Journal of Applied Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjas20

Isotonic regression for metallic microstructure data: estimation and testing under order restrictions

Martina Vittorietti, Javier Hidalgo, Jilt Sietsma, Wei Li & Geurt Jongbloed

To cite this article: Martina Vittorietti, Javier Hidalgo, Jilt Sietsma, Wei Li & Geurt Jongbloed (2021): Isotonic regression for metallic microstructure data: estimation and testing under order restrictions, Journal of Applied Statistics, DOI: 10.1080/02664763.2021.1896685

To link to this article: https://doi.org/10.1080/02664763.2021.1896685

Published online: 05 Mar 2021.

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2021.1896685

Taylor & Francis
Taylor & Francis Group

Isotonic regression for metallic microstructure data: estimation and testing under order restrictions

Martina Vittorietti [a,b], Javier Hidalgo[c], Jilt Sietsma [c], Wei Li[c] and Geurt Jongbloed[a]

[a]Department of Applied Mathematics, Delft University of Technology, Delft, Netherlands; [b]Materials Innovation Institute (M2i), Delft, Netherlands; [c]Department of Materials Science and Engineering, Delft University of Technology, Delft, Netherlands

ABSTRACT

Investigating the main determinants of the mechanical performance of metals is not a simple task. Already known physically inspired qualitative relations between 2D microstructure characteristics and 3D mechanical properties can act as the starting point of the investigation. Isotonic regression allows to take into account ordering relations and leads to more efficient and accurate results when the underlying assumptions actually hold. The main goal in this paper is to test order relations in a model inspired by a materials science application. The statistical estimation procedure is described considering three different scenarios according to the knowledge of the variances: known variance ratio, completely unknown variances, and variances under order restrictions. New likelihood ratio tests are developed in the last two cases. Both parametric and non-parametric bootstrap approaches are developed for finding the distribution of the test statistics under the null hypothesis. Finally an application on the relation between geometrically necessary dislocations and number of observed microstructure precipitations is shown.

ARTICLE HISTORY
Received 7 February 2020
Accepted 23 February 2021

KEYWORDS
Isotonic regression; alternating iterative method; likelihood ratio test; bootstrap; order restrictions; geometrically necessary dislocations

Abbreviations

AIM; alternating iterative method
EBSD; electron backscatter diffraction
GND; geometrically necessary dislocations
KAM; Kernel average misorientation
LRT; likelihood ratio test
MLE; maximum likelihood estimation
PAVA; pool-adjacentviolators algorithm

1. Introduction

Understanding the intrinsic nature of the mechanical properties of metals is usually not an easy task. In order to get insight into what gives desired mechanical performance to a

CONTACT M. Vittorietti m.vittorietti@tudelft.nl Department of Applied Mathematics, Delft University of Technology, Van Mourik Broekmanweg, 6, Delft, 2628 XE, Netherlands

metal, a deep and detailed analysis of the metal microstructure characteristics is needed. For instance, it is known in literature that dislocations, i.e. line defects in the crystalline arrangement of the atoms [17], play a fundamental role in the mechanical behavior of metal alloys. More specifically, the appearance of geometrically necessary dislocations[1] (GNDs) during plastic deformation of the material contributes to the hardening of the material. Detecting GNDs from 2D microstructure images is often challenging. One widely accepted way is to use the so-called Kernel average misorientation (KAM) [26]. The KAM, measured in electron backscatter diffraction (EBSD), quantifies the average misorientation around a measurement point with respect to a defined set of a nearest or nearest plus second-nearest neighbor points [8].

In [6,22,31], studies on the relation between GNDs and microstructure properties such as grain size and carbides size are presented. The relation between GNDs and grain size has both theoretical and experimental confirmation and it can be related to the well-known macroscopic physical Hall-Petch relation [15,29]. In fact, the Hall-Petch relation, in its original version, describes the negative dependence of yield stress (mechanical property) on grain size; loosely speaking the smaller grains are, the stronger the material is. More specifically in [18], the authors give as an explanation of the relation between GNDs and grain size that as the grain size decreases the grain boundary layer in which GNDs typically accumulate, occupies a greater volume fraction of the material, therefore it is reasonable to think that the smaller are the grains, the more GNDs will be observed.

Still unclear is, instead, the relation between carbides and GNDs. In fact, since the 1940s, several studies on how carbides affect the mechanical behavior of metals have been conducted. In [30], the authors state that the primary carbides and their distribution have a major influence on the wear resistance and the toughness of the material. However, carbides tend to precipitate along the grain boundaries, that as said before, are the locations in which GNDs typically accumulate. Until now, no direct physical relationship has been found between carbides and GNDs. Therefore, isolating carbides effect and assessing the conjecture on the positive relation between carbides and GNDs is a problem of interest.

In [16], a descriptive statistical analysis with response variable KAM, used as a proxy of GNDs and as explanatory variables the number of grains, the number of carbides and the position of carbides revealed an almost monotone trend of the response variable according to the increments of the explanatories.

Therefore, in order to take into account the already known direction of the physical relation, we want to propose an approach that incorporates this information and a procedure for testing the prementioned conjectures on a new dataset.

In this context, isotonic regression comes to aid. In fact, the idea at the basis of isotonic regression is taking order restrictions into account for improving the efficiency of the statistical analysis by reducing the error or the expected error of estimates and increasing the power of the testing procedures, provided that the hypothesized order restriction actually holds. The first papers about isotonic regression appeared in the 1950's [1,41] and books [2,33] are well-known references for frequentist statistical inference under order restrictions. Isotonic regression proves its power in different fields such as epidemiology in testing the effects of different treatments or in dose-finding [34,40] but also in genetics [23], business [19], biology [3]. There are not many examples of isotonic regression use in Materials Science. Throughout this paper, special attention is given to the peculiar data structure. Nowadays, developments toward multivariate isotonic regression, isotonic regression in

inverse and censoring problems [13,14] are ongoing. Also in the Bayesian framework, the problem of estimating and testing under order constraints has been addressed and well developed [7,10,20,21]. But also in the most basic frequentist framework, there is still something missing.

In this paper, starting off with the most basic case, univariate isotonic regression of means under normality assumptions with known variances, we guide the reader into frequentist estimation and testing order restriction assumptions, considering different conditions on the variances.

Three different scenarios are considered. In all three cases, we focus on maximum likelihood as estimation procedure and likelihood ratio test (LRT) as test statistic for hypothesis testing.

The first case is the basic case in which 'the variances' are known or unknown but their ratio is known. This instance is considered extensively in [2,33] and results for estimation and testing order restrictions are already known.

The second scenario is from an applications point of view the most common scenario in which the variances are unknown. In [37], the authors derive a two steps estimating procedure for means and variances and interesting results on existence and uniqueness of the maximum likelihood estimates are derived under special conditions. Another iterative method, proposed in [38], is extended to the unknown variances case. The derivation of the test statistic and of its distribution in this scenario is not trivial. In fact, the estimate of the mean under the null hypothesis is also affected by the non-knowledge of the variances. We propose the LRT statistic and two different bootstrap approaches, one parametric and one non-parametric, for obtaining the test statistic distribution.

The last model considers not only the means under order restrictions but also the variances. This case has not often been faced probably because it is not common to have prior knowledge on the order of both means and variances. As in the unknown variances scenario, a two-step procedure for estimating means and variances is derived in [36] and similar results on existence and uniqueness under specific conditions on the empirical variances are given. In [38], an improved algorithm called alternating iterative method (AIM) and more general results about convergence are derived. For testing in this case we derive the LRT taking into account the order of variances also under the null hypothesis and apply a parametric and non-parametric bootstrap approach in line with the one derived in the unknown variance case to obtain approximate p-values.

The paper structure is as follows. In Section 2, we explain the estimation procedure of the isotonic means in the three different cases (Sections 2.1, 2.2, 2.3). In Section 3, the focus is on the LRT. We present it in the three different cases (Sections 3.1, 3.2, 3.3) and in Section 4, we propose both a parametric and non-parametric bootstrap approach for approximating the distribution of the test statistics under the null hypothesis. Finally, in Section 5, we come back to the application and we illustrate step-by-step how to deal with a real problem and more precisely how to perform isotonic regression and test for monotonicity of KAM with respect to the number of carbides. The paper ends with conclusions in Section 6.

2. Estimating restricted means in the normal case

We first introduce isotonic regression and the notation used in the rest of the paper in a more general context. Normality is assumed throughout this section.

Let y_{ij}, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$ be the jth observation of the response variable Y corresponding to the ith level of the explanatory variable X.

We assume Y_{ij} to be independent random variables, normally distributed with means μ_i and variances σ_i^2, $i = 1, \ldots, k, j = 1, 2, \ldots, n_i$.

