

1996

- Magder L**, Zeger S. A smooth nonparametric estimate of a mixing distribution using mixtures of Gaussians. *Journal of the American Statistical Association*. 1996;91:1141-51.
- Qu Y, **Tan M**, Kutner MH. Random-effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics* 52:797–810, 1996.
- Tan M**, Xiong X. Continuous and group sequential conditional probability ratio tests for phase II clinical trials. *Stat Med* 1996;15(19):2037–2051.
- Tan M**, Xiong X. Stopping a clinical trial very early based on sequential conditional probability ratio test. *Controlled Clinical Trials* 16:94–95, 1996.

1997

- Magder L**. Logistic Regression with the outcome is measured with uncertainty. *American Journal of Epidemiology*. 1997;146(2):195-203.

1998

- Qu Y, **Tan M**. Analysis of clustered ordinal data with sub-clusters via a Bayesian hierarchical model. *Communications in Statistics A: Theory & Method* 27:1461–1475, 1998.
- Tan M**, Xiong X, Kutner MH. Clinical trial designs based on sequential conditional probability ratio tests and reverse stochastic curtailing. *Biometrics* 54 (2): 682-95, 1998.

1999

- Tan M**, Qu Y, Mascha E, Schubert A. A Bayesian hierarchical model for multi-level repeated ordinal data: analysis of oral practice examinations in a large anaesthesiology training programme. *Stat Med* 18(15):1983–1992, 1999.
- Tan M**, Qu Y, Rao JS. Robustness of the latent variable model for correlated binary data. *Biometrics* 55(1):258–263, 1999.
- Tan M**, Xiong X. Implications of using the group sequential conditional probability ratio test in the beta-blocker heart attack trial. *Controlled Clinical Trials* 20:35, 1999.

2000

- Liang JJ, Li RZ, **Fang HB**, Fang KT. Testing multinormality based on low-dimensional projection. *Journal of Statistical Planning and Inference* 86:129–141, 2000.
- Liu A, **Tan M**, Boyett JM, Xiong X. Testing secondary hypotheses following sequential clinical trials. *Biometrics* 56(2):640–4, 2000.
- Qu Y, **Tan M**, Rybicki L. A unified approach to estimating association measures via a joint generalized linear model for paired binary data. *Communications in Statistics: Theory & Method* 29:143–156, 2000.
- Song PX, **Tan M**. Marginal models for longitudinal continuous proportional data. *Biometrics* 56(2):496–502, 2000.

2001

- Fang HB**, Sun J. Consistency of nonparametric maximum likelihood estimation of a distribution based on doubly interval-censored failure time data. *Statistics & Probability Letters* 53(3):311–318, 2001.

2002

- Fang HB**, Fang KT, Kotz S. The meta-elliptical distributions with given marginals. *Journal Multivariate Analysis* 2002;82:1–16.
- Fang HB**, Sun J, Lee ML. Nonparametric survival comparisons for interval-censored continuous data. *Statistica Sinica* 2002;12:1073–1083.
- Kotz S, **Fang HB**, Wei G. Multiple square tray distributions. *Statistics & Probability Letters* 2002;59:195–207.
- Liu Z**, Xu Y. Adaptive support vector classification. In: Abraham A, Ruiz-Del-Solar J, Koppen M, eds. *Soft computing systems: Design, management and applications*. Burke, VA: IOS Press, pp. 331–339, 2002.
- Qiu Z, Song PX, **Tan M**. Bayesian hierarchical models for multi-level repeated ordinal data using WinBUGS. *J Biopharm Stat* 2002;12(2):121–135.
- Tan M**, **Fang HB**, **Tian GL**, Houghton PJ. Experimental design and sample size determination for testing synergism in drug combination studies based on uniform measures. *Stat Med* 2003;22(13):2091–2100.

- Tan M, Fang HB, Tian GL**, Houghton PJ. Small-sample inference for incomplete longitudinal data with truncation and censoring in tumor xenograft models. *Biometrics* 2002;58(3):612–620.
- Xiong X, **Tan M**, Kutner MH. Computational methods for evaluating sequential tests and post-test estimation via the sufficiency principle. *Statistica Sinica* 2002;12(4):1027–1041.

