USING ISOTONIC REGRESSION TO IMPROVE ESTIMATION IN FACTORIAL EXPERIMENTS WITH ORDERED FACTOR LEVELS

Matt Strand
Jim Higgins

Follow this and additional works at: http://newprairiepress.org/agstatconference

Part of the Agriculture Commons, and the Applied Statistics Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Recommended Citation
USING ISOTONIC REGRESSION TO IMPROVE ESTIMATION IN FACTORIAL EXPERIMENTS WITH ORDERED FACTOR LEVELS

By Matt Strand and Jim Higgins

Abstract

In many designed experiments in agriculture and the life sciences, a researcher can anticipate the direction that responses will take when treatments are varied. For example, in a 2-way factorial, a researcher may know that increasing the levels of nitrogen and phosphorus will increase yields of a crop. Classical analysis of variance does not take into account a known ordering among population means. However, it can be shown that by restricting the estimates of means to have the same ordering as the anticipated ordering of population means, a reduction in mean-squared errors of estimators will likely occur, often by more than 50%. A procedure used to create such estimates is called isotonic regression.

In this article, a recently proposed method of isotonic regression for lattice-ordered means will be presented after first reviewing well-established methods. The newer method will be illustrated using data from an entomology experiment. In addition, standard errors of the estimators will be approximated using a bootstrap procedure.

Key words: lattice order, parametric bootstrap procedure

I. Introduction

Designed experiments with factorial treatment structures are carried out and analyzed quite routinely, even by people who are not regular practitioners of statistics. Most do not give much thought about which statistical techniques to use for data produced by such experiments. The common test for some difference in treatment means is ANOVA, which is sometimes followed by more specific multiple comparison tests. Treatment means are generally estimated with corresponding sample means.

In many experiments it may be reasonable to assume that the treatment means will have a particular order. To illustrate, consider an experiment in agriculture where the yield of a certain crop is measured in the absence or presence of two fertilizers, nitrogen and phosphorus. The design is completely randomized with a 2x2 factorial treatment structure. Let \( \mu_{ij} \) represent the true treatment mean for \( i=1 \) (absence of nitrogen), \( 2 \) (presence of nitrogen) and \( j=1 \) (absence of phosphorus), \( 2 \) (presence of phosphorus). If the amount of fertilizer given for “presence” is not too great, then it seems reasonable to assume that a crop given one fertilizer would have higher mean yield than one given no fertilizer (\( \mu_{11} < \mu_{12} \), \( \mu_{11} < \mu_{21} \)). Also, a crop given both fertilizers would have higher mean yield than one given either only one fertilizer (\( \mu_{12} < \mu_{22} \), \( \mu_{21} < \mu_{22} \)) or no fertilizer (\( \mu_{11} < \mu_{22} \)). It is more uncertain how the mean yields would compare between a crop given only phosphorus and another given only nitrogen (\( \mu_{12} < \mu_{21} \)). This particular type of order on means is referred to as lattice order.

The formal definition for lattice-ordered means in a 2-factor experiment follows naturally.
Let $i$ and $j$ be indices for the first and second factors, respectively, where $1 \leq i \leq m_1$ and $1 \leq j \leq m_2$. Then the means are lattice-ordered if $\mu_{ij} \leq \mu_{i'j'}$ for each $(ij)$ and $(i'j')$ such that $1 \leq i \leq i' \leq m_1$ and $1 \leq j \leq j' \leq m_2$. (This definition can be generalized for experiments with more than two factors.) Lattice-ordered treatment means will occur naturally in many designed experiments in which the factors have ordinal levels. The researchers themselves may often have the best insight as to whether the assumption is appropriate for their experiment.

