

THEORY OF RANK TESTS

**Jaroslav Hajek
Zbynek Sidak
Pranab K. Sen**

(Second Edition-)



ACADEMIC PRESS

San Diego London Boston New York
Sydney Tokyo Toronto

Contents

Preface to the Second Edition	v
Preface to the First Edition	vii
Introduction and coverage	1
1.1 THE BACKGROUND.	1
1.2 ORGANIZATION OF THE PRESENT TREATISE	5
Preliminaries	10
2.1 BASIC NOTATION.	10
2.1.1 Probability space and observations.	10
2.1.2 Statistics, σ -fields and \mathcal{A} -fields.	11
2.1.3 Martingales and related sequences.	13
2.2 FAMILIES OF ONE-DIMENSIONAL DENSITIES.	13
2.2.1 Absolutely continuous densities.	14
2.2.2 Strongly unimodal densities.	14
2.2.3 Densities with finite Fisher information.	16
2.2.4 The η -functions.	17
2.2.5 Types of densities and the measurement of their distance.	21
2.3 TESTING HYPOTHESES.	22
2.3.1 Statement of the problem.	22
2.3.2 The envelope power function.	25
2.3.3 Most powerful test	26
2.3.4 The Neyman-Pearson lemma	27
2.3.5 Least favourable densities.	28
2.4 AUXILIARY RESULTS FOR NORMAL SAMPLES.	29
2.4.1 χ^2 -distribution for correlated normal random variables.	29
2.4.2 Most powerful tests.	30
Problems and complements to Chapter 2.	31

Elementary theory of rank tests	35
3.1 RANKS AND ORDER STATISTICS.	35
3.1.1 Ranks and order statistics.	35
3.1.2 Hypotheses H_0 and H_1 (randomness).	37
3.1.3 Hypothesis H_1 (symmetry).	39
3.1.4 Hypothesis H^* (independence).	40
3.1.5 Hypothesis H_3 (random blocks).	41
3.2 PERMUTATION, INVARIANT, AND RANK TESTS.	41
3.2.1 Permutation tests.	41
3.2.2 Tests invariant under changes of location and scale.	45
3.2.3 Rank tests.	52
3.2.4 Lehmann's alternatives.	54
3.3 EXPECTATIONS AND VARIANCES OF LINEAR RANK STATISTICS	57
3.3.1 Linear rank statistics for H_0 .	57
3.3.2 Linear rank statistics for H_1 , H_2 and H_3 . Antiranks.	62
3.4 LOCALLY MOST POWERFUL RANK TESTS.	64
3.4.1 Definition of locally most powerful tests.	64
3.4.2 A convergence theorem.	64
3.4.3 Scores.	65
3.4.4 H_0 against two samples differing in location.	67
3.4.5 H_0 against two samples differing in scale.	69
3.4.6 H_0 against regression in location.	69
3.4.7 H_0 against regression in scale.	70
3.4.8 H_0 against a general alternative.	70
3.4.9 H_1 against the location shift.	74
3.4.10 H_1 against two samples differing in scale.	75
3.4.11 H_2 against dependence.	76
3.5 STATISTICAL FUNCTIONALS.	79
3.5.1 L -statistics and U -statistics.	80
3.5.2 Differentiable statistical functionals.	83
3.5.3 Generalized statistical functionals.	84
Problems and complements to Chapter 3.	85
Selected rank tests	94
4.1 TWO-SAMPLE TESTS OF LOCATION	95
4.1.1 Linear rank tests.	95
4.1.2 Tests based on exceeding observations.	98
4.1.3 Tests of Kolmogorov-Smirnov types.	99
4.2 TWO-SAMPLE TESTS OF SCALE	103
4.2.1 Linear rank tests.	103
4.2.2 Tests based on exceeding observations.	107
4.2.3 Tests of Kolmogorov-Smirnov types.	108
4.3 REGRESSION.	108
4.3.1 Linear rank tests.	109
4.3.2 The E -test.	110
4.3.3 Tests of Kolmogorov-Smirnov types.	111
4.4 THREE OR MORE SAMPLES.	112

