

THE STATISTICAL ANALYSIS OF SHAPE DATA

by

Ian Leslie Dryden

Submitted in accordance with the requirements  
for the degree of Doctor of Philosophy.

The University of Leeds  
Department of Statistics.

July 1989

## ABSTRACT

In this thesis we consider some topics connected with a statistical shape analysis of point set data. We first give the exact shape distribution for a finite number of points which are independent isotropic bivariate normally distributed in a plane. Various properties of the distribution are investigated, including an asymptotic large variation distribution and a normal approximation for small variations. Connections with previous work are made, and moments, marginal distributions and invariances are also considered. The shape density for triangles is examined in particular detail.

The exact shape distribution is then used in a likelihood based inference procedure, which can be implemented on a computer. Various estimates are compared in a simulation study and the approximate inverse Fisher information matrix is also given. Likelihood ratio testing for shape change is examined and in particular we describe testing for a uniform shear - the simplest possible shape change. Inference is illustrated with a mouse vertebrae study from anatomy.

As a natural extension, the most general multinormal model in a plane is proposed. The exact shape probability density function under this model is given in a closed form. Although the density is quite complicated, it simplifies considerably in certain cases. Various properties are considered, including a useful normal approximation. Likelihood based inference with this general model is not straightforward, although we shall consider a simple anisotropic model for the mouse vertebrae data.

Some practical issues are discussed and an algorithm for semi-automatic landmark location is proposed. Finally, as a suitable summary for describing a shape change we consider an alternative to the biorthogonal grids for visualizing size and shape change.

#### ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Professor K.V. Mardia for his guidance and encouragement throughout the last three years. I am also grateful to Dr. P. O'Higgins and Dr. D.R. Johnson, of the Department of Anatomy at Leeds, for providing the vertebrae data.

I wish to thank the staff and postgraduates in the School of Mathematics for creating a pleasant working environment, and I would also like to thank my family and Susan for their continuing support and encouragement.

This work was funded by a S.E.R.C. research grant.

---

CONTENTS.

---

	<u>PAGE</u>
Abstract	(ii)
Acknowledgements	(iii)
Contents	(iv)
List of Figures	(viii)
List of Tables	(xii)
Abbreviations	(xiii)
 CHAPTER 1: INTRODUCTION	 1
1.1 GENERAL INTRODUCTION	1
1.2 THE SHAPE SPACE	9
1.3 SHAPE DISTRIBUTIONS	15
 CHAPTER 2: THE EXACT SHAPE DISTRIBUTION FOR INDEPENDENT ISOTROPIC BIVARIATE NORMALLY DISTRIBUTED POINTS IN $\mathbb{R}^2$ .	 18
2.1 INTRODUCTION	18
2.2 THE EXACT SHAPE DISTRIBUTION	19
2.2.1 The shape density	19
2.2.2 The density of complex shape variables	29
2.3 INVARIANCES	31
2.3.1 Baseline choice and label invariance	31
2.3.2 The closure property under rotation	33
2.3.3 Alternative shape variables	34

2.4	PROPERTIES	37
2.4.1	The asymptotic ( $\tau \rightarrow \infty$ ) distribution	37
2.4.2	Approximate normality ( $\tau \rightarrow 0+$ )	41
2.4.3	Moments	43
2.4.4	Particular cases	45
2.4.5	Marginal distributions	51
CHAPTER 3:	INFERENCE USING THE INDEPENDENT ISOTROPIC NORMAL MODEL	58
3.1	INTRODUCTION	58
3.2	BOOKSTEIN'S INFERENCE	59
3.3	MAXIMUM LIKELIHOOD ESTIMATION	60
3.3.1	Exact MLE	60
3.3.2	Normal MLE	62
3.3.3	Fisher information for the normal approximation	64
3.4	SIMULATION STUDIES	67
3.4.1	Estimation under the isotropic model	67
3.4.2	Fisher information and standard errors	78
3.4.3	A note on the parameterization	80
3.5	LIKELIHOOD RATIO TESTS	80
3.5.1	Standard tests	80
3.5.2	Uniform shape change	83
3.6	PRACTICAL APPLICATIONS	86
3.6.1	The mouse vertebrae data	86
3.6.2	The T1 vertebrae	88
3.6.3	The T2 vertebrae	95
3.6.4	Further analysis	100
3.6.5	The neural spine (T2)	105

