Statistical Methods in Paleontology: Examples from Upper Cretaceous Continental Mollusks

Marron Bingle-Davis
University of North Dakota
and
Sunshine Valley Petroleum Corporation
Introduction

Statistics provide a quantitative foundation to qualitative observations

- Defining taxonomy
- Analyzing species change and spatial relationships

There is no way to know which test provides the most accurate results, so choose the test that best fits your question

- Discriminant Analysis
- Cluster Analysis
- Analysis of Variance
- Chi² Tests
Hell Creek Formation Type Area: Western Williston Basin

Study Area
Typical Study Area Section

HC = Hell Creek Formation
FH = Fox Hills Formation
B = Bearpaw Shale

Image from Hartman and Bingle (2003)
DI = Dakota Isthmus

Paleogeography

Hell Creek Type Area

Western Interior Seaway

(Modified from Erickson, 1999; Hartman and Bingle, 2003)
Whitfield first described these species (1903, 1907) and placed them all into the genus *Unio*.

Russell (1976) reevaluated the descriptions and placed the species into more specialized modern and fossil genera.
The Unionoids

*Plethobasus*

Traditionally, these species are distinguished by the presence of one or two rows of nodular sculpture curving posteroventrally from the umbo.
The elongate forms of this genus are traditionally distinguished by a distinct posterior marginal shape.

The short forms of this genus were also originally distinguished by a different posterior marginal shape.
The Unionoids

Proparreysia

Proparreysia verrucosiformis

Proparreysia letsoni

These species were traditionally distinguished on the basis of either a linear or random node pattern.
**Bivalve Character Traits**

- **Maximum Length (mm)**
  - S4526 *Plesielliptio postbiplicatus*

- **Maximum Height (mm)**
  - 10 mm

- **Shell Convexity (mm)**
  - 10 mm

- **Outline Shape**
  - Orbicular
  - Ovate
  - Ovate Trigonal
  - Ovate Elliptical
  - Elongate Elliptical

- **Anterior Margin Shape**
  - 0 = Slight curvature
  - 1 = Moderate curvature
  - 2 = Strong curvature

- **Tooth Width**
  - S4526 *Plesielliptio postbiplicatus*

- **Tooth Orientation**
  - Orthogonal
  - Oblique Anterior
  - Curved Anterior
  - Oblique Posterior
  - Curved Posterior
Tests the separation of two groups
- Null hypothesis = The groups are the same
- Mahalanobis distance ($D^2$) = The multivariate distance between the means of each group
  - The greater the distance, the more distinct the groups
- Null hypothesis is rejected when $F_{calc} > F_{table}$
- Accomplished using two programs, DISCRIM and PAST
- This study attempts to discriminate closely related sister species as valid
Both programs determined that these species were distinct on the basis of size parameters.

DISCRIM: $F_{18,24} = 15.8^*$

Although PAST resulted in a significant separation based on shape parameters, DISCRIM did not.

DISCRIM: $F_{14,17} = 0.641$

* Significant at $\alpha = 0.05$
**Plesielliptio gibbosoides** and **Plesielliptio whitfieldi**

Both programs determined that these species were distinct on the basis of size parameters.

DISCRIM: $F_{15, 24} = 1447.8^*$

Neither program produced a significant separation based on shape.

DISCRIM: $F_{15, 24} = 0.66$

* Significant at $\alpha = 0.05$
Plesielliptio postbiplicatus and Plesielliptio brachyopisthus

Although DISCRIM showed significant separation, PAST did not classify all specimens into the correct species.

DISCRIM: $F_{15,47} = 275.73^*$

Neither program produced a significant separation based on shape.

DISCRIM: $F_{14,53} = 1.87$

* Significant at $\alpha = 0.05$
Proparreysia letsoni and Proparreysia verrucosiformis

Both programs determined that these species were distinct on the basis of size parameters.

DISCRIM: $F_{16,37} = 78.02^*$

Contrary to the other species sets, both programs produced a significant separation based on shape parameters, which is mostly due to the shape of the posteroventral sculpture.