4.4.1	Rank tests of χ^2 -types	112
4.4.2	Other rank tests and some general ideas	114
4.5	TESTS OF SYMMETRY	117
4.5.1	Linear rank tests	117
4.5.2	Kolmogorov-type tests	119
4.6	TESTS OF INDEPENDENCE	121
4.6.1	Linear rank tests	121
4.6.2	Other rank tests	126
4.6.3	Kolmogorov-type tests	126
4.7	RANDOM BLOCKS	128
4.7.1	Rank tests of χ^2 -types	128
4.8	TREATMENT OF TIES	130
4.8.1	Randomization, average statistics and scores, mid-ranks	131
4.8.2	Other methods of dealing with ties	135
4.9	RANK TESTS FOR CENSORED DATA	136
4.9.1	Type I censoring	136
4.9.2	Type II censoring	139
4.9.3	Progressive censoring	141
4.9.4	Interval censoring	144
4.9.5	Random censoring	148
4.10	MULTIVARIATE RANK TESTS	152
4.10.1	Bivariate sign tests	152
4.10.2	Multivariate signed-rank tests	153
4.10.3	Multivariate linear rank statistics	155
4.10.4	LMPR property for multivariate tests	159
	Problems and complements to Chapter 4	160

5 Computation of null exact distributions 165

5.1	DIRECT USE OF DISTRIBUTION OF RANKS	165
5.1.1	Tests for regression, independence and random blocks	165
5.1.2	Two-sample tests	167
5.1.3	Tests for three or more samples	167
5.1.4	Tests of symmetry	168
5.2	EXPLICIT FORMULAS FOR DISTRIBUTIONS	168
5.2.1	Statistics using the scores 0, \sim , 1	168
5.2.2	The Haga test and the E-test statistics	170
5.2.3	Kolmogorov-Smirnov statistics for two samples of equal sizes	170
5.3	RECURRENCE FORMULAS FOR DISTRIBUTIONS	173
5.3.1	The Wilcoxon and the Kendall test statistics	173
5.3.2	Statistics of Kolmogorov-Smirnov types for two samples of unequal sizes	174
5.3.3	The Kruskal-Wallis and the Friedman test statistics	176
5.4	IMPROVEMENTS OF LIMITING DISTRIBUTIONS	177
5.4.1	Corrections for continuity	177
5.4.2	Interpolation for numerical computations	178
5.4.3	Transformations of test statistics and the use of other limiting distributions	179
	Problems and complements to Chapter 5	179

Limiting null distributions	183
6.1 SIMPLE LINEAR RANK STATISTICS.	183
6.1.1 Convergence in indexed sets of statistics.	183
6.1.2 A special central limit theorem.	184
6.1.3 A convergence theorem.	186
6.1.4 Further preliminaries.	186
6.1.5 Locally optimum rank-test statistics for H_0 .	190
6.1.6 General simple linear rank statistics for H_0 .	194
6.1.7 Rank statistics for H_1 .	197
6.1.8 Simple linear rank statistics for H_2 .	199
6.1.9 Martingale characterizations.	200
6.2 RANK STATISTICS OF χ^2 -TYPES.	203
6.2.1 Convergence in distribution for random vectors.	203
6.2.2 Statistics-for testing H_0 against k samples differing in location or scale.	205
6.2.3 Statistics for H_0 against differences within blocks.	208
6.3 STATISTICS OF KOLMOGOROV-SMIRNOV TYPES.	209
6.3.1 Probability distributions in the space of continuous functions.	209
6.3.2 Convergence in distribution in $C[0,1]$.	215
6.3.3 Weak convergence in the Skorokhod metric.	217
6.3.4 The Brownian bridge.	219
6.3.5 Kolmogorov's inequality for dependent summands.	222
6.3.6 The basic convergence theorem.	224
6.3.7 The Kolmogorov-Smirnov statistics.	227
6.3.8 The Rerryi statistics.	228
6.3.9 The Cramer-von Mises statistics.	230
6.4 FUNCTIONAL CENTRAL LIMIT THEOREMS.	230
6.4.1 Weak invariance-principles for martingales.	231
6.4.2 Almost sure invariance principles for martingales.	232
6.4.3 Invariance principles for empirical processes.	233
6.4.4 Invariance principles for rank statistics.	234
Problems and complements to Chapter 6	240
Limiting non-null distributions	249
7.1 CONTIGUITY.	249
7.1.1 'Asymptotic methods.- Contiguity.	249
7.1.2 LeCam's first lemma.	251
7.1.3 LeCam's second lemma.	253
7.1.4 LeCam's third lemma.	257
7.1.5 LAQ family and the convolution theorem.	260
7.1.6 Contiguity and Hellinger distance.	261
7.2 SIMPLE LINEAR RANK STATISTICS.	262
7.2.1 Location alternatives for H_0 .	262
7.2.2 Scale alternatives for H_0 .	265
7.2.3 Rank statistics for F_{l0} -against two samples.	267
7.2.4 Rank statistics for H_0 against regression.	269
7.2.5 Rank statistics for H_1 .	274
7.2.6 Rank statistics for H_2 .	275