For many different experimental conditions, it has been shown that estimators that correctly adjust for lattice-ordered means have reduced MSE over those that do not take this order restriction into account (Strand, 1998). The reduction is often by more than 50%, averaged over all treatments. A method of estimation used for the treatment means when they can be ordered is isotonic regression. The word “isotonic” is used to denote that the estimates produced by this method necessarily have the same order as is assumed on the population means. Section 2 presents an example along with the standard statistical analysis. Certain methods of isotonic regression for lattice-ordered means are introduced in Section 3, and then Section 4 applies a method of isotonic regression to the data first presented in Section 2.

II. Example in Entomology: Classical Approach

Table 1 presents data collected from an entomology experiment. The response is oviposition of horse ticks (number of days female ticks lay eggs), subjected to various levels of temperature and humidity (Despins, 1992). The experiment consisted of a 4x4 factorial treatment structure in a completely randomized design, with unequal but similar sample sizes (in parentheses). In certain ranges of temperature and humidity it may be reasonable to assume that as temperature is decreased and/or humidity is increased, then the oviposition period will increase. Indeed, the sample means reflect a general increase from left to right and top to bottom. The increase is more noticeable stepping downward as opposed to stepping to the right, suggesting that temperature has a stronger effect on the response than humidity. In this case the mean responses across treatments is hypothesized to be lattice-ordered when temperature is ordered from high to low and humidity is ordered from low to high.

The sample standard deviations for the treatments range from 2.33 to 6.09, with no real patterns of increase or decrease from left to right and top to bottom. Assuming that a common standard deviation exists for all replicates among all treatments, call it $\sigma$, then a pooled estimate of this parameter is $\hat{\sigma}=4.24$. The estimated standard error (SE) of a sample mean based on 20 replicates would then be $\hat{\sigma}_x=(4.24/\sqrt{20})=0.95$. Similarly, the SE's are 0.97 and 1.00 for sample sizes of 19 and 18, respectively. Using the normal theory approach, one can then construct 95% confidence intervals for the treatment means using $\bar{X}_{ij}\pm2\cdot\text{SE}$, for all $i$ and $j$. These standard errors and confidence intervals are based on unrestricted treatment means; no assumption of lattice-ordered means is used.

III. Isotonic Regression for Lattice-Ordered Means

In a designed experiment, if the treatment means are assumed to be lattice-ordered, then isotonic regression can be applied to the corresponding sample means to ensure that the estimates
are also lattice ordered. Thus isotonic regression can be thought of as the 2nd step of an estimation procedure, where the 1st step is summary of the treatments with the respective sample means (or some other consistent measure of the corresponding population means).

The most common isotonic regression is least squares isotonic regression (LSIR). In fact, many use the term “isotonic regression” to specifically mean LSIR. In this article, the term is used more generally. As the name implies, LSIR produces estimates that minimize the sum of squared distances between the original estimates (treatment sample means) and all possible lattice-ordered solutions. Using mathematical notation, let the LSIR estimators be denoted as \{\hat{h}_{ij}: i=1,2,\ldots,m_1, j=1,2,\ldots,m_2\}, or more simply, \{\hat{h}_{ij}\}. Let \{h_{ij}\} be any real-valued, lattice-ordered array. (That is, \(h_{ij} \leq h_{i'j'}\) for all \(i \leq i'\) and \(j \leq j'\).) Then the LSIR estimators satisfy the following:

\[
\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (\bar{X}_{ij} - \hat{\mu}_{ij})^2 w_{ij} \leq \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (\bar{X}_{ij} - h_{ij})^2 w_{ij}
\]

where \{\bar{X}_{ij}\} are the treatment sample means, and \{w_{ij} = n_{ij}/\sigma_{ij}^2\} are weights. If \{\bar{X}_{ij}\} is lattice ordered, then the sum of squares are minimized at zero and the LSIR will have no effect. There are different methods that can be used to obtain a least squares isotonic solution for lattice-ordered means. One such method was introduced by Dykstra and Robertson (1982) which incorporates the Pool-adjacent-violators algorithm (PAVA). This method can be carried out for a given data set with relative ease, using computer programs.

