to AT-PDF.

Statement A.4.2. If a necessary uniform addition auto-transformation is a filling one, then the value of its side addition is not more than f_{mean}.

Proof. The value of, e.g., the filling left addition $\Delta f_{un.left}$ to AT-PDF is

$$\Delta f_{un.left}=\frac{\delta_{out.left}}{b-a}\le\frac{\delta_{out}}{b-a}\le f_{mean}\,. \qquad \square$$

A4.3. Vanishing Symmetric Auto-Transformations

Remember that the vanishing auto-transformations are referred to as the ATs such that the addition parts Δf of the AT-PDFs start at the corresponding boundaries of ATIs with non-zero values and end with the zero values.

Vanishing ATs can be of convex, triangle, concave, and mixed types. The above stepwise one-step ATs can be considered in a sense as the upper (unattainable) limit cases for convex vanishing ATs.

Symmetric auto-transformations of symmetric PDFs can be defined as ATs such that their AT-intervals are symmetric with respect to the centers of symmetry of the MF-PDFs.

Let us prove three interrelated statements.

Consider the lengths of triangle ATs.

Statement A.4.3.1. If the boundary addition value Δf_{bound} of a necessary symmetric triangle AT of a symmetric PDF is not less than f_{mean}, then the length of the triangle is not more than the length of the AT-interval.

Proof. The length of, e.g., the left triangle b_t-a is

$$b_t - a = \frac{2\delta_{out.left}}{\Delta f_{bound.left}} = \frac{\delta_{out}}{\Delta f_{bound.left}} \le \frac{\delta_{out}}{f_{mean}} \le b - a\,. \qquad \square$$

Consider the boundary addition value of a triangle AT.

Statement A.4.3.2. If the length of the triangle of a necessary symmetric triangle AT of a symmetric PDF is not less than the length of the ATI (if AT is at least filling), then its boundary addition value is not more than f_{mean}.

Proof. The boundary addition value, e.g. $\Delta f_{bound.left}$, is

$$\Delta f_{bound.left} = \frac{2\delta_{out.left}}{b_t - a} = \frac{\delta_{out}}{b_t - a} \le \frac{\delta_{out}}{b - a} \le \frac{1}{2}\frac{1}{b - a} \le f_{mean}\,. \qquad \square$$

Consider the concavity of vanishing ATs for symmetric PDFs.

Statement A.4.3.3. The addition of a necessary vanishing filling symmetric AT of a symmetric MF-PDF, such that $\Delta f_{bound} \ge f_{mean}$, is concave (at least partially) except in the case of both $\delta_{out} = 1/2$ and $\Delta f_{bound} = f_{mean}$.

Proof. Under these conditions, the boundary value Δf_{bound} of the addition to the AT-PDF is, similar to the preceding case,

$$\Delta f_{bound.left} = \frac{2\delta_{out.left}}{b_t - a} = \frac{\delta_{out}}{b_t - a} \le \frac{\delta_{out}}{b - a} \le f_{mean}\,.$$

Therefore, the addition of this AT is evidently concave (at least partially) except in the case of both $\delta_{out} = 1/2$ and $\Delta f_{bound} = f_{mean}$ when it is triangle. $\square$

Note. If the lengths of the triangles of a symmetric triangle AT of a symmetric PDF is equal to b-a, then evidently this AT produces the uniform addition AT. If these lengths are less than b-a, then a certain flat portion of Δf_{AT} will evidently exist in the middle of the AT-interval.

References

[1] Harin, A., 2012. "Data dispersion in economics (II) – Inevitability and consequences of restrictions." *Rev. Econ. Fin.*, vol. 2, pp. 24–36.

[2] Harin, A., 2015. "General bounds in economics and engineering at data dispersion and risk." In *Proceedings of the 13th MASR School, no. 13 in Modeling and Analysis of Safety and Risk in Complex Systems.* pp. 105–17.

[3] Harin, A., 2018. "Forbidden zones for the expectation. New mathematical results for behavioral and social sciences." Available: https://mpra.ub.uni-muenchen.de/86650/1/MPRA_paper_86650.pdf

[4] Cacoullos, T., 1982. "On upper and lower bounds for the variance of a function of a random variable." *Ann. Probab.*, vol. 10, pp. 799–809.