The log-likelihood is then given by

$$l(\boldsymbol{\mu}, \boldsymbol{\sigma^2}) = \sum_{i=1}^{k} \left\{ -\frac{n_i}{2} \ln \sigma_i^2 - \frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} (y_{ij} - \mu_i)^2 \right\} + \mathrm{c} \tag{1}$$

where c is a constant which does not depend on the parameters $\boldsymbol{\mu} = (\mu_1, \ldots \mu_k)'$ and $\boldsymbol{\sigma^2} = (\sigma_1^2, \ldots, \sigma_k^2)'$.

Furthermore, we assume that $\boldsymbol{\mu}$ satisfies

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k. \tag{2}$$

A k-dimensional vector $\boldsymbol{\mu}$ is said to be isotonic if $t \leq s$ implies $\mu_t \leq \mu_s$.

Let D be the set of all the isotonic vectors in $\mathbb{R}^k$,

$$D = \{\boldsymbol{\mu} \in \mathbb{R}^k; \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k\} \tag{3}$$

In this section, we are interested in the maximum likelihood estimator of $(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$, where $\boldsymbol{\mu}$ is isotonic and $\sigma_i^2 > 0$. Depending on the information on $\boldsymbol{\sigma^2}$, different MLEs have been derived.

In the following three sections, the three different cases are considered.

2.1. Isotonic regression of means with known variance ratio

This first case constitutes the most basic case in which all variances are either known or unknown but they differ according to some known multiplicative constants c_i. This means that the variance σ_i^2 of the response variable Y_i is given by:

$$\sigma_i^2 = c_i \sigma^2, \quad 1 \leq i \leq k.$$

This specific case is already covered in [2,33], but we hereafter report the main results. The problem of maximizing log-likelihood (1) in μ can be rewritten equivalently as solving:

$$\min_{\mu \in D} \sum_{i=1}^{k} (\bar{y}_i - \mu_i)^2 w_i \tag{4}$$

where $\bar{y}_i = \frac{\sum_j y_{ij}}{n_i}$ and $w_i = \frac{n_i}{c_i}$. Note that this objective function does not depend on σ^2. The solution, $\hat{\boldsymbol{\mu}}^I$, is called the isotonic regression of $\bar{\boldsymbol{y}} = (\bar{y}_1, \ldots, \bar{y}_k)$ with weights $\boldsymbol{w} = (w_1, \ldots, w_k)$ [37]. For obtaining the solution to (4), different algorithms have been proposed in the literature ([2,33]). In this paper, the the intuitive and skilfully implemented 'Pool-Adjacent Violators Algorithm' (PAVA) is used ([1,5]).

More details about the algorithm are provided in Appendix 1.

2.2. Isotonic regression of means with unknown variances

In this second case, no assumptions on the variances are made. They are unknown and for obtaining the maximum likelihood estimate of $\boldsymbol{\mu}$, they need to be estimated as well. In [37], the authors consider this case and interesting results on existence and uniqueness of the MLE are achieved. We hereby recall the main results. The approach is to maximize the log-likelihood (1), with $\boldsymbol{\mu} \in D$ and $\boldsymbol{\sigma^2} \in \mathbb{R}_+^k$.

For any fixed $\boldsymbol{\sigma^2} \in \mathbb{R}_+^k$, the maximizer $\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}$ of $l(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$ over $\boldsymbol{\mu} \in D$ is the isotonic regression of $\bar{\boldsymbol{y}}$ with weights $\boldsymbol{w} = (w_1, \dots w_k)'$ and $w_i = \frac{n_i}{\sigma_i^2}$.

On the other hand, for any fixed $\boldsymbol{\mu} \in D$, the maximizer $\boldsymbol{\sigma^2}$ of $l(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$ over $\boldsymbol{\sigma} \in \mathbb{R}_+^k$ is $\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}(\boldsymbol{\mu}) = (\hat{\sigma}_1^2(\mu_1), \dots \hat{\sigma}_k^2(\mu_k))'$, where $\hat{\sigma}_i^2(\mu_i) = \frac{\sum_j^{n_i}(y_{ij}-\mu_i)^2}{n_i}$.

Substituting $\hat{\sigma}^2(\mu)$ into (1), we can express the profile log-likelihood of $\boldsymbol{\mu}$ as

$$l(\boldsymbol{\mu}) = \sum_{i=1}^{k} -n_i \ln[\bar{\sigma}_i^2 + (\bar{y}_i - \mu_i)^2] + c \tag{5}$$

where $\bar{\sigma}_i^2 = \frac{\sum_j^{n_i}(y_{ij}-\bar{y}_i)^2}{n_i}$ is the sample variance of the ith normal population and c a constant that does not depend on $\boldsymbol{\mu}$. Note that $l(\boldsymbol{\mu}) \to -\infty$ if $\mu_k \to \infty$ or $\mu_1 \to -\infty$. Hence, maximizing l over D is equivalent to maximizing l over a compact subset of D of type $D_a = \{\mu \in D : \mu_1 \geq -a, \mu_k \leq a\}$. As l is continuous on D_a, a maximizer over D exists.

As previously said, the authors in [37] discuss also uniqueness of the MLE of $(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$. They state that l is not a concave function in general and that for guaranteeing uniqueness the following condition suffices (see Theorem 2.3 [37]):

Condition 2.1: For $i = 1, \dots, k$, $\bar{\sigma}_i^2 > \max\{(\bar{y}_i - \min(\bar{\boldsymbol{y}}))^2, (\bar{y}_i - \max(\bar{\boldsymbol{y}}))^2\}$.

For finding a maximizer of (5), a two-step iterative algorithm based on PAVA has been proposed in [37]. From an initial guess for $\boldsymbol{\mu}$, the associated maximizer in $\boldsymbol{\sigma^2}$ is computed and after that the maximizer in $\boldsymbol{\mu}$ based on this $\boldsymbol{\sigma^2}$ and so on. This iterative procedure stops when the maximum difference between the estimated means at step $l-1$ and at step l is less than an arbitrary small threshold value, e.g.

$$\max_{1 \leq i \leq k} |\mu_i^{I(l-1)} - \mu_i^{I(l)}| \leq 10^{-m},$$

where m is taken to be equal to 3 in our case. In [38], the authors propose a new algorithm called AIM. The procedure is based on the minimization of a semi-convex function. In particular, restating the problem in terms of $(\boldsymbol{\mu}, \boldsymbol{\nu})$, where $\boldsymbol{\nu} = (1/\sigma_1^2, \dots, 1/\sigma_k^2)'$ and given D_a is a convex subset of $\mathbb{R}^k$ and V a convex subset of $\mathbb{R}_+^k$, $V = \{\nu \in \mathbb{R}_+^k : 0 \leq 1/\max_i(\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta)) \leq \nu_i \leq 1/\min_i(\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta))\}$, $L(\boldsymbol{\mu}, \boldsymbol{\nu})$ is a semi-convex function because: i) $L(\boldsymbol{\mu}, \boldsymbol{\nu})$ is defined on $D_a \times V$; ii) for any given $\boldsymbol{\mu} \in D_a$, $L(\boldsymbol{\mu}, \cdot)$ is strictly convex on V and, for any given $\boldsymbol{\nu} \in V$, $L(\cdot, \boldsymbol{\nu})$ is strictly convex on D_a. The algorithm originally proposed for the simultaneous order restrictions of means and variances can be easily extended to the unknown variance case. The iteration method works in alternating the search of the minimum point, $\boldsymbol{\mu}^{(l)}$, of $L(\boldsymbol{\mu}, \boldsymbol{\nu}(\boldsymbol{\mu}^{(l-1)}))$ on a compact subset

D_a and the search of the minimum point, $\boldsymbol{\nu}^{(l)}$, of $L(\boldsymbol{\mu}(\boldsymbol{\nu}^{(l-1)}), \boldsymbol{\nu})$ on V. Proof of the convergence of the algorithm does not require additional conditions [38]. The iterative procedure stops when the difference between the likelihoods at step $l-1$ and at step l is less than an arbitrary small threshold value:

$$|L(\boldsymbol{\mu}^{(\boldsymbol{l-1})}, \boldsymbol{\nu}^{(\boldsymbol{l-1})}) - L(; \boldsymbol{\mu}^{(\boldsymbol{l})}, \boldsymbol{\nu}^{(\boldsymbol{l})})| \leq 10^{-m} \tag{6}$$

A more detailed version of both algorithms is reported in Appendix A.2.