2003

- Liu Z**, Bozdogan H. Improving the performance of radial basis function classification using information criteria. In: bozdogan H, ed. *Statistical Data Mining and Knowledge Discovery*. Danvers: CRC, pp. 193–212, 2003.
- Liu Z**, Bozdogan H. Kernel PCA for feature extraction with information complexity. In: Bozdogan H, ed. *Statistical Data Mining and Knowledge Discovery*. Danvers: CRC, pp. 309–322, 2003.
- Liu Z**, Bozdogan H. Radial basis function networks for classification using new kernel functions. *International Journal of Neural Parallel and Scientific Computations* 2003;11:43–54.
- Sun J, **Fang HB**. A nonparametric test for panel count data. *Biometrika* 2003;90:199–208.
- Tan M, Fang HB, Tian GL**, Houghton PJ. Experimental design and sample size determination for testing synergism in drug combination studies based on uniform measures. *Stat Med* 2003;22(13):2091–2100.
- Tan M, Tian GL, Fang HB**. Estimating restricted normal means using the EM-type algorithms and IBF sampling. In: *Development of Modern Statistics and Related Topics*, New Jersey, 2003.
- Tan M, Tian GL**, Ng KW. A non-iterative sampling method for computing posteriors in the structure of EM-type algorithms. *Statistica Sinica* 2003;13(3):625–639.
- Tian GL**, Ng KW, Geng Z. Bayesian computation for contingency tables with incomplete cell-counts. *Statistica Sinica* 2003;13(1):189–206.
- Tian GL, Tan M**. Exact statistical solutions using the inverse Bayes formulae. *Statistics & Probability Letters* 2003;62(3):305–315.
- Xiong X, **Tan M**, Boyett J. Sequential conditional probability ratio tests for normalized test statistic on information time. *Biometrics* 2003;59(3):624–631.
- Xiong X, Yao W, **Tan M**, Yue G. A stochastic model for assessing synchronization of spike trains in a population of neuron motor units. *Methodology and Computing in Applied Probability* 2003;5:355–367.

2004

- Chen D, **Liu Z**, Ma X, Hua D. Selecting genes by test statistics. *Journal of Biomedicine and Biotechnology* 2004;2:132–138.
- Chen Z, Zhang Z, **Liu Z**, Cheng X. An ensemble method of discovering sample classes using gene expression profiling. In: Pardalos P, Principe J, eds. *Data Mining in Biomedicine XVIII*. Springer, pp. 39–46, 2004.
- Fang HB**, Li G, Sun J. Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. *Scandinavian Journal of Statistics* 2005;32:59–75.
- Liu Z**, Chen D, Bensmail H. Gene expression data classification with kernel principal component analysis. *Journal of Biomedicine and Biotechnology* 2004;2:155–159.
- Liu Z**, Chen D, Densmail H, Xu Y. Clustering gene expression with kernel principal component analysis. *Journal of Bioinformatics Research and Applications* 2004;1(2):169–180.
- Liu Z**, Chen D, Xu Y, Liu J. Logistic support vector machine with applications to microarray data. *International Journal of Bioinformatics Research and Applications* 2004;1(2):169–180.
- Liu Z**. CpG island identification with higher order and variable order Markov Models. In: Pardalos P, Principe J, eds. *Data Mining in Biomedicine XVIII*. Springer, pp. 47–58, 2004.
- Song P, Qiu Z, **Tan M**. Modeling heterogeneous dispersion in marginal simplex models for continuous longitudinal proportional data. *Biometrical Journal* 2004;46:540–553.

2005

- Fang HB**, Kotz S, Wei G. Dependence patterns of random variables: geometric properties of copulas. In: Fan J, Li G, eds. *Contemporary multivariate analysis and experimental designs*. Singapore: World Scientific Publishing Co, Inc., pp. 233–252, 2005.
- Fang HB**, Li G, Sun J. Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. *Scandinavian Journal of Statistics* 32:59–75, 2005.
- Fang HB, Tian GL**, Xiong X, **Tan M**. A multivariate random-effects model with restricted parameters: application to assessing radiation therapy for brain tumours. *Stat Med* 2005;25(11):1948–1959.
- Lin S, Ding J, Dong C, **Liu Z**, Ma Z, Wan S, Xu Y. Comparisons of methods for linkage analysis and haplotype reconstruction using extended pedigree data. *BMC Genetics* 2005;6(Suppl 1):S76.

