

KUN CHEN

CONTACT 314 Philip E. Austin Building Phone: (860) 486-4847
Department of Statistics E-mail: kun.chen@uconn.edu
University of Connecticut Web: kun-chen.uconn.edu

215 Glenbrook Rd. U-4120
Storrs, CT 06269-4120

EDUCATION *Ph.D. in Statistics* 2011
Department of Statistics and Actuarial Science, University of Iowa, Iowa
City, Iowa, USA
Thesis: Regularized Multivariate Stochastic Regression
Advisor: Dr. Kung-Sik Chan

M.S. in Statistics 2007
Department of Mathematics and Statistics, University of Alaska Fairbanks,
Fairbanks, Alaska, USA

B.Econ. in Finance 2003
Dual B.S. in Computer Science and Technology 2003
Department of Statistics and Finance, University of Science and Technol-
ogy of China, Hefei, Anhui, China

EMPLOYMENT

08/2018– *Associate Professor (Tenured)*, Department of Statistics, University of Con-
necticut, Storrs, CT

01/2017– *Affiliated Faculty*, Institute for Collaboration on Health, Intervention, and Pol-
icy (InCHIP), University of Connecticut, Storrs, CT

05/2014– *Research Fellow*, Center for Population Health & Health Policy, University of
Connecticut Health Center, Farmington, CT

10/2013– *Affiliated Faculty*, Center for Environmental Sciences and Engineering, Uni-
versity of Connecticut, Storrs, CT

08/2013–08/2018 *Assistant Professor*, Department of Statistics, University of Connecticut, Storrs,
CT

08/2011–08/2013 *Assistant Professor*, Department of Statistics, Kansas State University, Man-
hattan, KS

HONORS & AWARDS

- Recognition for Teaching Excellence, UConn Spr. 2015, Fall 2015, Fall 2016, Fall 2018
- Elected Member of International Statistical Institute (ISI) Since 2015

- Travel Award, IMS New Researchers' Conference. 2014
- Travel Award, Junior Researchers' Workshop at ENAR. 2012
- ENAR Distinguished Student Paper Award, International Biometric Society 2011
- Henry L. Rietz Award, University of Iowa 2009
- Provincial Outstanding College Graduate, Anhui Province, China 2003
- Outstanding Graduate Award, University of Science and Technology of China 2003

RESEARCH INTERESTS

- Multivariate statistical learning; integrative statistical learning
- Machine learning and big data
- Statistical computing and optimization
- Population health and healthcare analytics
- Agricultural, ecological and environmental statistics

PUBLICATIONS

THEORY & METHODOLOGIES (*Student/Post-doc author[†]; corresponding author**)

- [1] Liu[†], X., Ma, S., and Chen*, K. (2021) Multivariate functional regression via a nested reduced-rank regularization. *Journal of Computational & Graphical Statistics*. Accepted. [arXiv:2003.04786](https://arxiv.org/abs/2003.04786).
- [2] Li[†], Y., Yu, C., Zhao, Y., Aseltine, R., Yao, W., and Chen*, K. (2021) Pursuing sources of heterogeneity in modeling clustered population. *Biometrics*. Accepted. (2020 ENAR Distinguished Student Paper Award). [arXiv:2003.04787](https://arxiv.org/abs/2003.04787).
- [3] Halder, A., Mohammed, S., Chen, K., and Dey, D. (2021) Spatial Tweedie exponential dispersion models: An application to insurance rate-making. *Scandinavian Actuarial Journal*. In press. [arXiv:2003.06299](https://arxiv.org/abs/2003.06299).
- [4] Mishra[†], A., Chen, Y., Dey, D. K., and Chen, K. (2021) Generalized co-sparse factor regression. *Computational Statistics & Data Analysis*, 157:107–127.
- [5] Vaughan[†], G., Aseltine, R., Chen, K., and Yan, J. (2020) Efficient interaction selection via stagewise generalized estimation equations. *Statistics in Medicine*, 39(22):2855–2868.
- [6] Sun[†], Z., Xu, W., Cong, X., Li, G., and Chen*, K. (2020) Log-contrast regression with functional compositional predictors: Linking preterm infant's gut microbiome trajectories to neurobehavioral outcome. *Annals of Applied Statistics*, 14(3):1535–1556. (2020 John van Ryzin Award).
- [7] Wang[†], W., Aseltine, R., Chen*, K., and Yan, J. (2020) Integrative survival analysis with uncertain event times in application to a suicide risk study. *Annals of Applied Statistics*, 14(1):51–73.
- [8] Uematsu, Y., Fan, Y., Chen, K., Lv, J., and Lin, W. (2019) SOFAR: large-scale association network learning. *IEEE Transactions on Information Theory*, 65(8):4924–4939.