2.3. Isotonic regression of means and variances simultaneously

We now assume that both mean and variances are restricted by simple orderings. Therefore, in addition to assumption (2), we assume also:

$$\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_k^2 > 0 \tag{7}$$

The reason for taking decreasing order relates to our application considered in Section 5; increasing variances can be dealt with analogously. In [36], maximum likelihood estimation under simultaneous order restrictions on mean and variances from a normal population is studied. Some of the most important results are hereby recalled. The approach is to maximize the log-likelihood (1) with $\boldsymbol{\mu} \in D$ and $\boldsymbol{\sigma^2} \in \bar{G}$, where $\bar{G}$ is the closure of

$$G = \{\boldsymbol{\sigma^2} \in \mathbb{R}_+^k : \sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_k^2 > 0\}. \tag{8}$$

This means that the maximizer will have positive σ^2-values if there is variation within the groups. Then, for any fixed $\boldsymbol{\sigma^2} \in G$, the maximizer $\boldsymbol{\mu^I}$ of $l(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$ over $\boldsymbol{\mu} \in D$ is the isotonic regression of $\bar{\boldsymbol{y}}$ with weights $\boldsymbol{w} = (w_1, \ldots w_k)'$ and $w_i = \frac{n_i}{\sigma_i^2}$.

Furthermore, for any $\boldsymbol{\mu} \in D$, the maximizer $\hat{\boldsymbol{\sigma}}^{\boldsymbol{2I}}(\boldsymbol{\mu})$ of $l(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$ is the so-called antitonic regression (isotonic regression with reversed order [13]) of $\boldsymbol{s^2} = (s_1^2, \ldots s_k^2)'$, $s_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2}{n_i}$, with weights $\boldsymbol{N} = (n_1, \ldots, n_k)'$. Existence is guaranteed noticing that $\boldsymbol{\sigma^2} \in [\min_i(\min_{\min_i(\bar{\boldsymbol{y}}) \leq \theta \leq \max_i(\bar{\boldsymbol{y}})} s_i^2(\theta)), \max_i(\min_{\min_i(\bar{\boldsymbol{y}}) \leq \theta \leq \max_i(\bar{\boldsymbol{y}})} s_i^2(\theta))]$, $s_i^2(\theta) = \sum_{j=1}^{n_i}(y_{ij} - \theta)^2/n_i$ (see [36, Theorem 2.1]).

Uniqueness is proved under the following condition (see Theorem 2.2 [36])

Condition 2.2: For $i = 1, \ldots k$ the sample variance $\bar{\sigma}_i^2$ satisfies $\bar{\sigma}_i^2 > 2(b-a)$, where b and a are the maximal and the minimal means respectively.

As in the unknown variances case, a two-step iterative algorithm is proposed for finding the solution for both means and variances under order restrictions. The proof of the convergence of the algorithm is given under Condition 2.2.

Later, in [38], as mentioned in the previous section, the authors show that restating the problem in terms of $(\boldsymbol{\mu}, \boldsymbol{\nu})$, where $\boldsymbol{\nu} = (1/\sigma_1^2, \ldots, 1/\sigma_k^2)'$ Condition 2.2 is not needed for proving that the algorithm converges. In fact, also in this case the proposed AIM algorithm can be employed. Since $L(\boldsymbol{\mu}, \boldsymbol{\nu})$ has continuous second-order partial derivatives and the Hessian matrix with respect to μ $H(\boldsymbol{\mu}, \boldsymbol{\nu}) = \text{diag}(n_1\nu_1, \ldots, n_k\nu_k)$ is a positive definite diagonal matrix for any fixed $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_k)' \in V_0$, $V_0 = \{\nu \in \mathbb{R}^k :$

$0 \leq 1/\max_i(\min_{\min_i(\bar{y})\leq\theta\leq\max_i(\bar{y})} s_i^2(\theta)) \leq \nu_1 \leq \cdots \leq \nu_k \leq 1/\min_i(\min_{\min_i(\bar{y})\leq\theta\leq\max_i(\bar{y})} s_i^2(\theta))\}$ then by Theorem 4 in [38] the iterative sequence of solutions to $L(\boldsymbol{\mu}, \boldsymbol{\nu})$, $\{(\boldsymbol{\mu^{(n)}}, \boldsymbol{\nu^{(n)}})\}$ converges to the MLE solution and consequently the sequence $\{(\boldsymbol{\mu^{(n)}}, \boldsymbol{\sigma^{2(n)}})\}$ as well.

As in the previous case, the alternating iterative procedure is stopped when the maximum difference between the likelihoods at step $l-1$ and at step l is less than an arbitrary small threshold value (see (6)).

A pseudo code of the algorithms can be found in Appendix A.3.

3. LRT: constant μ against monotonicity

We are interested in testing hypotheses of monotonicity in μ under the various assumptions on the variances discussed in Section 2. There exists extensive literature on testing hypotheses on means. In most cases, a standard testing procedure entails testing the hypothesis of equality of means against the hypothesis that they are different. In this paper, we consider the same null hypothesis but the alternative is different: monotonicity of the means. As in the previous section, we consider three different testing frameworks according to the different assumptions on the variances. In all three different scenario, the test statistic of interest is the LRT, an intuitive and powerful tool in hypothesis testing. In both [2] and [33], an entire chapter is dedicated to LRT developments and its use for testing order restrictions hypothesis under the normality assumption and known variance ratio. Using the same notation used in Section 2, we wish to test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

against monotonicity of means

$$H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k. \tag{9}$$

The LRT for H_0 against H_1 can be defined as:

$$\Lambda = \frac{\max_{(\boldsymbol{\mu}\in H_0;\boldsymbol{\sigma}^2)} L(\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_k}; \boldsymbol{\mu}, \boldsymbol{\sigma^2})}{\max_{(\boldsymbol{\mu}\in H_1;\boldsymbol{\sigma}^2)} L(\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_k}; \boldsymbol{\mu}, \boldsymbol{\sigma^2})} \tag{10}$$

where $\boldsymbol{y_i} = (y_{i1}, \ldots, y_{in_i})'$, $\boldsymbol{\mu} = (\mu_1, \ldots \mu_k)'$ and $\boldsymbol{\sigma^2} = (\sigma_1^2, \ldots \sigma_k^2)'$. It rejects the null hypothesis for small values of Λ or alternatively for large values of $-2\log\Lambda$. The convenience in using this other form lies on the analogy with the χ^2 statistic used to test against the alternative hypothesis $\bar{H}_0$, that not all μ_i's, $i = 1, \ldots, k$, are the same.

In the following sections, more explicit expressions for Λ are given depending on the specific assumptions on means and variances.

3.1. LRT with known variance ratio

As in Section 2.1, let y_{ij} $j = 1, 2, \ldots n_i$, $i = 1, 2, \ldots k$ be independent observations, normally distributed with unknown mean μ_i and variances $\sigma_i^2 = c_i\sigma^2$ with c_i known and σ^2 unknown. Under H_0, the maximum likelihood estimate of $\mu_1 = \mu_2 = \cdots = \mu_k$ is given

by:

$$\hat{\mu}_{H_0} = \frac{\sum_{i=1}^{k} w_i \bar{y}_i}{\sum_{i=1}^{k} w_i} \tag{11}$$

with $w_i = \frac{n_i}{c_i}$. Under H_1 the MLE of $\boldsymbol{\mu}$ is $\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}_{\boldsymbol{H_1}}$, the isotonic regression of $\bar{\boldsymbol{y}}$, with weights $\boldsymbol{w} = (w_1, \ldots, w_k)'$, with respect to the simple order defined in (9).

The LRT for H_0 against H_1, if the variances are known and $c_i = 1$ boils down to rejecting H_0 for large values of

$$-2\log\Lambda = \frac{1}{\sigma^2}\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}_{H_0})^2 - \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}^I_{iH_1})^2\right] \tag{12}$$

It is easy to check that the test is equivalent to rejecting H_0 for large values of:

$$\bar{\chi}^2 = \frac{\sum_{i=1}^{k} \bar{\chi}_i^2}{\sigma^2} \tag{13}$$

where $\bar{\chi}_i^2 = n_i(\hat{\mu}^I_{iH_1} - \hat{\mu}_{H_0})^2$ and σ^2 is the (known) common value of the variance.

Now, let us consider the more general case, $\sigma_i^2 = c_i\sigma^2$ with $c_1, c_2, \ldots c_k$ known and σ^2 unknown. The estimator of σ^2 under the null hypothesis is

$$\hat{\sigma}^2_{H_0} = \frac{\sum_{i=1}^{k} c_i^{-1} \sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}_{H_0})^2}{N} \tag{14}$$

and under H_1

$$\hat{\sigma}^2_{H_1} = \frac{\sum_{i=1}^{k} c_i^{-1} \sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}^I_{iH_1})^2}{N} \tag{15}$$

The LRT rejects H_0 for small values of $\Lambda = (\frac{\hat{\sigma}^2_{H_1}}{\hat{\sigma}^2_{H_0}})^{N/2}$ or equivalently, taking $\bar{E}^2 = 1 - \Lambda^{2/N}$, for large values of

$$\bar{E}^2 = \frac{\sum_{i=1}^{k} c_i^{-1} \bar{\chi}_i^2}{\sum_{i=1}^{k} c_i^{-1} \sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}_{H_0})^2} \tag{16}$$

An extension to the multivariate case with covariance matrix Σ unknown but common can be found in [28,35].

