Why CALFED Needs Ecological Detectives

Marc Mangel
University of California
Santa Cruz, CA 95064

Outline of the talk:

¥What We’d Like to Know and What The Usual Methods Give Us

¥Water Flow and Fish Growth

¥Dealing with Contradictory Data
The Earth is Round (p< 0.05)

“I argue herein that NHST (null hypothesis significance testing) has not only failed to support the advance of psychology as a science but also has seriously impeded it” (J. Cohen. 1994. American Psychologist. 49:997-1003)

What is a null hypothesis?

Null hypotheses entertain the possibility that nothing has happened, that a process has not occurred, or that change has not been produced by a cause of interest. They are reference points against which alternatives should be contrasted.

However, since it is often impossible to prove that something has occurred, we construct a null hypothesis that is the complement of the hypothesis of interest and use the accumulated data to assess the probability that the null hypothesis is true.
Examples of Null Hypotheses That Were Rejected

• The occurrence of sheep remains in coyote scats does not vary across season (p=0.03)

• Duckling body mass does not vary across years (p<0.0001)

• The density of large trees is not greater in unlogged than logged forests (p=0.02)


• Driving cessation [in the elderly] leads to a decline in out-of-home activity (p<0.001)

What’s wrong with NHST?

Well, among many other things, it does not tell us what we want to know, and we so much want to know what we want to know that, out of desperation, we nevertheless believe that it does.

What we want to know is

Given these data, what is the probability that the null hypothesis is true?

But, as most of us know, what it tells us is

Given that the null hypothesis is true, what is the probability of these (or more extreme) data

(Cohen pg 997)

What we want for scientific understanding

\[ \Pr\{H|D\} \neq \Pr\{D|H\} \]

NHST has “caused scientific research workers to pay undue attention to the results of the tests of significance they perform on their data..and too little to the estimates of the magnitude of the effects they are estimating”

But Sometimes You Might (Jagger and Richards, op. cit.)

Likelihood Shows the Way to Get What You Want

“...the discipline of statistics has neglected a key question for which it is responsible: when does a given set of observations support one statistical hypothesis over another?”

The principle fundamental to providing this answer is the law of likelihood, which “provides the explicit objective quantitative concept of evidence that is missing.”

**Law of Likelihood:** If hypothesis A implies that the probability that a random variable X takes the value x is \( p_A(x) \), while hypothesis B implies that the probability is \( p_B(x) \), then the observation \( X=x \) is evidence supporting A over B only if \( p_A(x) > p_B(x) \), and the likelihood ratio, \( p_A(x)/p_B(x) \), measures the strength of that evidence.

Ecological Detection in Environmental Problem Solving

”Method of multiple working hypotheses”
--T.C. Chamberlain (1890)

The three questions we need to ask

Given two (or more) hypotheses/models and an observation (data) we can ask

• What do I believe, now that I have this observation?

• What should I do, now that I have this observation?

• How should I interpret this observation as evidence regarding the different models/hypotheses?
An Illustrated Guide to the Principles of Ecological Detection

Example: Water flow and fish growth rates

Null hypothesis: There is no relationship between water flow and fish growth rates

Multiple models:

Model 1: No relationship
Model 2: Linear
Model 3: Asymptotic
Model 4: Peak
Graphical Representation of the Four Models

- Model 1
- Model 2
- Model 3
- Model 4

Normalized growth rate vs. Normalized flow rate
ED Principle #1: Always Plot Your Data
ED Principle #2: Explore with Standard Statistical Methods

\[ y = 0.71461 + 1.3123x \quad R = 0.68169 \]
One way analysis of variance:

$$F = 34.377, p < 0.0001$$

**Conclusion:** Reject the null hypothesis that fish growth does not depend upon flow rate.
ED Principle #3: Take Advantage of Models

Model 1: Growth = $q_0$  
(Already rejected)

Model 2: Growth = $q_0$ Flow  
One parameter

Model 3: Growth = $q_0$ Flow/(Flow + $f_0$)  
Two parameters

Model 4: Growth = $q_0$ Flow - $f_1$Flow$^2$  
Two parameters
ED Principle #4: Confront Different Models with Data

Confrontation judged by

¥ Mallows $C_p$ (Residual sum of squares)

¥ AIC (Assumption about the structure of observation error)
ED Principle #4: Confront Different Models with Data

<table>
<thead>
<tr>
<th>Model</th>
<th>Best Parameters</th>
<th>SSQ*</th>
<th>Mallows Cp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$q_0 = 0.025$</td>
<td>45.9</td>
<td>0.389</td>
</tr>
<tr>
<td>3</td>
<td>$q_0 = .035, f_0 = 0.4$</td>
<td>14.4</td>
<td>0.125</td>
</tr>
<tr>
<td>4</td>
<td>$q_0 = 0.5, f_1 = 2$</td>
<td>12.9</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Conclusion: Models 3 and 4 are very good competitors. Can they be separated?