DISCRIM: $F_{14,34} = 3.17^*$

* Significant at $\alpha = 0.05$
The Deccan Traps

- ~1.5 million km³ volume & ≥ 500,000 km² area (est originally 10 million km³ & 10 million km²)
- 10–60 m thick; Flows thicker to west
- ≤ 15 flows; variable
- Uppermost Cretaceous-Paleocene
- ~6 Ma (debated) of volcanism with largest pulse at ~67 Ma
Paleogeography of India at ~ 65 Ma

During this time India was moving at ~18-19.5 cm/yr

Source of Volcanism – Breakup of India and Seychelles
Current Interpretation

- Mostly lacustrine clay, silt, and carbonate
- Channel deposits rare
- Sediments often converted to chert
- Infra- and Intertrappean beds are often thin and discontinuous
- Sequence rests on Gondwanan units or Precambrian rocks
- Traps can be amygdaloidal or nodular
Previous Studies: James Sowerby (1840) and Stephen Hislop (1860)
Continental and a few brackish infra- and intertrappean molluscan localities, central India.

(geology – GSI, 1998, 1:2,000,000) (base – Gizi Map, 1:3,000,000)
Int 3: Sindhi
Int 3: Butera
Int 2: Kalmeshwar
Int 1: Takli
Inf: Pijdura

Sindhi

from Wilson, 2010
Fossils are mostly steinkerns found as float – Infratrappean is red claystone with some sandstone lenses

from Wilson, 2010
Fossils are steinkerns and original shell; microsnails dominant – Intertrappean is green claystone.
Fossils are mostly steinkerns; microsnails dominant - Intertrappean is light colored siltstone and brown claystone
Fossils are mostly steinkerns; micro- and macrosnails - Intertrappean is light colored siltstone and chert.
Fossils are steinkern cross-sections; microsnails dominant—Intertrappean is black colored chert

from Samant and Mohabey, 2009
Gastropod Character Traits

**Quantitative**

- Basic Shell Parameters
- Number of Whorls
- Aperture Angle

**Qualitative**

- Suture Depression
- Umbilicus Type
- Sculpture

- Slight
- Some
- Regular
- Strong

- Open Umbilicus
- Closed Umbilicus

- Revolving Sculpture
Cluster Analysis

- Groups observations into the most homogeneous and distinct clusters
  - Based on overall similarities in the characteristics of the members
- Displayed as dendrogram
  - Shorter the arms, the greater the similarity
- Correlation coefficient (R) calculated to show strength of results (goodness of fit)
  - Between 0 and 1; R = 0.8 considered very good fit
- Uses hierarchical clustering = most similar observations clustered first
- Accomplished using StatistiXL for Microsoft Excel
- This study attempts to identify morphotypes (species) based on grouping patterns
16 Morphotypes

R = 0.71
**Viviparidae**
- *Bellamya lattooformis* (vivA)
- *Bellamya normalis* (vivB)

**Pomatiopsidae**
- *Tricula virapai* (hydA)
- *Tricula hislopi conoidea* (hydB1)
- *Tricula hislopi takliensis* (hydB2)
- *Tricula sankeyi* (hydC)

**Thiaridae**
- *Thiara quadrilineata* (hydD)
**Valvatidae**
*Valvata multicarinata* (valA)
*Valvata unicarinifera chiknaformis* (valB1)
*Valvata unicarinifera golata* (valB2)

**Subulinidae**
*Paleozootecus burji* (styA)
*Subulina subcylindracea* (styB)
*Subulina pyramis* (styC)
Lymnaeidae
*Lymnaea oviformis* (lymA)
*Lymnaea pokhariensis* (lymB)
*Lymnaea subulata* (lymC)

Planorbidae
*Platyphysa prinsepii elongata* (phyA)
*Platyphysa prinsepii normalis* (phyB)
Analysis of Variance (ANOVA)

- Tests the separation of more than two groups
  - Null hypothesis = The groups are the same
- Based on the variance in variables
  - One continuous variable and one or more categorical variables
  - 2-way and multiway also test the interaction of variables
- Null hypothesis is rejected when $F_{\text{calc}} > F_{\text{table}}$
- Post-hoc Tukey’s test (if significant) determines how groups are related
- Accomplished using R
- This study attempts to differentiate localities based on the characteristics of their gastropod fauna
  - Ratio data = Morphology changes
  - Raw data = Specimen size changes
**Tricula sankeyi**