3.2. LRT with unknown variances

In this second case, no assumptions on the variances are made. They are unknown and possibly unequal. Using the notation of Section 2.2, let y_{ij}, $j = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, k$ be independent observations from a univariate normal distribution with unknown mean vector μ_i and completely unknown variances $\sigma_i^2 > 0$. Let $\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}$ be the solution of the isotonic regression of $\bar{\boldsymbol{y}}$ with weights $\boldsymbol{w} = (w_1, \ldots, w_k)'$, $w_i = \frac{n_i}{\sigma_i^2}$ found used Algorithm (2.2) in Appendix 1.

The first example of testing when all the variances are unknown can be found in [4] and the univariate version of the test proposed by the author is:

$$\sum_{i=1}^{k} \frac{(\hat{\mu}_i^I - \bar{y})^2 n_i}{s_i^2} \tag{17}$$

where $\bar{y} = \frac{\sum_{i=1}^{k} n_i \bar{y}_i}{\sum_{i=1}^{k} n_i}$ and $s_i^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2}{n_i-1}$. This test is clearly inspired by the LRT but it is not.

Let us consider first the maximum likelihood solution $(\hat{\mu}_{H_0}, \hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}{}_{H_0})$, $\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}_{\boldsymbol{H_0}} = (\hat{\sigma}^2_{1H_0}, \ldots, \hat{\sigma}^2_{kH_0})'$ under the null hypothesis. The log-likelihood under the null hypothesis is

$$l(\mu, \boldsymbol{\sigma}^{\boldsymbol{2}}) = \sum_{i=1}^{k} \left\{ -\frac{n_i}{2} \ln \sigma_i^2 - \frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 \right\} + \text{c.} \tag{18}$$

Differentiating this log-likelihood with respect to μ and σ_i^2, the following $k+1$ score equations in $k+1$ unknowns emerge:

$$\begin{cases} \mu_{H_0} = \dfrac{\sum_{i=1}^{k} n_i \sigma_{iH_0}^{-2} \bar{y}_i}{\sum_{i=1}^{k} n_i \sigma_{iH_0}^{-2}} \\ \sigma^2_{iH_0} = \sum_{j=1}^{n_i} n_i^{-1} (y_{ij} - \mu_{H_0})^2 \quad 1 \le i \le k \end{cases} \tag{19}$$

Substituting $\boldsymbol{\sigma}^{\boldsymbol{2}}_{\boldsymbol{H_0}}(\boldsymbol{\mu})$ in (18), the profile likelihood of μ is:

$$l(\mu) = -\sum_{i=1}^{k} \frac{n_i}{2} \ln \left(\sum_{j=1}^{n_i} n_i^{-1} (y_{ij} - \mu)^2 \right) + c. \tag{20}$$

Theorem 1: *A maximizer of* (20) *over* $\mathbb{R}^d$ *exists and it is contained in* $[\min_i \bar{y}_i, \max_i \bar{y}_i]$. *Moreover, if* $[\min_i \bar{y}_i, \max_i \bar{y}_i] \in [\max_{1\le i\le k}(\bar{y}_i - \bar{\sigma}_i), \min_{1\le i\le k}(\bar{y}_i + \bar{\sigma}_i)]$ *then the maximizer is unique.*

Proof: Maximizing profile likelihood of μ (20) boils down to maximize the sum of functions

$$-\frac{n_i}{2} \ln(n_i(\bar{\sigma}_i^2 + (\bar{y}_i - \mu)^2)), \quad i = 1, \ldots, k. \tag{21}$$

Functions of type (21) are unimodal with mode at $\bar{y}_i$ and strictly concave on $[\bar{y}_i - \bar{\sigma}_i; \bar{y}_i + \bar{\sigma}_i]$. As the sum of unimodal functions is decreasing to the right of the rightmost mode (since all terms are decreasing) and from $-\infty$ to the leftmost mode, the sum is increasing (as all of the functions are increasing on that set). Therefore, any maximizer of l, if it exists, belongs to the interval $[\min_i \bar{y}_i, \max_i \bar{y}_i]$. As l is continuous on $[\min_i \bar{y}_i, \max_i \bar{y}_i]$, existence of a maximizer is guaranteed.

Then if we consider the (possibly empty) interval where all the functions in (21) are strictly concave, on that interval the sum is also strictly concave. As for each i the function (21) is strictly concave on $I_i = [\bar{y}_i - \bar{\sigma}_i; \bar{y}_i - \bar{\sigma}_i]$, (20) is strictly concave on $\bigcap_{i=1}^{k} I_i$. If

$[\min_i \bar{y}_i, \max_i \bar{y}_i]$ is contained in this intersection, l is strictly concave on $[\min_i \bar{y}_i, \max_i \bar{y}_i]$. Hence l has a unique maximizer on $\mathbb{R}^d$. ■

Remark 3.1: Remark in a setting with real data, it is easy to check whether $[\min_i \bar{y}_i, \max_i \bar{y}_i] \in [\max_{1\leq i\leq k}(\bar{y}_i - \bar{\sigma}_i), \min_{1\leq i\leq k}(\bar{y}_i + \bar{\sigma}_i)]$ and hence to determine whether the maximum is unique.

However, as seen from (19), the MLE estimate $(\mu, \boldsymbol{\sigma^2})$ has no closed form expression. Therefore, in [11] and [27], two different methods for finding the optimal solution are proposed. The first is an iterative procedure based on the Newton-Raphson method. A reasonable initial value for $\hat{\mu}_{H_0}^{(0)}$ is the so-called Graybill-Deal estimator [12] $\hat{\mu}_{(GD)} = \frac{\sum_{i=1}^{k}(n_i\bar{y}_i)/\bar{s}_i^2}{\sum_{i=1}^{k} n_i/\bar{s}_i^2}$ with $\bar{s}_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2}{n_i-1}$. The convergence speed of the algorithm strongly depends on the initial values. The second method is based on the profile likelihood approach. The authors in [27] propose the bisection method for finding the zero of the profile likelihood with respect to μ_{H_0}. Under H_1 we use as estimates of $(\boldsymbol{\mu_{H_1}}, \boldsymbol{\sigma^2_{iH1}})$, $(\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}, \hat{\boldsymbol{\sigma}}^{\boldsymbol{2}})$ found using the iterative procedure described in Section 2.2.

The LRT when the variances are completely unknown can be expressed as:

$$\tilde{\Lambda} = \prod_{i=1}^{k} \left(\frac{\hat{\sigma}_{iH_0}^2}{\hat{\sigma}_{iH_1}^2} \right)^{-\frac{n_i}{2}}$$

Therefore, as in the previous case, the test rejects for small values of $\tilde{\Lambda}$ or equivalently for large values of $-2\log\tilde{\Lambda}$.

3.3. LRT with ordered variances

Using the notation of Section 2.3, let y_{ij}, $j = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, k$ be independent observations from Normal distributions with mean vector μ_i and variances σ_i^2. As in the previous case, the first step is the estimation of $(\mu, \boldsymbol{\sigma^2})$ under the null hypothesis. In this case, we need to maximize (18) under the restriction

$$\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_k^2 > 0. \tag{22}$$

Theorem 2: *Suppose that for* $1 \leq i \leq k$, $\bar{\sigma}_i^2 > 0$. *Then there exists a maximizer of* (18) *under constraints* (22).

Proof: First consider the situation for fixed $\boldsymbol{\sigma^2}$ with $\sigma_i^2 > 0$ for all i. Differentiating (18) with respect to μ yields the equation

$$\sum_{i=1}^{k} \frac{n_i(\bar{y}_i - \mu)}{\sigma_i^2}$$

This shows, that for this $\boldsymbol{\sigma^2}$, the (unique) maximizer of (18) in μ is given by the following weighted sum of level-means,

$$\hat{\mu}(\boldsymbol{\sigma^2}) = \frac{\sum_{i=1}^{k} n_i\sigma_i^{-2}\bar{y}_i}{\sum_{i=1}^{k} n_i\sigma_i^{-2}}$$

Consequently, $\min_i \bar{y}_i \leq \hat{\mu}(\boldsymbol{\sigma}^{\boldsymbol{2}}) \leq \max_i \bar{y}_i$, bounding the set of possible maximizers of (18) in μ irrespective of the precise value of $\boldsymbol{\sigma}^{\boldsymbol{2}}$.