Least absolute deviation isotonic regression (LADIR) is similar to LSIR, but the solution minimizes the sum of absolute distances. A method which incorporates the Max-min formulas (see Barlow, 1972 and Robertson, 1988) can be used to perform LADIR. (This method can also be used to perform LSIR.) Unfortunately, this method is inefficient for even modest sized data sets \((k=2, m_1=6, m_2=6)\) using computers. This is due to the large number of combinations the program must search through to find the solution.

An alternative approach to finding an isotonic solution for lattice-ordered means is based on estimators presented by Mukerjee and Stern (1994). This method uses simple averages of maximums and minimums of subsets of the data. Formally, for a designed experiment, this set of estimators \{M_{ij}\} can be defined as

\[
M_{ij} = \frac{\text{Min}_{ij} + \text{Max}_{ij}}{2}
\]

for all \((i,j)\)

where

\[
\text{Min}_{ij} = \text{Min}\{ \bar{X}_{i'j'}: (i'j') \succ (i,j) \}
\]

and

\[
\text{Max}_{ij} = \text{Max}\{ \bar{X}_{i'j'}: (i'j') \preceq (i,j) \}.
\]

While necessarily being a lattice-ordered solution, these estimators satisfy the uniform maximum likelihood. I.e., if \(X_{ij} \sim \text{iid} \ U(\mu_{ij}-a, \mu_{ij}+a)\) for each \((i,j)\), then \{M_{ij}\} are maximum likelihood estimators. (Note that \(\bar{X}_{ij}\) is not likely to have a uniform distribution if for any \((i,j),\)
Due to the Central Limit Theorem. Hence this procedure that yields \( \{M_{ij}\} \) will be called Uniform Maximum Likelihood Isotonic Regression (UMLIR).

Simulation studies were performed to compare efficiencies of isotonic regression methods under a variety of experimental conditions (Strand, 1998). The conditions included distribution of \( \bar{X}_{ij} \) (varied by replicate distribution and treatment sample size), form of population treatment means, and distance between population treatment means. Reduction in simulated MSE (experiment-wide) that each isotonic regression method had over that of ordinary treatment sample mean estimation under the same conditions was determined for each simulation. In more than 200 simulations tried, the distribution of \( \bar{X}_{ij} \) had the strongest influence on relative performance of the isotonic regression methods. When each \( \bar{X}_{ij} \) was distributed uniformly, UMLIR had the best efficiency. Similarly, for normal to medium-tailed distributions, LSIR had the best efficiency. LADIR performed the best for heavy-tailed distributions, as expected.

For 4x4 experiments, the difference in efficiency between UMLIR and LSIR was minor for both the uniform and normal distributions. Given that the efficiency of the isotonic regression methods have roughly the same efficiency for light to moderate-tailed distributions, UMLIR has an advantage due to the fact that the estimates are easier to calculate than those of LSIR. In fact, UMLIR estimates can be determined by hand, as will be shown in the next section.

### IV. Example in Entomology: Order-Restricted Approach

Estimates of the treatment means for the entomology experiment are given in Table 2, based on the UMLIR procedure. Changes made to sample means are highlighted in grey. Bootstrap estimates of standard error are also given in the table. (Both treatment mean estimates and standard errors using LSIR were quite similar, and thus will not be presented.) To illustrate how an estimate is determined, consider the treatment associated with 35 °C temperature and 61% relative humidity and refer to the original sample means in Table 1. Now

\[
\begin{align*}
\text{Min}_{12} &= \text{Min}\{ \bar{X}_{ij}: i \geq 1, j \geq 2 \} \\
&= \text{Min}\{13.7, 14.4, 14.3, 15.1, 13.2, 15.7, 20.6, 23.4, 23.8, 29.1, 30.5, 30.2 \} \\
&= 13.2
\end{align*}
\]

and

\[
\begin{align*}
\text{Max}_{12} &= \text{Max}\{ \bar{X}_{ij}: i \leq 1, j \leq 2 \} \\
&= \text{Max}\{13.1, 13.7 \} \\
&= 13.7
\end{align*}
\]

and thus

\[
M_{12} = \frac{(\text{Min}_{12} + \text{Max}_{12})}{2} = \frac{(13.2 + 13.7)}{2} = 13.45.
\]

The other estimates can be determined similarly.