[5] Chernoff, H., 1981. "The identification of an element of a large population in the presence of noise." *Ann. Probab.*, vol. 9, pp. 533–36.

[6] Madansky, A., 1959. "Bounds on the expectation of a convex function of a multivariate random variable." *Ann. Math. Stat*, vol. 30, pp. 743–46.

[7] Moriguti, S., 1959. "A lower bound for a probability moment of any absolutely continuous distribution with finite variance." *Ann. Math. Stat*, vol. 23, pp. 286–89.

[8] Bhatia, R. and Davis, C., 2000. "A better bound on the variance." *Am. Math. Mon.*, vol. 107, pp. 353–57.

[9] Dokov, S. P. and Morton, D. P., 2005. "Second-order lower bounds on the expectation of a convex function." *Math. Oper. Res.*, vol. 30, pp. 662–77.

[10] Pinelis, I., 2011. "Exact lower bounds on the exponential moments of truncated random variables." *J. Appl. Probab.*, vol. 48, pp. 547–60.

[11] Prekopa, A., 1990. "The discrete moment problem and linear programming." *Discrete Appl. Math.*, vol. 27, pp. 235–54.

[12] Sharma, R. and Bhandari, R., 2014. "On variance upper bounds for unimodal distributions." *Commun. Stat. A-Theor.*, vol. 43, pp. 4503–13.

[13] Sharma, R., Gupta, M., and Kapoor, G., 2010. "Some better bounds on the variance with applications." *J. Math. Inequal.*, vol. 4, pp. 355–63.

[14] Biagini, S. and Frittelli, M., 2008. "A unified framework for utility maximization problems: an Orlicz space approach." *Ann. Appl. Probab.*, vol. 18, pp. 929–66.

[15] Choulli, T. and Ma, J., 2017. "Explicit description of HARA forward utilities and their optimal portfolios." *Theory Probab. Appl.*, vol. 61, pp. 57–93.

[16] von Neumann, J. and Morgenstern, O., 1944. *Theory of games and economic behavior*. Princeton: Princeton University Press.

[17] Steingrimsson, R. and Luce, R. D., 2007. "Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function." *J. Math. Psychol.*, vol. 51, pp. 29–44.

[18] Aczel, J. and Luce, D. R., 2007. "A behavioral condition for Prelec's weighting function on the positive line without assuming W(1) = 1." *J. Math. Psychol.*, vol. 51, pp. 126–29.

[19] Applebaum, D. and Siakalli, M., 2010. "Stochastic stabilization of dynamical systems using Levy noise." *Stoch. Dynam.*, vol. 10, pp. 509–27.

[20] Arnold, L., Crauel, H., and Wihstutz, V., 1983. "Stabilization of linear systems by noise." *SIAM J. Control Optim.*, vol. 21, pp. 451–61.

[21] Barbu, V., 2009. "The internal stabilization by noise of the linearised Navier-Stokes equation." *Contr. Op. Ca. Va.*, vol. 17, pp. 117–30.

[22] Cerrai, S., 2005. "Stabilization by noise for a class of stochastic reaction-diffusion equations." *Probab. Theory Rel.*, vol. 133, pp. 190–214.

[23] Flandoli, F. B. and Gess, M. S., 2017. "Synchronization by noise." *Probab. Theory Rel.*, vol. 168, pp. 511–56.

[24] Hua, G. M., Liub, X. M., and Songb, M., 2009. "Noise suppresses exponential growth under regime switching." *J. Math. Anal. Appl.*, vol. 355, pp. 783–95.

[25] Giacomin, G. and Poquet, C., 2015. "Noise, interaction, nonlinear dynamics and the origin of rhythmic behaviors." *Braz. J. Prob. Stat.*, vol. 29, pp. 460–93.

[26] Scheutzow, M., 1985. "Noise can create periodic behavior and stabilize nonlinear diffusions." *Stoch. Proc. Appl.*, vol. 20, pp. 323–31.

[27] Kahneman, D. and Thaler, R., 2006. "Anomalies: Utility maximization and experienced utility." *J. Econ. Perspect.*, vol. 20, pp. 221–34.