

Contents

1	Cinematics and R	1
1.1	Introduction	1
1.2	Cinematics	2
1.2.1	The idea	2
1.2.2	Pattern recognition in cinematics	3
1.2.3	Lies, damned lies and statistics (and fear)	4
1.3	Structure of the notes	5
2	Examples	6
2.1	Preamble	6
2.2	Examples	6
2.2.1	ASLs over time	6
2.2.2	Comparison of ASL distributions	7
2.2.3	Pattern and SL distributions	9
2.2.4	Internal pattern - individual films	11
2.2.5	Internal pattern - an example of global analysis	12
2.2.6	Shot-scale analysis	14
3	Getting R, getting started	22
3.1	Finding R	22
3.2	Data entry	22
3.2.1	General	22
3.2.2	Using the Cinematics database	23
3.3	Packages	23
3.4	Reading	24
4	Descriptive statistics	25
4.1	Introduction	25
4.2	Basics	25
4.3	Functions	26
4.4	Data manipulation	27
4.5	Illustrative graphical analyses	27
5	Graphical analysis – basics	33
5.1	Histograms	33
5.1.1	Basics – an example	33
5.1.2	Technicalities	34
5.1.3	Example continued - log-transformation	35
5.2	Kernel density estimates	36
5.3	Boxplots	39
5.3.1	Basics	39
5.3.2	Interpretation	40

5.3.3	Boxplots and outliers	41
6	Comparative graphical analysis	44
6.1	KDEs and Histograms	44
6.2	Boxplots and violin plots	47
6.3	Cumulative frequency diagrams	48
6.4	Comparison with reference distributions	49
6.4.1	Comparisons with the lognormal distribution	49
6.4.2	Aspects of the normal distribution	49
6.4.3	Normal probability plots	50
6.4.4	Using KDEs for SL comparisons	50
6.4.5	Examples	52
7	Time series analysis of SLs	56
7.1	Introduction	56
7.2	Polynomial models (trendlines)	57
7.3	Simple moving averages	60
7.4	Smoothed Moving averages	61
7.5	Loess smoothing	62
7.5.1	Examples	62
7.5.2	Some technicalities	64
7.6	Robust methods	65
7.6.1	Introduction	65
7.6.2	Moving medians	66
7.6.3	Moving average of ranked SLs	67
7.6.4	Order structure matrices	69
7.7	Kernel density estimation	71
7.8	Summary	72
7.8.1	General similarities	72
7.8.2	Specific differences between methods	73
7.8.3	Additional methodological comments	74
7.8.4	Discussion and conclusion	75
8	Enhanced plots	77
8.1	<i>Lights of New York</i> (1928) explored	77
8.1.1	Colour-coding shots	77
8.1.2	Adding ‘wallpaper’	79
8.1.3	Order structure matrices revisited	80
8.2	<i>Intolerance</i> revisited	82
8.3	Jeu d’esprit	83
8.3.1	<i>Halloween</i> (1928) – matching colour to content	83
8.3.2	Controlling for shot numbers - an idea	85
8.3.3	Code	87
8.3.4	Coda	88
9	Shot-scale data analysis	92
9.1	Correspondence analysis – initial example	92
9.2	What correspondence analysis does	95
9.3	Producing enhanced CA plots	96
9.4	CA and supplementary points	100
9.5	Other plotting methods	102
9.5.1	Analysis of ranks	102
9.5.2	Analysis of individual shot-scale types	104

A	The Lognormal distribution	109
A.1	Introduction	109
A.2	The lognormal and normal distributions	110
A.2.1	Definitions	110
A.2.2	Parameter interpretation	110
A.2.3	Models and estimates	111
A.3	Why logarithms?	113
A.4	Other normalising transformations	114
A.4.1	Box-Cox (BC) transformations	115
A.4.2	Yeo-Johnson (YJ) transformations	115