The standard errors of the estimates were determined using a parametric bootstrap method. Let \( k \) denote the replicate within a treatment and let \( n_{ij} \) denote the sample size within each treatment. It was assumed that the model
\[ Y_{ijk} = \mu_{ij} + \epsilon_{ijk}, \text{ where } \epsilon_{ijk} \sim N(0, \sigma^2) \text{ for } i=1, \ldots, m_1, j=1, \ldots, m_2 \text{ and } k=1, \ldots, n_{ij} \]

and where \( \mu_{ij} \leq \mu_{ij'} \) for \( i \leq i' \) and \( j \leq j' \)

was reasonably accurate for the data. Using the pooled estimate of \( \sigma^2 \), which was \( \hat{\sigma}^2 = 17.98 \), bootstrap errors, \( \epsilon_{ijk}^* \), were created by resampling from the normal distribution, \( N(0, 17.98) \). Bootstrap responses were then determined from the model

\[ Y_{ijk}^* = M_{ij}^* + \epsilon_{ijk}^* \]

and subsequently new bootstrap estimates, \( \{ M_{ij}^* \} \), were determined from \( \{ Y_{ijk}^* \} \). Repeating this 2000 times yielded \( M_{ij}^{*b}, b=1, \ldots, 2000 \) for each \((i,j)\). The estimated standard error for \( M_{ij} \) is then the standard deviation of \( \{ M_{ij}^{*b}, b=1, \ldots, 2000 \} \). Table 3 displays the bootstrap standard errors of \( \{ M_{ij} \} \). For a comparison, the theoretical standard errors of \( \{ \hat{X}_{ij} \} \) are displayed in parentheses. The standard errors for the order-restricted estimators are anywhere from 14% to 41% less than those of the sample means (comparing within treatments).

The Appendix contains a SAS program to calculate UMLIR estimates and bootstrap estimates of standard error for a two-factor experiment, given the sample means. The generalization to experiments with three or more factors is straightforward.

### V. Discussion

One might wonder why isotonic regression estimators are not used more often, based on some of the results presented and/or discussed in this article. Some difficulties may be due to general lack of knowledge and familiarity of the methods and lack of understanding of when the methods are appropriate.

The fact that isotonic regression estimators are not used due to people’s lack of knowledge of the methods is clear. However, some may know of the methods, but are reluctant to use them due to lack of familiarity. The standard analyses for designed experiment data is straightforward and is even taught in introductory statistics at some schools. People may not realize the gains that using isotonic regression has over the standard analyses, and decide to go with what is easy and well known as opposed to making that extra step to obtain more accurate estimators. Another legitimate reason that people might be hesitant to use isotonic regression is that they do not know whether the means are lattice-ordered, or are not willing to make this assumption. (This applies more generally to any type of order restriction.) To many, the fewer assumptions that are made, the more appeal that a method has.

There are many times when an order restriction on the treatment means is very reasonable and appropriate. Lattice-ordered treatment means will occur in many experiments where a dose-response exists, with respect to one or more factors. The agriculture and entomology scenarios presented herein are two such examples.