$F_{4,67} = 3.94, p = 0.0063^*$

$F_{4,67} = 1.06, p = 0.3819$

* Significant at $\alpha = 0.05$

---

**InS0126 Pijdura**

**InS0657 Takli**

**InS0762 Kalmeshwar**

**InS1186 Butera**

**InS1500 Sindhi**

*Scale = 1 mm*
**Lymnaea pokhariensis**

- **F**$_{4,53}$ = 13.8, \( p = <0.0001^* \)

- **F**$_{4,53}$ = 1.06, \( p = 0.3851 \)

* Significant at \( \alpha = 0.05 \)

**Images:**

- InS0182 Pijdura
- InS0261 Takli
- InS0944 Kalmeshwar
- InS1292 Butera
- InS1501b Sindhi

Scale = 1 mm
Platyphysa prinsepii normalis

$X^2_3 = 28.2, \quad p = <0.0001^*$

$F_{4, 48} = 10.4, \quad p = <0.0001^*$

* Significant at $\alpha = 0.05$
Subulina subcylindracea

$F_{3,37} = 2.75, p = 0.0567$

$F_{3,37} = 0.320, p = 0.8111$

* Significant at $\alpha = 0.05$
**Valvata unicarinifera**

F$_{2, 45}$ = 4.49, p = 0.0167*

F$_{2, 45}$ = 0.712, p = 0.4959

* Significant at $\alpha = 0.05$
Bellamya normalis

$\chi^2_{4} = 47.9, \quad p = <0.0001^*$

$F_{4,62} = 9.91, \quad p = <0.0001^*$

* Significant at $\alpha = 0.05$
Chi$^2$ Tests

- Tests the separation of groups based on count data
  - Null hypothesis = The groups are the same
- Number of items of each group compared to the expected
  - The expected is the average based on the total number of items
- Null hypothesis is rejected when $X^2_{\text{calc}} > X^2_{\text{table}}$
- Accomplished using Microsoft Excel
- This study attempts to differentiate localities based on the diversity and abundance of gastropod fauna
## Diversity

<table>
<thead>
<tr>
<th>Location</th>
<th>Number of Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pijdura</td>
<td>9</td>
</tr>
<tr>
<td>Takli</td>
<td>14</td>
</tr>
<tr>
<td>Kalmeshwar</td>
<td>14</td>
</tr>
<tr>
<td>Butera</td>
<td>14</td>
</tr>
<tr>
<td>Sindhi</td>
<td>8</td>
</tr>
</tbody>
</table>

Change in diversity not significant

\[ \chi^2_4 = 5.45 \]

## Abundance

<table>
<thead>
<tr>
<th>Location</th>
<th>Number of Specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pijdura</td>
<td>110</td>
</tr>
<tr>
<td>Takli</td>
<td>192</td>
</tr>
<tr>
<td>Kalmeshwar</td>
<td>225</td>
</tr>
<tr>
<td>Butera</td>
<td>114</td>
</tr>
<tr>
<td>Sindhi</td>
<td>25^</td>
</tr>
</tbody>
</table>

Change in abundance significant

\[ \chi^2_4 = 486.3^* \]
\[ \chi^2_3 = 148.5^* - \text{Sindhi removed} \]

[^ Small Amount of Material

* Significant at \( \alpha = 0.05 \)
Conclusions

Statistics can provide a quantitative backbone to paleontological studies to bolster observational results.

Tests can support taxonomic studies.

- Discriminant analysis differentiated closely related unionoid species of the Hell Creek Fm.
- Cluster analysis identified 16 different morphotypes in the Deccan Trap snails (4 families, 6 genera modified; 1 genus, 5 species, 4 subspecies new).

Tests can identify patterns in diversity and morphologic change.

- ANOVA identified that overall the size of Deccan Trap snails changed but not their morphology.
- The Chi2 test showed a change in snail abundance over the Deccan Trap sequence but no change in diversity.
Acknowledgements

- University of North Dakota
- Dr. Joseph Hartman
- Department of Energy
- Energy and Environmental Research Center
- Dr. Henning Scholz
- Geological Survey of India
- Dr. Dhananjay Mohabey
- Dr. Sunil Bajpai
- Dr. Ritu Sharma
Questions?