Now, given any $\mu \in \mathbb{R}$, the corresponding optimal $\boldsymbol{\sigma}^{\boldsymbol{2}}$ is the solution to the antitonic regression problem antireg($\bar{\boldsymbol{\sigma}}^{\boldsymbol{2}}_{\boldsymbol{H_0}}, \boldsymbol{N}$) where $\bar{\boldsymbol{\sigma}}^{\boldsymbol{2}} = (\bar{\sigma}_1^2, \ldots \bar{\sigma}_k^2)'$, $\bar{\sigma}_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij}-\mu)^2}{n_i}$, $\boldsymbol{N} = (n_1, \ldots, n_k)'$ (see [33, Example 1.5.5]). The vector to be projected has elements $\bar{\sigma}_i^2 + (\mu - \bar{y}_i)^2$. This means, that if μ is restricted to $[\min_i \bar{y}_i, \max_i \bar{y}_i]$, the coordinates to be projected all belong to the interval $[\min_i \bar{\sigma}_i^2, \max \bar{\sigma}_i^2 + (\max \bar{y}_i - \min \bar{y}_i)^2]$. So, if μ ranges over $[\min_i \bar{y}_i, \max_i \bar{y}_i]$, the optimal $\boldsymbol{\sigma}^{\boldsymbol{2}}$ is also contained in a the closed bounded region $[\min_i \bar{\sigma}_i^2, \max \bar{\sigma}_i^2 + (\max \bar{y}_i - \min \bar{y}_i)^2]^k$. By our assumption that all $\bar{\sigma}_i^2 > 0$, the MLE exists being a maximizer of a continuous function on a compact set in $\mathbb{R} \times \mathbb{R}^k$ ■

If we consider this case as a special case of the case considered in [36], the solution is unique if Condition 2.2 holds. Given that the solution is not in a closed form, we use an iterative procedure to approximate the solution. As a starting value $\hat{\mu}^{(0)}$, a modified version of the Graybill-Deal estimator of the common mean when the variances are subject to order restrictions proposed in [25] appears to be a good choice:

$$\hat{\mu}_{(I)} = \frac{\sum_{i=1}^{k} w_i \hat{\tau}_i \bar{y}_i}{\sum_{i=1}^{k} w_i \hat{\tau}_i} \tag{23}$$

where $\hat{\tau}_i$ is the isotonic regression of $(\boldsymbol{t}, \boldsymbol{N})$ where $\boldsymbol{t} = (t_1, \ldots t_k)'$, $t_i = \frac{1}{s_i^2}$.

Under H_1 we use as estimates of $(\boldsymbol{\mu}_{\boldsymbol{H_1}}, \boldsymbol{\sigma}^{\boldsymbol{2}}_{\boldsymbol{iH1}})$, $(\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}, \hat{\boldsymbol{\sigma}}^{\boldsymbol{2I}})$ found using the iterative procedure described in Section 2.3.

In contrast with the previous cases, it is not possible to further reduce the expression of the LRT because

$$\exp\left\{\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{(y_{ij}-\hat{\mu}_{H_0})^2}{\sigma^2_{H_0}}\right\}$$

does not reduce to a constant. The same holds under H_1. Therefore, the LRT in this case can be computed by substituting the solutions obtained via the iterative procedure under H_0 and H_1 in the generic expression given in (10):

$$\Lambda^I = \frac{L(\hat{\mu}_{H_0}, \hat{\boldsymbol{\sigma}}^{\boldsymbol{2I}}_{\boldsymbol{H_0}})}{L(\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}_{\boldsymbol{H_1}}, \hat{\boldsymbol{\sigma}}^{\boldsymbol{2I}}_{\boldsymbol{H_1}})} \tag{24}$$

4. Null hypothesis distribution of the test statistics: bootstrap approach

In order to determine the significance of the various test statistics proposed in the previous sections, we need the null hypothesis distribution of the test statistics. The main distributional results concerning $\bar{\chi}_k^2$ and $\bar{E}_k^2$, the test statistics derived in the known variance ratio case, are contained in [2, Theorems 3.1–3.2]. However, problems related to the value of k can arise in the analytical derivation of the p-values. Numerical approximation can be necessary, especially if $k > 4$ and if the variation in the range of the weights is not 'moderate' [32,39].

Furthermore, in the case of completely unknown variances, the null distribution depends on the unknown variances. When analytical derivation of the null distribution is particularly complex or not possible, bootstrap methodology is a good option. Therefore, we propose both a parametric and a non-parametric bootstrap approach that can be easily employed for finding approximate p-values taking into account the different assumptions on the variances. For overcoming the complex derivation when the variances are unknown, bootstrap procedures have been proposed in the literature [4,24].

In particular, in [24], an interesting review of the methods used to approximate the null distribution of the test statistic under H_0 and the restrictive normality assumption (with which we will not deal in this paper) is reported. Moreover, the authors propose both a parametric and non-parametric bootstrap approach for the LRT null distribution for one-sided hypothesis testing for means in a multivariate setting [24]. Also in [4], a bootstrap approach to test the homogeneity of order restricted mean vectors when the covariance matrices are unknown is used. In line with those previous approaches, here we propose two general bootstrap procedures, parametric and non-parametric, that can be used for testing the null hypothesis taking into account the various assumptions on the variances.

Parametric bootstrap

Algorithm 1

(1) Obtain the estimates $\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}_{\boldsymbol{iH_1}}$ and $\hat{\mu}_{H_0}$ using the original data and compute the observed value of the test statistic of interest $LRT^{(0)}$ ($\bar{\chi}^{2(0)}$, $\bar{E}^{2(0)}$, $\tilde{\Lambda}^{(0)}$ or $\Lambda^{I(0)}$).

(2) Generate, for $1 \leq i \leq k$, $1 \leq j \leq n_i$ $\boldsymbol{Y^*_{ij}} \sim N(\hat{\mu}_{H_0}, \sqrt{\sigma^2_{iH_0}})$, independently.

(3) For $(\boldsymbol{Y^*_i}, \ldots, \boldsymbol{Y^*_k})$ obtain the estimates $\hat{\boldsymbol{\mu}}^{\boldsymbol{I*}}_{\boldsymbol{i}}$ and $\hat{\mu}^*$ and compute the bootstrap test statistic of interest LRT^*

(4) Repeat **(2)**-**(3)** for a sufficient large number of times M

The bootstrap approximation of the p-value is the given by:

$$p \approx \frac{\#(LRT^* > LRT^{(0)})}{M} \tag{25}$$

and the null hypothesis is rejected whenever this p-value is less than the nominal level α. Step (2) is the key step, in which the assumption on the variances play a crucial role. It is interesting to notice that the above procedure can be further simplified. In fact, we can instead of generating individual observations, directly generate empirical means $\bar{y}_i = Z_i * \frac{\sigma}{\sqrt{n_i}}$, with Z_i Standard Normally distributed.

Non-parametric bootstrap

The non-parametric version of the bootstrap releases the normality assumption of the bootstrap samples. However, a relatively large sample size is required for the following approach.

Algorithm 2

(1) Obtain the estimates $\hat{\boldsymbol{\mu}}^{\boldsymbol{I}}_{\boldsymbol{iH_1}}$ and $\hat{\mu}_{H_0}$ using the original data and compute the observed value of the test statistic of interest $LRT^{(0)}$ ($\bar{\chi}^{2(0)}$, $\bar{E}^{2(0)}$, $\tilde{\Lambda}^{(0)}$ or $\Lambda^{I(0)}$).

(2) Standardize the original y_{ij} to obtain 'standardized residuals' $z_{ij} = \frac{y_{ij}-\bar{y}_i}{\sqrt{s_i^2}}$

(3) Combine all z_{ij} observations from $(1 \leq i \leq k; 1 \leq j \leq n_i)$ into a vector of length $\sum_{i=1}^{k} n_i$ and draw k simple random samples z^*_{ij} with replacement each of respective sizes $(n_1, n_2, \ldots, n_k)$

(4) Transform z^*_{ij} to $y^*_{ij} = z^*_{ij} \cdot \tilde{\sigma}_{iH_0} + \hat{\mu}_{H_0}$

(5) For each bootstrap sample y^*_{ij} $(1 \leq i \leq k; 1 \leq j \leq n_i)$ obtain the estimates $\hat{\boldsymbol{\mu}}^{\boldsymbol{I*}}_{\boldsymbol{iH_1}}$ and $\hat{\mu}^*_{H_0}$ and compute the bootstrap test statistic of interest LRT^*

(6) Repeat **(3)**-**(4)**-**(5)** for a sufficient large number of times M

The bootstrap approximated p-value is defined as in the parametric case (25).

5. Application

One of the most common ways for investigating strength and ductility of metallic materials is by performing a tensile test. Loosely speaking, a tensile test is an experiment in which force is applied to the test sample causing deformation of the material, temporarily (elastic behavior), permanently (plastic behavior) and eventually its fracture [9]. Data used in this paper are image data of the microstructure of the material subjected to a plastic strain (deformation) of 0.139 obtained performing a uniaxial tensile test in which force is applied to the test sample with respect to just one specific axis (Figure 1).