Although it may be easy to find a variable whose mean response is lattice ordered with respect to certain factors, violations to the assumption might occur if the levels of the factors are not considered carefully. For instance, in an extension of the agricultural example presented in the
introduction, consider yield of a crop measured in response to levels of nitrogen and phosphorus, each having zero, low, medium and high levels, for a 4x4 factorial experiment. The mean response may increase for most of the levels, but if the “high” level is too high, then the crop may be over-saturated and the mean yield may actually go down. In the entomology example, if levels of temperature and/or humidity that are extremely high are included in the experiment, then mean tick laying period may start to decrease. Similarly, the mean response may not follow lattice order if levels are chosen at an extremely low level. Even so, if the violations occur in only a few of the treatment means, then a reduction in MSE is possible (averaged over all treatments) over the unrestricted case. Thus like many other assumptions necessary to perform statistical methods, violations to the assumption of lattice order occur by degree.

Isotonic regression should be considered as an alternative estimation method for certain designed experiments with factorial treatment structures. If the researcher feels that the treatment means are lattice ordered with respect to the given factors, then more accurate estimators can be used through isotonic regression. It has been shown that both MSE and standard errors of estimators are significantly reduced when used in appropriate situations.
Table 1: Mean Oviposition Period for Horse Ticks for Certain Temperatures and Humidities, with Sample Sizes in Parentheses

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Relative Humidity (%)</th>
<th>40</th>
<th>61</th>
<th>75</th>
<th>91</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>13.1 (20)</td>
<td>13.7 (20)</td>
<td>14.4 (19)</td>
<td>14.3 (19)</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>14.7 (20)</td>
<td>15.1 (20)</td>
<td>13.2 (20)</td>
<td>15.7 (20)</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>20.1 (20)</td>
<td>20.6 (19)</td>
<td>23.4 (20)</td>
<td>23.8 (20)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>28.9 (20)</td>
<td>29.1 (18)</td>
<td>30.5 (20)</td>
<td>30.2 (20)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Lattice-Ordered Estimates of Horse Tick Data using UMLIR, with Bootstrap Standard Errors in Parentheses

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Relative Humidity (%)</th>
<th>40</th>
<th>61</th>
<th>75</th>
<th>91</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>13.15 (0.69)</td>
<td>13.45 (0.59)</td>
<td>13.80 (0.58)</td>
<td>14.35 (0.70)</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>13.95 (0.63)</td>
<td>14.15 (0.56)</td>
<td>14.15 (0.60)</td>
<td>15.7 (0.81)</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>20.1 (0.82)</td>
<td>20.6 (0.82)</td>
<td>23.4 (0.81)</td>
<td>23.8 (0.82)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>28.9 (0.78)</td>
<td>29.1 (0.72)</td>
<td>30.35 (0.69)</td>
<td>30.35 (0.75)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Standard Errors of \{ M_{ij} \}, with Standard Errors of \{ \overline{X}_{ij} \} in Parentheses

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Relative Humidity (%)</th>
<th>40</th>
<th>61</th>
<th>75</th>
<th>91</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>0.69 (0.95)</td>
<td>0.59 (0.95)</td>
<td>0.58 (0.97)</td>
<td>0.70 (0.97)</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.63 (0.95)</td>
<td>0.56 (0.95)</td>
<td>0.60 (0.95)</td>
<td>0.81 (0.95)</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.82 (0.95)</td>
<td>0.82 (0.97)</td>
<td>0.81 (0.95)</td>
<td>0.82 (0.95)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.78 (0.95)</td>
<td>0.72 (1.00)</td>
<td>0.69 (0.95)</td>
<td>0.75 (0.95)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix: SAS Program to Calculate UMLIR Estimates and Bootstrap Standard Errors