- Liu Z**, Chen D, Tian J. Cancer classification with partial least square algorithm and proteomic data. *Applied Signal Processing: Advanced Signal/Image Processing techniques for Bioinformatics* 2005;special edition:
- Liu Z**, Chen D, Tian J. Chromosome classification with entropy kernel and PLS algorithm, ICIC 2005, Part I, *Lecture Notes in Computer Science (LNCS)* Vol. 3644. Berlin Heidelberg: Springer-Verlag, pp. 543–551, 2005.
- Liu Z**, Lin S. Multiplies LD measure and tagging SNP selection with generalized mutual information. *Journal of Genetic Epidemiology* 2005;19(4):353–364.
- Tan M, Fang HB, Tian GL**, Houghton PJ. Repeated-measures models with constrained parameters for incomplete data in tumour xenograft experiments. *Stat Med* 2005;24(1):109–119.
- Tan M, Fang HB, Tian GL**. Statistical analysis for tumor xenograft experiments in drug development In: Fan J, Li GG, eds. *Contemporary multivariate analysis and experimental designs*. Singapore: The World Scientific Publisher Co. Inc., pp. 319–335, 2005.
- Tan M, Fang HB, Tian GL**, Wei G. Testing multivariate normality test for incomplete data of small sample size. *Journal of Multivariate Analysis* 2005;93(1):164–179.
- Tan M, Tian GL**, Ng KW. Hierarchical models for repeated binary data using the ibf sampler. *Computational Statistics and Data Analysis* 2005;50(5):1272–1286.
- Wan Y**, Chen T. A hidden Markov model based scoring function for mass spectrometry database search. *Proc of Research in Computational Molecular Biology*: 432–437, 2005.
- Wang J, **Fang HB**. Moments of the Kaplan-Meier estimator under noninformative censoring. *Far East Journal of Theoretical Statistics* 2005;16:337–350.

2006

- Chen H**, Cheng C-S. Doubling and Projection: A Method of Constructing Two-Level Designs of Resolution IV. *Annals of Statistics*. 2006;34:546–58.
- Fang H-B, Tian G**, Xiong X, **Tan M**. A multivariate random-effects model with restricted parameters: application to assessing radiation therapy for brain tumours. *Stat Med* 25(11):1948–1959 2006.
- Ng KW, Tang M-L, **Tan M, Tian G-L**. Grouped Dirichlet distribution: A new tool for incomplete categorical data analysis. *Journal of Multivariate Analysis*. 2008;99(3):490–509.
- Liu Z**, Lin S, **Tan M**. Genome-Wide Tagging SNPs with Entropy-Based Monte Carlo Method. *J Comput Biol*. 2006;13(9):1606–14.

2007

- Liu Z**, Jiang F, **Tian G-L**, Wang S, Sato F, Meltzer SJ, **Tan M**. Sparse Logistic Regression with Lp Penalty for Biomarker Identification. *Statistical Applications in Genetics and Molecular Biology*. 2007;6(1):Article 6.
- Shardell M**, Scharfstein DO, Bozzette SA. Survival curve estimation for informatively coarsened discrete event-time data. *Statistics in Medicine*. 2007;26(10):2184–202.
- Tan M, Fang HB**. Advanced statistical methods for the design and analysis of tumor xenograft experiments. In *Cancer models with applications to cancer screening, cancer treatment and risk assessment*. WY Tan, A Yakovlev, eds. World Scientific Publishing Co., New Jersey, 2007.
- Tan M, Tian GL, Fang HB**. An efficient MCEM algorithm for fitting generalized linear mixed models for correlated binary data. *Journal of Statistical Computation and Simulation*. 2007;77(11):929 - 43.
- Tan M, Tian GL, Fang HB**, Ng, KW. A fast EM algorithm for quadratic optimization subject to convex constraints. *Statistica Sinica* 17:945–964, 2007.
- Tian J, **Liu Z**. Coalescent random walks on the graphs. *Journal of Computational and Applied Math* 202 (1): 144–154, 2007.
- Xiong X, **Tan M**, Boyett J. A sequential procedure for monitoring clinical trials against historical controls. *Stat Med* 2007;26(7):1497–1511.

2008

- Liu Z, Tan M**. ROC-based utility function maximization for feature selection and classification with applications to high-dimensional protease data. *Biometrics*. 2008;64(4):1155–61.
- Liu Z**, Chen D, Tian J. Cancer classification with partial least square algorithm and proteomic data. *Journal of Bioinformatics Research and Applications*. 2008;4(1):1–10.
- Liu Z**, Chen D, Tian JP. Classification of proteomic data with multiclass Logistic Partial Least Squares algorithm. *Int J Bioinform Res Appl*. 2008;4(1):1–10.