Answer: Yes, but requires likelihood/AIC methods
## Results of Likelihood Analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>NLL*</th>
<th>Parameters</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>103.1</td>
<td>$q_0 = 0.025$</td>
<td>210.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma = 0.7$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>45.4</td>
<td>$q_0 = 0.035$</td>
<td>96.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma = 0.4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_0 = 0.4$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>40.46</td>
<td>$q_0 = 0.05$</td>
<td>86.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma = 0.4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_1 = 2$</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion: Models 3 and 4 are much better than Model 2.

Furthermore, although Models 3 and 4 have the same number of parameters, model 4 fits as if it had five fewer parameters than Model 4 (each parameter adds 2 to the AIC). It is the winner.
ED Principle #5: Look at The Curves/Surfaces Associated with Estimation
Summary: Principles of Ecological Detection

Fundamental principles:

• Assume there is an effect and try to estimate it.

• Variation is not noise; it is biology.

Specifics:
1. Always plot your data.
2. Use standard statistical procedures when possible.
3. Take advantage of models (there always is a model)
4. Confront different models with data.
5. Look at the curve/surface associated with the estimation.
6. Always test your methods with simulated data.
7. Use Bayesian methods to deal with contradictory data.
The literature

Entry level


Post-entry level


Dealing with Incompatible Data by Likelihood Methods

Understanding data better is always an unsolved problem in astrophysics. Whatever your choice of area, make the choice to live your professional life at a high level of statistical sophistication, and not at the level -- basically freshman lab level -- that is the unfortunate common currency of most astronomers. Thereby we will all move forward together.

Results of 20 Investigators Measuring the Growth-Flow Parameter

![Diagram showing the mean of $q_0$ for 20 investigators. The x-axis represents the investigator number, and the y-axis shows the mean values. The data points are scattered across the graph, indicating variability in the measurements.]
Standard Deviation of The Growth Flow Parameter

![Graph showing standard deviation of q_0 versus investigator number. The x-axis represents investigators, and the y-axis represents the standard deviation values ranging from 0.22 to 0.38. The data points are scattered across the graph.](image-url)
Question #1: Are These Data Compatible?

Model (we always have a statistical model, spoken or unspoken)

\[ q_0(i) = q_0 + \sigma(i)X(i) \]

¥ The estimate of \( q_0 \) is the average of the \( q_0(i) \)

¥ A standard Chi-squared distribution can be used to assess if the data are compatible with each other.

( I have outliers )
Question #2: What Should be Done if the Data Are Not Compatible?

Expand the model. If the measurement is correct then

\[ q_0(i) = q_0 + \sigma(i)X(i) \]

If the measurement is incorrect, then some other model applies, for example

\[ q_0(i) = q_0 + SX(i) \]

where \( S \gg \sigma(i) \)

Alternatively (e.g. Chen, Y. and D. Fournier. 1999. Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries. Canadian Journal of Fisheries and Aquatic Sciences 56:1525-1533)
Bayesian Computation of The Correctness of a Measurement

Assume

\[ p = \text{Prior probability that a measurement is correct} \]

Given a set of observations and an estimate

\[ \{q_0(i), \sigma(i)\} \quad \text{and} \quad q_{0,\text{est}} \]

We want to know

\[ \Pr\{\text{observation } i \text{ is correct} \mid p, q_0(i), \sigma(i), q_{0,\text{est}} \} \]

The devil is in the details.
Caveat: One must choose $p$ and $\sigma$ from some prior information  
Good news: It may not matter too much

<table>
<thead>
<tr>
<th>$p$</th>
<th>$S/\sigma$</th>
<th>Observations Picked as Incorrect (probability incorrect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>2</td>
<td>4 (1) 15 (0.78)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4 (1) 15 (0.88)</td>
</tr>
<tr>
<td>0.7</td>
<td>2</td>
<td>4 (1) 15 (0.93) 20 (0.54)</td>
</tr>
<tr>
<td>0.7</td>
<td>3</td>
<td>4 (1) 15 (0.96) 20 (0.52)</td>
</tr>
</tbody>
</table>