At a microstructure level, the deformation of the material corresponds to displacements in the lattice structure and in the possible appearance of GNDs. The material used in this paper is an annealed AISI420 stainless steel with $M_{23}C_6$ carbides and aim is investigating the carbide effect on the GNDs formation. KAM is used as a proxy of the GNDs. In Figure 2, the KAM is represented by filaments inside and among grains delimited by solid lines called grain boundaries; carbides are the black dots.

Modeling the relationship between KAM and carbides and more generally understanding its inhomogeneous distribution over the microstructure is now the main aim and it can be considered a starting point for finding a stochastic model for predicting mechanical properties from 2D microstructure images. We apply estimation procedures and perform tests under order restrictions, solving the three different univariate isotonic regressions problems (2.1, 2.2, 2.3) according to the assumption on the variances.

The first step for obtaining the data in the most suitable form for the analysis is 'overlaying' a grid over the image. In Figure 2, a 25×25 grid is added to the image. With y_{ij}, we denote the mean KAM value (expressed in misorientation degree $[0, 5]$) of the jth square of the grid of the image taken in which i carbides are observed. The explanatory variable X in all three isotonic regressions is the number of carbides observed in the grid squares. A plot of the data is shown in Figure 3.

We wish to test the null hypothesis that the expected KAM is the same in all the squares of the gird, regardless the numbers of carbides observed in the grid. The alternative

Figure 1. Tensile testing machine.

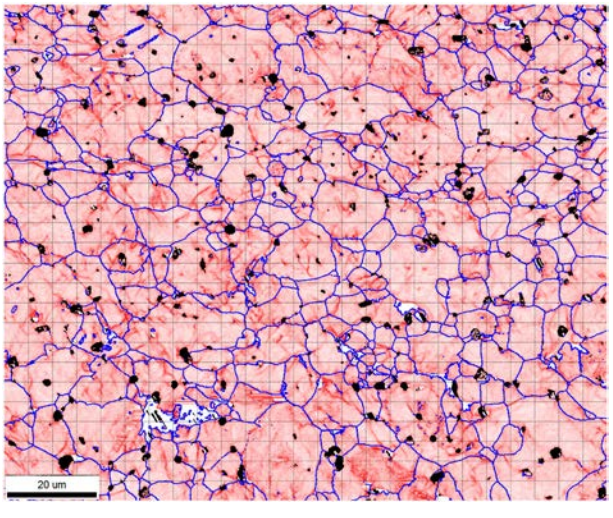

Figure 2. Microstructure image showing the KAM at strain level 13.9% (overlapped grid of 25 $\times$ 25).

hypothesis

$$\mu_0 \leq \mu_1 \leq \mu_2 \leq \mu_3. \tag{26}$$

represents the idea that KAM tends to be higher in areas where more carbides are observed. Moreover, in the ordered variances case, we assume that

$$\sigma_0^2 \geq \sigma_1^2 \geq \sigma_2^2 \geq \sigma_3^2 > 0. \tag{27}$$

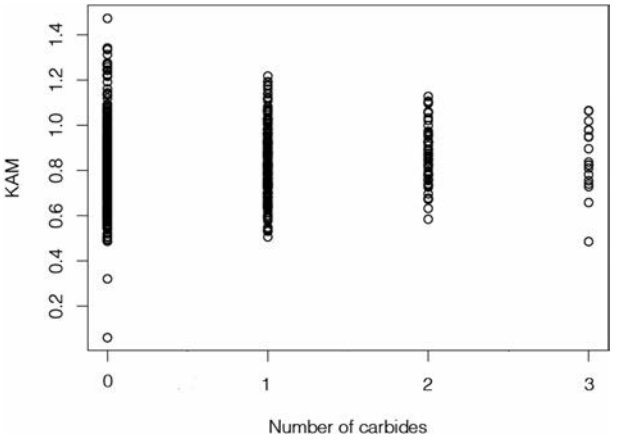

Figure 3. Plot of KAM and number of carbides for the 625 squared areas of Figure 2.

Table 1. Values of estimated means and variances of the KAM conditioned on the number of carbides visible in a square of a grid 25 × 25 according to different order restrictions assumptions (13.9% Strain).

	0	1	2	3
$\bar{y}$	0.815	0.833	0.870	0.854
$\bar{\sigma}^2$	0.035	0.024	0.017	0.022
s^2	0.035	0.024	0.017	0.023
$\hat{\mu}^I_{(1)}$	0.815	0.833	0.867	0.867
$\hat{\mu}^I_{(2)}$	0.815	0.833	0.867	0.867
$\hat{\sigma}_{(2)}$	0.035	0.024	0.017	0.022
$\hat{\mu}^I_{3)}$	0.815	0.833	0.866	0.866
$\hat{\sigma}^I_{(3)}$	0.035	0.024	0.018	0.018
n	340	211	54	18

This is in accordance with what we see in Figure 3. In fact, the idea behind this assumption is that in areas in which less carbides are observed GNDs have more freedom to move, resulting in increments in dispersion.

The results of the estimation in the three different scenarios faced in Sections 2.1–2.3 are summarized in Table 1. The values of $\bar{y}$, $\bar{\sigma}^2$ and s^2 represent the empirical group mean, group variance and group sample variance respectively. Comparing the solutions of the three different isotonic regression, (1) known variance ratio, (2) unknown variance and (3) variances under order restrictions, just very slightly differences can be noticed.

In Table 2, the results of testing the null hypothesis are shown. For computing $\chi^{2(0)}$ and $\bar{E}^{2(0)}$, the variance ratio is supposed to be known. In the specific case, we assume that the KAM total variance for the whole image is the real-known variance and that $c_i = \frac{\bar{\sigma}_i^2}{\sigma^2}$. For computing both the parametric and non-parametric p-values, the two different bootstrap approaches described in Section 4 have been used and M, the number of replications, is

Table 2. Estimated values for the four different LRT with the corresponding parametric and non-parametric *p*-values.

	$\hat{\mu}_{H_0}$	Test statistic	*p*-value (parametric)	*p*-value (non-parametric)
$\bar{\chi}^{2(0)}$	0.827	5.760	0.0323	0.0310
$\tilde{E}^{2(0)}$	0.831	0.0121	0.0112	0.0085
$-2\log\tilde{\Lambda}$	0.831	7.330	0.0178	0.0222
$\Lambda^{I(0)}$	0.831	7.105	0.0212	0.0251
	0	1	2	3
$\hat{\sigma}^2_{H_0}$	0.035	0.024	0.018	0.022
$\hat{\sigma}^{2I}_{H_0}$	0.035	0.024	0.019	0.019

taken equal to 20, 000. Independent of the knowledge or assumptions on the variances, the conclusion is the same and leads to the rejection of the null hypothesis.

6. Conclusions

This paper presents three different models involving order restrictions and within these models the ML estimators and LRTs for the homogeneity of the means against monotonicity are introduced and studied. Prior knowledge given by physical relations or intuition is not often exploited in statistical studies about materials and this can lead to less efficient methods that produce less accurate results. After having described the estimation procedures and highlighted how prior knowledge of the variances influence these, we propose the LRT as test statistic for testing homogeneity of means. In the unknown variances case and the ordered variances case, heteroskedasticity plays a crucial role also under the null hypothesis, leading to different estimates of the common mean under H_0. Results on existence and uniqueness of the maximum likelihood estimates in these last two cases are derived. Furthermore, two different bootstrap approaches are proposed for approximating the null distribution of the test statistic under the different assumptions on the variances. The proposed tests are applied to a real data example from Materials Science, showing evidence that the so-called KAM tends to be higher in regions of the microstructure where more carbides are observed. In fact, incorporating reasonable intuition about the order of means and the variances order in this context helps understanding the evolution of complicated structure of dislocations in metals and its effect on the hardening behavior of the material during deformation.

Note

1. Dislocations are usually classified into redundant and non-redundant dislocations, respectively, called Statistically Stored Dislocations (SSDs) and Geometrically Necessary Dislocations (GNDs). GNDs are dislocations with a cumulative effect and they allow the accommodation of lattice curvature due to non-homogeneous deformation. They control the work hardening individually by acting as obstacles to slip and collectively by creating a long-range back stress.

Acknowledgments

This research was carried out under project number S41.5.14547b in the framework of the Partnership Program of the Materials innovation institute M2i (www.m2i.nl) and the Technology

Foundation TTW (www.stw.nl), which is part of the Netherlands Organization for Scientific Research (www.nwo.nl).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was carried out under project number S41.5.14547b in the framework of the Partnership Program of the Materials innovation institute M2i (www.m2i.nl) and the Technology Foundation TTW (www.stw.nl), which is part of the Netherlands Organization for Scientific Research (www.nwo.nl).