- Liu Z**, Guo Z, **Tan M**. Constructing Tumor Progression Pathways and Biomarker Discovery with DNA Methylation Data. *Cancer Informatics*. 2008(1):1-8.
- Liu Z**, Guo Z, **Tan M**. Constructing tumor progression pathways and biomarker discovery with fuzzy kernel kmeans and DNA methylation data. *Cancer Inform*. 2008;6:1-7. PMID: PMC2623296.
- Liu Z**, Gartenhaus RB, **Tan M**, Jiang F, Jiao X. Gene and pathway identification with Lp penalized Bayesian logistic regression. *BMC Bioinformatics*. 2008;9(1):412. PMID: 2567351.
- Liu Z**, Jiang F. Gene Identification and survival prediction with supervised and unsupervised learning. *International Journal of Data Mining and Bioinformatics*. 2008(to appear).
- Shardell M**, Scharfstein DO, Vlahov D, Galai N. Inference for cumulative incidence functions with informatively coarsened discrete event-time data. *Stat Med*. 2008;27(28):5861-79. PMID: PMC2796438.
- Shardell M**, Miller RR. Weighted estimating equations for longitudinal studies with death and non-monotone missing time-dependent covariates and outcomes. *Stat Med*. 2008;27(7):1008-25. PMID: PMC2792882.
- Shardell M**, Scharfstein DO, Vlahov D, Galai N. Sensitivity Analysis Using Elicited Expert Information for Inference With Coarsened Data: Illustration of Censored Discrete Event Times in the AIDS Link to Intravenous Experience (ALIVE) Study. *Am J Epidemiol*. 2008;168(12):1460-9.
- Tan M**, **Fang H-B**, **Tian G-L**. Dose and Sample Size Determination for Multi-Drug Combination Studies. *Statistics in Biopharmaceutical Research*. 2009;1(3):301-16.
- Tan M**. Conditional Power in Clinical Trial Monitoring. In: D'Agostino R, editor. Wiley Encyclopedia of Clinical Trials Wiley-Interscience 2008.
- Tian GL**, Tang ML, **Fang HB**, **Tan M**. Efficient methods for estimating constrained parameters with applications to lasso logistic regression. *Comput Stat Data Anal*. 2008;52(7):3528-42. PMID: 2352165.
- Tian G-L**, Ng KW, **Tan M**. EM-type algorithms for computing restricted MLEs in multivariate normal distributions and multivariate t-distributions. *Computational Statistics & Data Analysis*. 2008(52):4768-78.
- Ng KW, **Tian G-L**, Tang ML, **Tan M**. Grouped Dirichlet distribution: A new tool for incomplete categorical data analysis. *Journal Multivariate Analysis*. 2008(99):490-509.

2009

- Chen H**, Cheng C-S. Some results on 2^{n-m} designs of resolution IV with (weak) minimum aberration. *Annals of Statistics*. 2009; 37(6A):3600-15.
- Deng D, **Fang HB**. On nonparametric maximum likelihood estimations of multivariate distribution function based on interval-censored data. *Communications in Statistics A: Theory & Method*. 2009(38):54-74.
- Fang H-B**, **Tian G-L**, Li W, **Tan M**. Design and Sample Size for Evaluating Combinations of Drugs of Linear and Loglinear Dose-Response Curves. *Journal of Biopharmaceutical Statistics*. 2009;19(4):625 - 40. PMID: In process.
- Liu Z**, **Tan M**, Jiang F. Regularized F-Measure Maximization for Feature Selection and Classification. Journal of Biomedicine and Biotechnology [serial on the Internet]. 2009 [cited 2009 Apr 27]; JBB. PMID: 19421401.
- Liu Z**, **Tan M**, Jiang F. Regularized F-Measure Maximization for Feature Selection and Classification. Journal of Biomedicine and Biotechnology [serial on the Internet]. 2009 [cited 2009 Apr 27]; JBB: Available from: <http://www.hindawi.com/journals/jbb/2009/617946.html>.
- Liu Z**, Chen X, Gartenhaus RB, Howell CD, **Tan M**. Survival Prediction and Gene Identification with Penalized Global AUC Maximization. *Journal of Computational Biology*. 2009;16(12):1661-70. PMID: In Process.
- Shardell M**, El-Kamary SS. Sensitivity Analysis of Informatively Coarsened Data Using Pattern Mixture Models. *Journal of Biopharmaceutical Statistics*. 2009; 19(6):1018 - 38.
- Tao GL, Tang ML, **Liu Z**, **Tan M**, Tang NS. Sample size determination for the non-randomised triangular model for sensitive questions in a survey. *Stat Methods Med Res*. 2009;Feb 16. [Epub ahead of print].
- Tan M**, **Fang H-B**, **Tian G-L**. Dose and Sample Size Determination for Multi-Drug Combination Studies. *Statistics in Biopharmaceutical Research*. 2009;1(3):301-16.
- Tan M**, **Tian G-L**, Tang M-L. Sample Surveys with Sensitive Questions: A Non-randomized Response Approach. *The American Statistician*. 2009;63(1):9-16.
- Tan M**, **Tian G-L**, Ng KW. Bayesian Missing Data Problems: EM, Data Augmentation and Non-iterative Computation. Jones B, Liu J-P, Peace KE, Turnbull BW, editors. Boca Raton: Chapman and Hall CRC; 2009.
- Tan M**, **Fang HB**, **Tian GL**. Dose and sample size determination for multi-drug combination studies. *Statistics in Biopharmaceutical Research*. 2009;1(3):301-16.