CHAPTER 4: THE EXACT SHAPE DISTRIBUTION FOR MULTINORMALLY	
DISTRIBUTED POINTS IN A PLANE.	107
4.1 INTRODUCTION	107
4.2 THE EXACT SHAPE DISTRIBUTION	108
4.2.1 The shape density	108
4.2.2 The parameterization for unequal means	113
4.3 PARTICULAR CASES	114
4.3.1 The complex normal case	114
4.3.2 The equicorrelation case	119
4.3.3 The equal means case	119
4.4 PROPERTIES	122
4.4.1 Large variations	122
4.4.2 The normal approximation	123
4.4.3 The triangle case	124
4.5 INFERENCE	128
4.5.1 The anisotropic model	128
4.5.2 The mouse vertebrae using the anisotropic model	130
4.5.3 A problem with the triangle case	133
4.5.4 General inference and other models	138
CHAPTER 5: SOME PRACTICAL ISSUES.	143
5.1 PRE-PROCESSING	143
5.1.1 Outline acquisition	143
5.1.2 Landmark location	147
5.2 VISUALIZING SHAPE CHANGE	153
5.2.1 Shape change descriptions	153
5.2.2 Bookstein's triangle shape change space	155
5.2.3 Biorthogonal and inverse distance grids	156
5.2.4 The mouse vertebrae	165

CHAPTER 6: CONCLUSIONS AND COMMENTS ON FUTURE WORK.

170

REFERENCES.

177



## LIST OF FIGURES.

	<u>PAGE</u>
<u>Fig.1.1.1:</u> Galileo's picture of bones from two different animals.	4
<u>Fig.1.1.2:</u> Two images of chromosomes which have the same shape.	4
<u>Fig.1.1.3:</u> The geometric interpretation of Bookstein's (modified) shape variables.	12
<u>Fig.1.1.4:</u> Part of the shape space for triangles.	13
<u>Fig.2.1:</u> A surface view of the asymptotic ( $\tau \rightarrow \infty$ ) distribution together with a contour plot.	38
<u>Fig.2.2:</u> The exact shape density and a simulation for an example triangle.	46
<u>Fig.2.3:</u> Comparing contour plots of the exact shape density and the normal approximation, for example triangles.	47
<u>Fig.2.4:</u> Comparing contour plots of the exact and the asymptotic large variation shape densities, for example triangles.	49
<u>Fig.2.5:</u> The exact marginal shape density and the normal approximation.	56
<u>Fig.2.6:</u> The exact marginal shape density and the asymptotic large variation distribution.	57
<u>Fig.3.1:</u> Comparison of exact and normal MLE's in a simulation study for triangles.	70
<u>Fig.3.2:</u> Comparison of exact and normal MLE's for a triangle in a more symmetric case.	74



<u>Fig.3.3:</u> Comparison of various estimates in a simulation study for quadrilaterals.	75
<u>Fig.3.4:</u> Comparison of various estimates in a simulation study for hexagons.	76
<u>Fig.3.5:</u> Contour plots showing that the parameterization with $\tau$ is sensible as it reflects a poor choice of baseline.	79
<u>Fig.3.6:</u> D'Arcy-Thompson's (1917) famous example of a uniform shear from one species of fish to another.	84
<u>Fig.3.7:</u> An example of a uniform shape change.	84
<u>Fig.3.8:</u> The digitized outline of a T1 bone from the Large group, pictured in the shape space.	89
<u>Fig.3.9:</u> The outlier T1 from the Small group.	90
<u>Fig.3.10:</u> Marginal scatter plots of the shape variables for the T1's.	91
<u>Fig.3.11:</u> Uniform shear fitting for the T1 bones with 4 landmarks.	93
<u>Fig.3.12:</u> The digitized outline of a T2 bone from the Large group, pictured in the shape space.	96
<u>Fig.3.13:</u> The outlier T2 in the Small group.	97
<u>Fig.3.14:</u> Marginal scatter plots of the shape variables for the T2's.	98
<u>Fig.3.15:</u> Uniform shear fitting for the T2 bones with 6 landmarks.	101