ORCID

Martina Vittorietti http://orcid.org/0000-0002-7695-3209
Jilt Sietsma http://orcid.org/0000-0001-8733-4713

References

[1] M. Ayer, H.D. Brunk, G.M. Ewing, W.T. Reid, and E. Silverman, *An empirical distribution function for sampling with incomplete information*, Ann. Math. Stat. 26 (1955), pp. 641–647.
[2] R.E. Barlow and H.D. Brunk, *The isotonic regression problem and its dual*, J. Am. Stat. Assoc. 67 (1972), pp. 140–147.
[3] S. Barragán, M.A. Fernández, C. Rueda, and S.D. Peddada, *isocir: an R package for constrained inference using isotonic regression for circular data, with an application to cell biology*, J. Stat. Softw. 54 (2013), pp. 00–00.
[4] A. Bazyari, *Bootstrap approach to test the homogeneity of order restricted mean vectors when the covariance matrices are unknown*, Commun. Stat.-Simul. Comput. 46 (2017), pp. 7194–7209.
[5] M.J. Best and N. Chakravarti, *Active set algorithms for isotonic regression: A unifying framework*, Math. Program. 47 (1990), pp. 425–439.
[6] M.W. Bird, T. Rampton, D. Fullwood, P.F. Becher, and K.W. White, *Local dislocation creep accommodation of a zirconium diboride silicon carbide composite*, Acta. Mater. 84 (2015), pp. 359–367.
[7] F. Böing-Messing, M.A. van Assen, A.D. Hofman, H. Hoijtink, and J. Mulder, *Bayesian evaluation of constrained hypotheses on variances of multiple independent groups*, Psychol. Meth. 22 (2017), pp. 262.
[8] M. Calcagnotto, D. Ponge, E. Demir, and D. Raabe, *Orientation gradients and geometrically necessary dislocations in ultrafine grained dual-phase steels studied by 2D and 3D EBSD*, Mater. Sci. Eng.: A 527 (2010), pp. 2738–2746.
[9] J.R. Davis, *Tensile Testing*, Materials Park, ASM international, Ohio, 2004.
[10] M.R. Danaher, A. Roy, Z. Chen, S.L. Mumford, and E.F. Schisterman, *Minkowski-Weyl priors for models with parameter constraints: an analysis of the bioCycle study*, J. Am. Stat. Assoc. 107 (2012), pp. 1395–1409.
[11] E.Y. Gökpinar and F. Gökpinar, *A test based on the computational approach for equality of means under the unequal variance assumption*, Hacettepe J. Math. Stat. 41 (2012), pp. 605–613.
[12] F.A. Graybill and R.B. Deal, *Combining unbiased estimators*, Biometrics 15 (1959), pp. 543–550.
[13] P. Groeneboom and G. Jongbloed, *Nonparametric Estimation Under Shape Constraints: Estimators*, Algorithms and Asymptotics (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge University Press, Cambridge, 2014, Vol. 38.
[14] A. Guntuboyina and B. Sen, *Nonparametric shape-restricted regression*, Stat. Sci. 33 (2018), pp. 568–594.

[15] E.O. Hall, *The deformation and ageing of mild steel: III discussion of results*, Proc. Phys. Soc., B 64 (1951), pp. 747–753.
[16] J. Hidalgo, M. Vittorietti, H. Farahani, F. Vercruysse, R. Petrova, and J. Sietsma, *Influence of large M23C6 carbides on the heterogeneous strain development in annealed 420 stainless steel*, preprint (2019).
[17] D. Hull and D.J. Bacon, *Introduction to Dislocations*, 4th ed. Butterworth-Heinemann, Oxford, 2001.
[18] J. Kadkhodapour, S. Schmauder, D. Raabe, S. Ziaei-Rad, U. Weber, and M. Calcagnotto, *Experimental and numerical study on geometrically necessary dislocations and non-homogeneous mechanical properties of the ferrite phase in dual phase steels*, Acta. Mater. 59 (2011), pp. 4387–4394.
[19] A. Keshvari and T. Kuosmanen, *Stochastic non-convex envelopment of data: applying isotonic regression to frontier estimation*, Eur. J. Oper. Res. 231 (2013), pp. 481–491.
[20] I. Klugkist, O. Laudy, and H. Hoijtink, *Inequality constrained analysis of variance: a Bayesian approach*, Psychol. Meth. 10 (2005), pp. 477.
[21] W. Li and H. Fu, *Bayesian isotonic regression dose–response model*, J. Biopharm. Stat. 27 (2017), pp. 824–833.
[22] P.D. Littlewood, T.B. Britton, and A.J. Wilkinson, *Geometrically necessary dislocation density distributions in Ti-6Al-4V deformed in tension*, Acta. Mater. 59 (2011), pp. 6489–6500.
[23] R. Luss, S. Rosset, and M. Shahar, *Efficient regularized isotonic regression with application to gene-gene interaction search*, Ann. Appl. Stat. 6 (2012), pp. 253–283.
[24] A.T. Minhajuddin, W.H. Frawley, W.R. Schucany, and W.A. Woodward, *Bootstrap tests for multivariate directional alternatives*, J. Stat. Plan. Inference. 137 (2007), pp. 2302–2315.
[25] N. Misra and E.C. van der Meulen, *On estimation of the common mean of $k(\geq 2)$ normal populations with order restricted variances*, Stat. Probab. Lett. 36 (1997), pp. 261–267.
[26] C. Moussa, M. Bernacki, R. Besnard, and N. Bozzolo, *Statistical analysis of dislocations and dislocation boundaries from EBSD data*, Ultramicroscopy 179 (2017), pp. 63–72.
[27] H.T. Mutlu, F. Gökpinar, E. Gökpinar, H.H. Gül, and G. Güven, *A new computational approach test for one-way ANOVA under heteroscedasticity*, Commun. Statist.-Theor. Meth. 46 (2017), pp. 8236–8256.
[28] M.D. Perlman, *One-sided testing problems in multivariate analysis*, Ann. Math. Stat. 40 (1969), pp. 549–567.
[29] N.J. Petch, *The cleavage strength of polycrystals*, J. Iron Steel Instit. 174 (1953), pp. 25–28.
[30] E. Pippel, J. Woltersdorf, G. Pöckl, and G. Lichtenegger, *Microstructure and nanochemistry of carbide precipitates in high-speed steel S 6-5-2-5*, Mater. Charact. 43 (1999), pp. 41–55.
[31] C. Revilla, B. López, and J.M. Rodriguez-Ibabe, *Carbide size refinement by controlling the heating rate during induction tempering in a low alloy steel*, Mater. Des. 62 (2014), pp. 296–304.
[32] T. Robertson and F.T. Wright, *On approximation of the level probabilities and associated distributions in order restricted inference*, Biometrika 70 (1983), pp. 597–606.
[33] T. Robertson, F. Wright, and R.L. Dykstra, *Order Restricted Statistical Inference*, New York, Wiley, 1988.
[34] G. Salanti and K. Ulm, *Multidimensional isotonic regression and estimation of the threshold value*, Ph.D. diss., 2001.
[35] S. Sasabuchi, *More powerful tests for homogeneity of multivariate normal mean vectors under an order restriction*, Sankhy?: The Indian J. Stat. 69 (2007), pp. 700–716.
[36] N.Z. Shi, *Maximum likelihood estimation of means and variances from normal populations under simultaneous order restrictions*, J. Multivar. Anal. 50 (1994), pp. 282–293.
[37] N.Z. Shi and H. Jiang, *Maximum likelihood estimation of isotonic normal means with unknown variances*, J. Multivar. Anal. 64 (1998), pp. 183–195.
[38] N.Z. Shi, G.R. Hu, and Q. Cui, *An alternating iterative method and its application in statistical inference*, Acta Mathematica Sinica 24 (2008, English Series), pp. 843–856.
[39] M.J. Silvapulle and P.K. Sen, *Constrained Statistical Inference: Inequality, Order and Shape Restrictions*, 1st ed. John Wiley & Sons, Hoboken, NJ, 2005.

[40] M. Stylianou and N. Flournoy, *Dose finding using the biased coin up and down design and isotonic regression*, Biometrics 58 (2002), pp. 171–177.
[41] C. Van Eeden, *Note on two methods for estimating ordered parameters of probability distributions*, Indagationes Mathematicae (Proceedings); North-Holland; 1957, Vol. 60, pp. 506–512.

Appendix 1

A.1 ALGORITHM 2.1

0. INITIALIZATION Let $\boldsymbol{\mu}^{(0)} = \bar{\boldsymbol{y}} = (\bar{y}_1, \dots \bar{y}_k)'$ and $\boldsymbol{w} = (w_1, \dots w_k)'$, $w_i = \frac{n_i}{c_i}$
1. QUESTION Is $\bar{y}_1 \leq \bar{y}_2 \leq \cdots \leq \bar{y}_k$?
1.1 YES $\boldsymbol{\mu}^* = \bar{\boldsymbol{y}}$ is the solution
1.2 NO $\bar{y}_i > \bar{y}_{i+1}$
Replace $\bar{y}_i$ and $\bar{y}_{i+1}$ by

$$m_{i,i+1} = \frac{w_i \bar{y}_i + w_{i+1} \bar{y}_{i+1}}{w_i + w_{i+1}} \tag{A1}$$

Repeat until QUESTION 1 is satisfied.