- Tian GL, Fang HB, Liu Z, Tan M.** Regularized (bridge) logistic regression for variable selection based on ROC criterion. *Statistics and its Interface*. 2009; 2:493-502.
- Tian G-L, Tang M, Yuen K, Tan M.** Bayesian non-randomized response models for surveys with sensitive questions. *Statistics and Its Interface*. 2009;2(1):13-25.
- Tian G-L, Tan M, Tang ML, Tang NS.** Sample size determination for the non-randomized triangular model for sensitive questions in survey. *Statistical Methods in Medicine Research*. 2009(0):1-15.
- Tian G-L, Tan M, Ng WN, Tang ML.** A unified method for checking compatibility and uniqueness for finite discrete conditional distributions. *Communications in Statistics: Theory and Methods*. 2009;38(1-2):115-29.
- Tian GL, Ng KW, Li K-C, Tan M.** Non-iterative sampling based Bayesian methods for identifying changpoints in the sequence of cases of Haemolytic uraemic syndrome. *Computer Statistics and Data Analysis*. 2009;53:3314-23.
- Tian GL, Fang HB, Tan M, Qin H.** Uniform distributions in a class of convex polyhedrons with applications to drug combination studies. *Journal of Multivariate Analysis*. 2009;100(8):1854-65.
- Tian GL, Fang HB, Liu Z, Tan M.** Regularized (bridge) logistic regression for variable selection based on ROC criterion. *Statistics and its Interface*. 2009;2:493-502.
- Ng KW, Tang M-L, **Tian GL, Tan M.** The nested dirichlet distribution and incomplete categorical data analysis. *Statistica Sinica*. 2009;19(1):251-71.

2010

- Fang H-B, Yu T, Tan M.** Efficient experiment design and nonparametric modeling of drug interaction. *Front Biosci*. 2010;E2:258-65.
- Fang H-B, Yu T, Tan M.** Efficient experiment design and nonparametric modeling of drug interaction. *Front Biosci*. 2010;E2:258-65.
- Liu Z, Lin S, Tan M.** Sparse Support Vector Machines with L_p Penalty for Biomarker Identification. *IEEE/ACM Trans Comput Biol Bioinformatics*. 2010;7(1):100-7.
- Liu Z, Lin S, Tan M.** Sparse Support Vector Machines with L_p Penalty for Biomarker Identification. *IEEE/ACM Trans Comput Biol Bioinformatics*. 2010;7(1):100-7.
- Liu Z, Chen D, Tian G-L, Tang M-L, Tan M.** Efficient Support Vector Machine Method for Survival Prediction with SEER Data. In: Arabnia HR, editor. *Advances in Computational Biology*. New York: Springer; 2010.
- Tan M, Fang HB, Tian GL.** Efficient experimental design and nonparametric modeling: a re-analysis of joint effects of trimetrexate combined with AG2034. *Frontiers in Bioscience (Elite Ed)* 2010(2):258-65.
- Tan M, Xiong X.** A Flexible and Optimized Multi-Stage Design for Phase II Trials and Toxicity Monitoring *Pharmaceutical Statistics*. 2010 (in press).
- Tan M, Fang HB, Tian GL.** Efficient experimental design and nonparametric modeling: a re-analysis of joint effects of trimetrexate combined with AG2034. *Frontiers in Bioscience (Elite Ed)* 2010(2):258-65.
- Tan M, Xiong X.** A flexible multi-stage design for phase II oncology trials. *Pharm Stat. Pharmaceutical Statistics*. 8 Dec 2010. DOI: 10.1002/pst.478. 9:478-, 2010.
- Tang L, **Tan M**, Zhou X-H. A Sequential Conditional Probability Ratio Test Procedure for Comparing Diagnostic Tests. *Journal of Applied Statistics*. 2010 (in press).
- Tian G-L, Gopalakrishnan M and Tan M.** Analyzing zero-inflated count data using noniterative Bayesian sampling. *Communications in Statistics: Theory and Methods* (in review). 2010.