/*Assign values to variables and declare arrays*/

options ps=60 ls=80;
data first;
  orow=4;
  ocol=4;
  reps=2000;
array x(4,4) x11-x14 x21-x24 x31-x34 x41-x44;
  x11=13.1; x12=13.7; x13=14.4; x14=14.3;
  x21=14.7; x22=15.1; x23=13.2; x24=15.7;
  x31=20.1; x32=20.6; x33=23.4; x34=23.8;
  x41=28.9; x42=29.1; x43=30.5; x44=30.2;
array nn(4,4) nn11-nn14 nn21-nn24 nn31-nn34 nn41-nn44;
  nn11=20; nn12=20; nn13=19; nn14=19;
  nn21=20; nn22=20; nn23=20; nn24=20;
  nn31=20; nn32=19; nn33=20; nn34=20;
  nn41=20; nn42=18; nn43=20; nn44=20;
maxn=max(of nn11-nn14 nn21-nn24 nn31-nn34 nn41-nn44);
array maxx( 4,4) maxx 11-maxx 14 maxx21-maxx24 maxx31-maxx34 maxx41-maxx44;
array minn(4,4) minn11-minn14 minn21-minn24 minn31-minn34 minn41-minn44;
array m(4,4) m11-m14 m21-m24 m31-m34 m41-m44;
array eboot(20) eboot1-eboot20;
array yboot(20) yboot1-yboot20;
array ybav(4,4) ybav11- ybav 14 ybav21- ybav24 ybav 31- ybav34 ybav41- ybav44;
array mb(4,4) mb11-mb14 mb21-mb24 mb31-mb34 mb41-mb44;
/* Obtain UMLIR estimates */
do i=1 to orow;
do j=1 to ocol;
  maxx(i,j)=x(i,j);
  minn(i,j)=x(i,j);
  do ii=1 to i;
    do jj=1 to j;
      if x(ii,jj) ge maxx(i,j) then maxx(i,j)=x(ii,jj);
    end;
  end;
  do ii=i to orow;
    do jj=j to ocol;
      if x(ii,jj) le minn(i,j) then minn(i,j)=x(ii,jj);
    end;
  end;
m(i.j)=(maxx(i,j)+minn(i,j))/2;
end;
end;
/* Obtain bootstrap standard errors */

do k=1 to reps;
  do i=1 to nrow;
    do j=1 to ncol;
      do h=1 to maxnn;
        if h le nn(i,j) then
          do;
            eboot(h)=4.24*rannor(42323);
            yboot(h)=m(i,j)+eboot(h);
          end;
        end;
        if h > nn(i,j) then
          do;
            eboot(h)=.; yboot(h)=.;
          end;
        end;

        ybav(i,j)=mean(of yboot1-yboot20);
      end;
    end;
  end;
end;

do i=1 to nrow;
  do j=1 to ncol;
    maxx(i,j)=ybav(i,j);
    minn(i,j)=ybav(i,j);
    do ii=1 to i;
      do jj=1 to j;
        if ybav(ii,jj) ge maxx(i,j) then maxx(i,j)=ybav(ii,jj);
      end;
    end;
    do ii=i to nrow;
      do jj=j to ncol;
        if ybav(ii,jj) le minn(i,j) then minn(i,j)=ybav(ii,jj);
      end;
    end;
    mb(i,j)=(maxx(i,j)+minn(i,j))/2;
  end;
end;
output;

proc means noprint;
  var mb11-mb14 mb21-mb24 mb31-mb34 mb41-mb44;
  output out=second
    std=stdb11-stdb14 stdb21-stdb24 stdb31-stdb34 stdb41-stdb44;
run;
/* Obtain the final data and write the output. Note "m" represents
the UMLIR estimates and "stdb" represents the bootstrap standard errors.
The last two numbers on each of the variables denotes the treatment. For
example, "stdb13" is the bootstrap standard error for the row 1 and column
3 treatment. */

data last;
  set first(keep=reps m11-m14 m21-m24 m31-m34 m41-m44);
  set second(drop= _TYPE_ _FREQ_);
  run;
  proc print noobs;
  run;

The Bibliography

Barlow, R., Bartholomew, J., Bremner, J. and Brunk, H. (1972). Statistical Inference under

Despins, J. (1992). "Effects of Temperature and Humidity on Ovispositional Biology and Egg
Development of the Tropical Horse Tick, Dermacentor (Anocentor) Nitens." Journal of Medical
Entomology, 29, 332-336.


Wiley & Sons.