A.2 ALGORITHM 2.2

Two steps Iterative procedure.

0. INITIALIZATION Let $\boldsymbol{\mu}^{(\boldsymbol{0})} = \bar{\boldsymbol{y}} = (\bar{y}_1, \dots \bar{y}_k)'$, $\boldsymbol{\sigma}^{\boldsymbol{2(0)}} = \bar{\boldsymbol{\sigma}}^{\boldsymbol{2}} = (\bar{\sigma}_1^2, \dots \bar{\sigma}_k^2)'$, $\bar{\sigma}_i^2 = \frac{\sum_j (y_{ij} - \bar{y}_i)^2}{n_i}$ and $\boldsymbol{w}^{(\boldsymbol{0})} = (w_1^{(0)}, \dots w_k^{(0)})'$, $w_i^{(0)} = \frac{n_i}{\sigma^{2(0)}}$
1. QUESTION Is $\bar{y}_1 \leq \bar{y}_2 \leq \cdots \leq \bar{y}_k$?
1.1 YES $\boldsymbol{\mu}^{\boldsymbol{*}} = \boldsymbol{\mu}^{(\boldsymbol{0})} = \bar{\boldsymbol{y}}$ and $\boldsymbol{\sigma}^{\boldsymbol{2}} = \boldsymbol{\sigma}^{\boldsymbol{2(0)}} = \bar{\boldsymbol{\sigma}}^{\boldsymbol{2}}$ are the solutions
1.2 NO Use Step 1.2 Algorithm 2.1 to compute $\boldsymbol{\mu}^{(\boldsymbol{l})}$ with weights $\boldsymbol{w}^{(\boldsymbol{l-1})}$
1.2.1 Compute $\boldsymbol{\sigma}^{\boldsymbol{2}(\boldsymbol{l})} = \boldsymbol{s}^{\boldsymbol{2}(\boldsymbol{l})}$, $s_i^{2(l)} = \frac{\sum_j (y_{ij} - \mu_i^{(l)})^2}{n_i}$ and $\boldsymbol{w}^{(\boldsymbol{l})} = (w_1^{(l)}, \dots w_k^{(l)})'$, $w_i^{(l)} = \frac{n_i}{\sigma^{2(l)}}$
1.2.2 Go back to QUESTION 1 using $\boldsymbol{w}^{(\boldsymbol{l})}$.
Repeat until

$$\max_{1 \leq i \leq k} |\mu_i^{*(l-1)} - \mu_i^{*(l)}| \leq 10^{-m}$$

Alternating Iterative Method.

0. INITIALIZATION Let $\boldsymbol{\nu}^{(\boldsymbol{0})} = (1/\bar{\sigma}_1^2, \dots 1/\bar{\sigma}_k^2)'$, $\bar{\sigma}_i^2 = \frac{\sum_j (y_{ij} - \bar{y}_i)^2}{n_i}$
1. FIND $\boldsymbol{\mu}^{(\boldsymbol{l})}$ the isotonic regression on D_a using weights $\boldsymbol{w}^{(\boldsymbol{l-1})} = (w_1^{(l-1)}, \dots w_k^{(l-1)})'$, $w_i^{(l-1)} = n_i \nu^{(l-1)}$;
2. FIND $\boldsymbol{\nu}^{(\boldsymbol{l})}$ maximizing the profile likelihood $L(y; \boldsymbol{\mu}^{(\boldsymbol{l})}; \boldsymbol{\nu})$ on V_0, $V_0 = \{\nu \in \mathbb{R}^k : 0 \leq 1/\max_i (\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta)) \leq \nu_1 \leq \cdots \leq \nu_k \leq 1/\min_i (\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta))\}$.
Repeat (1)-(2) until

$$|L(y; \boldsymbol{\mu}^{(\boldsymbol{l-1})}, \boldsymbol{\nu}^{(\boldsymbol{l-1})}) - L(y; \boldsymbol{\mu}^{(\boldsymbol{l})}, \boldsymbol{\nu}^{(\boldsymbol{l})})| \leq 10^{-m}$$

A.3 ALGORITHM 2.3

Two steps Iterative procedure.

0. INITIALIZATION Let $\boldsymbol{\mu^{(0)}} = \bar{\boldsymbol{y}} = (\bar{y}_1, \dots \bar{y}_k)'$, $\boldsymbol{\sigma^{2(0)}} = \bar{\boldsymbol{\sigma}}^{\boldsymbol{2}} = (\bar{\sigma}_1^2, \dots \bar{\sigma}_k^2)'$, $\bar{\sigma}_i^2 = \frac{\sum_j (y_{ij} - \bar{y}_i)^2}{n_i}$ and and $\boldsymbol{w^{(0)}} = (w_1^{(0)}, \dots w_k^{(0)})'$, $w_i^{(0)} = \frac{n_i}{\sigma^{2(0)}}$.

1. QUESTION Is $\bar{y}_1 \leq \bar{y}_2 \leq \cdots \leq \bar{y}_k$?

1.1 YES $\boldsymbol{\mu^*} = \boldsymbol{\mu^{(0)}} = \bar{\boldsymbol{y}}$ go to QUESTION 2.

1.2 NO Use Step 1.2 Algorithm 2.1 to compute $\boldsymbol{\mu^{(l)}}$ with weights $\boldsymbol{w^{(l-1)}}$

1.2.1 Compute $\boldsymbol{\sigma^{2(l)}} = \boldsymbol{s^{2(l)}}$, $s_i^{2(l)} = \frac{\sum_j (y_{ij} - \mu_i^{(l)})^2}{n_i}$ and $\boldsymbol{w^{(l)}} = (w_1^{(l)}, \dots w_k^{(l)})'$, $w_i^{(l)} = \frac{n_i}{\sigma^{2(l)}}$

2. QUESTION Is $\sigma_1^{2(l)} \geq \sigma_2^{2(l)} \geq \cdots \geq \sigma_k^{2(l)}$?

2.1 YES $\boldsymbol{\mu^*} = \boldsymbol{\mu^{(l)}}$ and $\boldsymbol{\sigma^{2*}} = \boldsymbol{\sigma^{2(l)}}$ are the solutions.

2.2 NO $\sigma_i^{2(l)} < \sigma_j^{2(l)}$

Replace $\sigma_i^{2(l)}$ and $\sigma_{i+1}^{2(l)}$ by

$$\sigma_i^{2(l+1)} = \sigma_{i+1}^{2(l+1)} = \bar{s}_{i,i+1} = \frac{n_i \sigma_i^{2(l)} + n_{i+1} \sigma_{i+1}^{2(l)}}{n_i + n_{i+1}} \tag{A2}$$

Repeat until QUESTION 2 is satisfied.

2.2.1 Go back to QUESTION 1 using $\boldsymbol{w^{(l)}}$.

Repeat until

$$\max_{1 \leq i \leq k} |\mu_i^{*(l-1)} - \mu_i^{*(l)}| \leq 10^{-m} \quad \text{and} \quad \max_{1 \leq i \leq k} |\sigma_i^{2*(l-1)} - \sigma_i^{2*(l)}| \leq 10^{-m}$$

Alternating Iterative Method.

0. INITIALIZATION Let $\boldsymbol{\nu^{(0)}} = (1/\bar{\sigma}_1^2, \dots 1/\bar{\sigma}_k^2)'$, $\bar{\sigma}_i^2 = \frac{\sum_j (y_{ij} - \bar{y}_i)^2}{n_i}$

1. FIND $\boldsymbol{\mu^{(l)}}$ use Step 1 AlM Algorithm 2.2;

2. FIND $\boldsymbol{\nu^{(l)}}$ the isotonic regression on V_0, $V_0 = \{\nu \in \mathbb{R}^k : 0 \leq 1/\max_i(\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta)) \leq \nu_1 \leq \cdots \leq \nu_k \leq 1/\min_i(\min_{\min(\bar{\boldsymbol{y}}) \leq \theta \leq \max(\bar{\boldsymbol{y}})} s_i^2(\theta))\}$ with weights $\boldsymbol{N} = (n_1, \dots, n_k)'$ Repeat (1)–(2) until

$$|L(y; \boldsymbol{\mu^{(l-1)}}, \boldsymbol{\nu^{(l-1)}}) - L(y; \boldsymbol{\mu^{(l)}}, \boldsymbol{\nu^{(l)}})| \leq 10^{-m}$$

Arxiv Preprint: arXiv:2002.00993