

Contents

1	Introduction	1
1.1	Some Motivating Applications	1
1.2	Weighted Correlation and Applications	2
1.3	Organization of the Book	6
2	The Weighted Rank Correlation Coefficient r_W	9
2.1	Introduction	9
2.2	Rank Correlation	10
2.3	Weighted Rank Measure of Correlation	11
2.4	Properties of the Distribution of r_W Under the Null Hypothesis of Independence	13
2.4.1	Linear Rank Statistics	13
2.4.2	Exact and Asymptotic Distribution of r_W Under the Null Hypothesis of Independence	14
2.4.3	Simulations	15
2.4.4	Comparison Between r_W and r_S	17
2.5	The Asymptotic Distribution of r_W for the General Case	20
2.5.1	$\sum_{i=1}^3 A_{in}$ is Asymptotically Normal Distributed	22
2.5.2	B_{1n}^* is Asymptotically Negligible	24
2.5.3	r_W is Asymptotically Normal Distributed	24
2.6	Examples of Application of r_W	24
2.7	Conclusions About r_W	27
3	The Weighted Rank Correlation Coefficient r_{W2}	29
3.1	Introduction	29
3.2	The Formula of the Coefficient r_{W2} in the Case of Ties	33
3.3	Comparison Between the Three Coefficients r_{W2} , r_W and r_S	36
3.4	A New Way of Developing Weighted Correlation Coefficients	38

4	A Weighted Principal Component Analysis, WPCA1; Application to Gene Expression Data	39
4.1	Introduction.	39
4.2	A New Weighted Version of PCA	41
4.3	Preliminary Version of Weighted Principal Component Analysis Using r_w	42
4.4	Weighted Principal Component Analysis, WPCA1, Using r_{w2}	44
4.4.1	Robustness to Outliers and Noise.	46
4.5	A New Method for Selecting Relevant Genes in Microarray Data.	48
4.5.1	Support Vector Machines Classification: Results for Genes Chosen with SAM, PDeig, and with Our Method for WPCA and PCA	50
4.5.2	Analysis of the Chosen Genes.	51
4.6	Conclusions.	52
5	A Weighted Principal Component Analysis (WPCA2) for Time Series Data	55
5.1	Introduction.	55
5.2	Motivation and Definition.	56
5.3	Weight Functions: Examples	59
5.4	Applications	60
5.5	Final Remarks.	67
6	Weighted Clustering of Time Series	69
6.1	Introduction.	69
6.2	Applications	72
	Appendix	75
	References.	87