(x)

<u>Fig.3.16:</u> Uniform shear fitting for the T2 bones with 4 landmarks.	103
<u>Fig.3.17:</u> The marginal scatter plot of the shape variables for the neural spine triangle in the Small group of experiment A.	106
<u>Fig.4.1:</u> Contour plots of the shape densities for a triangle of points multinormally distributed in a plane, with $\sigma = 0.1$ .	126
<u>Fig.4.2:</u> Contour plots of the shape densities for a triangle of points multinormally distributed in a plane, with $\sigma = 0.03$ .	127
<u>Fig.4.3:</u> Anisotropic estimation for the T2 bones. The plots show the anisotropic MLE's with approximate 95% confidence regions.	134
<u>Fig.4.4:</u> A scatter plot for Bookstein's shell data.	136
<u>Fig.4.5:</u> The log-likelihood for Bookstein's shell data.	140
<u>Fig.5.1:</u> A video-image of a T1 bone and a T2 bone.	144
<u>Fig.5.2:</u> A binary digitized image (2 grey levels) of a T2 bone.	146
<u>Fig.5.3:</u> The calculation of the $k$ -curvature.	146
<u>Fig.5.4:</u> Example bones from the Large group with the approximate $k$ -curvature functions.	149
<u>Fig.5.5:</u> Example results from the landmark location algorithm. a) 4 landmarks on a T1 bone. b) 6 landmarks on a T2 bone.	150
<u>Fig.5.6:</u> Two examples of ambiguity in using the landmark location algorithm.	152

- Fig.5.7: Pseudo-landmarks provide more information about the outline. 152
- Fig.5.8: Examples of landmark locations. 154
- Fig.5.9: The shape change from a) triangle *A* to b) triangle *B* is represented by c) the scaled principal strains. 157
- Fig.5.10: An explanation through similar triangles of the shape change at landmark *i*. 157
- Fig.5.11: The shape change for males from age 8-14. a) The biorthogonal grid and b) the inverse distance grid. 161
- Fig.5.12: The shape change for females from age 8-14. a) The biorthogonal grid and b) the inverse distance grid. 162
- Fig.5.13: The shape change from a square to a kite, shown by the biorthogonal grid and the inverse distance grids. 164
- Fig.5.14: The inverse distance grids for the T1 vertebrae. a) Control to Large. b) Control to Small. 166
- Fig.5.15: The inverse distance grids for the T2 vertebrae. a) Control to Large. b) Control to Small. 168

## LIST OF TABLES.

	<u>PAGE</u>
<u>Table 3.1:</u> Comparison of standard deviations of estimates from a simulation study, with the inverse of the approximate Fisher information matrix.	78
<u>Table 3.2:</u> Exact maximum likelihood estimates (and approximate standard errors) for the T1 bones.	92
<u>Table 3.3:</u> Exact maximum likelihood estimates (and approximate standard errors) for the T2 bones.	99
<u>Table 4.1:</u> Shape mean square error for estimates under various simulations.	130
<u>Table 4.2:</u> Exact anisotropic maximum likelihood estimates (and approximate standard errors) for the T1 bones.	131
<u>Table 4.3:</u> Exact anisotropic maximum likelihood estimates (and approximate standard errors) for the T2 bones.	132

ABBREVIATIONS

MLE	-	Maximum Likelihood Estimate
MSE	-	Mean Square Error
NAG	-	Numerical Algorithms Group, Oxford.
pdf	-	probability density function
T1	-	First Thoracic Vertebra
T2	-	Second Thoracic Vertebra