7.3	FAMILIES OF SIMPLE LINEAR RANK STATISTICS	276'
7.3.1	Finite families (vectors).	276
7.3.2	Continuous families (processes).	279
7.4	ASYMPTOTIC POWER.	281
7.4.1	One-sided S-tests.	281
7.4.2	Two-sided S-tests.	282
7.4.3	Q-tests.	283
7.4.4	Tests of Kolmogorov-Smirnov types.	283
7.4.5	Local behaviour of the asymptotic power (one-sided tests).	284
7.4.6	Local behaviour of asymptotic power (multi-sided tests).	287
7.4.7	Progressive censoring.	288
7.5	NON-CONTIGUOUS ALTERNATIVES.	289
7.5.1	A historical note.	289
7.5.2	Variance inequality for linear rank statistics.	290
7.5.3	The Hajek projection approximation.	291
7.5.4	Asymptotic normality of linear rank statistics.	292
7.5.5	Integration with contiguity based approaches.	293
	Problems and complements to Chapter 7.	294
8	Asymptotic optimality and efficiency	299
8.1	ASYMPTOTICALLY OPTIMUM TESTS.	299
8.1.1	Introduction.	299
8.1.2	Asymptotic sufficiency of the vector of ranks.	302
8.1.3	Asymptotically optimum one-sided tests.	307
8.1.4	Asymptotically optimum multi-sided tests.	310
8.2	ASYMPTOTIC EFFICIENCY OF TESTS.	316
8.2.1	One-sided tests.	316
8.2.2	Multi-sided tests.	320
8.2.3	Local asymptotic efficiency.	321
	Bibliographical notes.	323
8.3	BAHADUR EFFICIENCY.	325
8.3.1	Exact Bahadur efficiency.	325
8.3.2	Kullback-Leibler information.	330
8.3.3	Best exact slope of two-sample rank statistics.	332
8.3.4	Approximate Bahadur efficiency.	336
8.3.5	Relation between Bahadur efficiency and Pitman efficiency.	336
8.4	HODGES-LEHMANN DEFICIENCY	336
8.4.1	Asymptotic expansions for distributions of rank statistics.	337
8.4.2	Hodges-Lehmann's concept of deficiency.	340
8.5	ADAPTIVE RANK TESTS.	341
8.5.1	Restrictive adaptive tests.	342
8.5.2	Non-restrictive adaptive tests.	345
	Problems and complements to Chapter 8.	352
9	Rank estimates and asymptotic linearity	358
9.1	$\hat{\Lambda}$ -ESTIMATES OF LOCATION AND REGRESSION.	358
9.1.1	$\hat{\Lambda}$ -estimates of shift between two samples.	358
9.1.2	$\hat{\Lambda}$ -estimates of shift of one sample.	361

9.1.3	$\hat{\beta}$ -estimates of regression parameters.	362
9.2	ASYMPTOTIC LINEARITY OF RANK STATISTICS IN REGRESSION PARAMETERS.	366
9.2.1	Rank statistics, one-dimensional regression.	366
9.2.2	Rank statistics, multidimensional regression.	377
9.2.3	Signed-rank statistics.	378
9.3	RANK ESTIMATION OF REGRESSION PARAMETERS	378
9.3.1	Basic methods.	378
9.3.2	Linearized rank estimates.	381
	Problems and complements to Chapter 9.	382
10	Miscellaneous topics in regression rank tests	383
10.1	ALIGNED RANK TESTS.	383
10.1.1	Aligned rank tests for two-way layouts.	384
10.1.2	Aligned rank statistics for subhypothesis testing.	389
10.2	REGRESSION RANK [SCORES.	394
10.2.1	The duality of RQ and RRS.	395
10.2.2	RRS estimates and connection to location model.	397
10.3	RANK VERSUS OTHER ROBUST PROCEDURES.	399
	Problems and complements to Chapter 10.	404
	Bibliography	407
	Subject index	425
	Author index	429
	Index of mathematical symbols	433
	Titles in this series	437