- [9] Li, G., Liu[†], X., and Chen*, K. (2019) Integrative multi-view regression: Bridging group sparse and low-rank models. *Biometrics*, 75(2):593–602.
- [10] He[†], L., Chen*, K., Xu[†], W., Zhou, J., and Wang, F. (2018) Boosted sparse and low-rank tensor regression. In *Advances in Neural Information Processing Systems (NeurIPS) 31*, pages 1009–1018. Curran Associates, Inc.
- [11] Liang[†], J., Chen*, K., Lin, M., Zhang, C., and Wang, F. (2018) Robust finite mixture regression for heterogeneous targets. *Data Mining & Knowledge Discovery*, 32:1509–1560.
- [12] Zhang, X., He, L., Chen, K., Luo, Y., Zhou, J., and Wang, F. (2018) Information fusion with multi-view graph convolutional network and its applications on neuroimage analysis for patients with Parkinson’s disease. *American Medical Informatics Association (AMIA) Annual Symposium (Regular Paper)*, pages 1147–1156.
- [13] Luo[†], C., Liang[†], J., Li, G., Wang, F., Dey, D. K., and Chen*, K. (2018) Leveraging mixed-type and incomplete outcomes via a generalized reduced rank regression. *Journal of Multivariate Analysis*, 167:378–394.
- [14] Chen*, K., Mishra, N., Smyth, J., Bar, H., Schifano, E., Kuo, L., and Chen, M.-H. (2018) A tailored multivariate mixture model for detecting proteins of concordant change in the pathogenesis of *Necrotic Enteritis*. *Journal of the American Statistical Association*, 113:546–559.
- [15] Vaughan[†], G., Aseltine, R., Chen, K., and Yan, J. (2017) Stagewise generalized estimation equations with grouped variables. *Biometrics*, 73:1332–1342.
- [16] Mishra[†], A., Dey, D. K., and Chen*, K. (2017) Sequential co-sparse factor regression. *Journal of Computational & Graphical Statistics*, 26(4):814–825.
- [17] She, Y. and Chen*, K. (2017) Robust reduced-rank regression. *Biometrika*, 104(3):633–647.
- [18] Goh[†], G., Dey, D. K., and Chen, K. (2017) Bayesian sparse reduced rank multivariate regression. *Journal of Multivariate Analysis*, 157:14–28.
- [19] Lachos, V. H., Moreno, E. J., Chen, K., and Cabral, C. R. B. (2017) Finite mixture modeling of censored data using the multivariate student-*t* distribution. *Journal of Multivariate Analysis*, 159:151–167.
- [20] Chen*, K. and Ma, Y. (2017) Analysis of double single index models. *Scandinavian Journal of Statistics*, 44(1):1–20.
- [21] Yu[†], C., Yao, W., and Chen, K. (2017) A new method for robust mixture regression and outlier detection. *Canadian Journal of Statistics*, 45(1):77–94.
- [22] Chen*, K., Hoffman, E. A., Seetharaman[†], I., Lin, C.-L., and Chan, K.-S. (2016) Linking lung airway structure to pulmonary function via composite bridge regression. *Annals of Applied Statistics*, 10(4):1880–1906.
- [23] Chen*, K. (2016) Model diagnostics in reduced rank estimation. *Statistics and Its Interface*, 9(4):469–484.
- [24] Luo[†], C., Liu, J., Dey, D. K., and Chen*, K. (2016) Canonical variate regression. *Biostatistics*, 17(3):468–483.

- [25] Chen*, K. and Chan, K.-S. (2016) A note on rank reduction in sparse multivariate regression. *Journal of Statistical Theory and Practice*, 10(1):100–120.
- [26] Bai[†], X., Chen, K., and Yao, W. (2016) Mixture of linear mixed models using multivariate t distribution. *Journal of Statistical Computation and Simulation*, 86(4):771–787.
- [27] Gan, G. and Chen, K. (2016) A soft subspace clustering algorithm with log-transformed distances. *Big Data and Information Analytics*, 1(1):93–109.
- [28] Mukherjee[†], A., Chen, K., Wang, N., and Zhu, J. (2015) On the degrees of freedom of reduced-rank estimators in multivariate regression. *Biometrika*, 102(2):457–477.
- [29] Yu[†], C., Chen, K., and Yao, W. (2015) Outlier detection and robust mixture modeling using nonconvex penalized likelihood. *Journal of Statistical Planning and Inference*, 164:27–38.
- [30] Dong, H., Chen, K., and Linderoth, J. T. (2015) Regularization vs. relaxation: A conic optimization perspective of statistical variable selection. *CoRR*, abs/1510.06083.
- [31] Chen*, K., Chan, K.-S., and Stenseth, N. C. (2014) Source-sink reconstruction through regularized multicomponent regression analysis—with application to assessing whether North Sea cod larvae contributed to local fjord cod in Skagerrak. *Journal of the American Statistical Association*, 109:560–573.
- [32] Shi, J., Chen, K., and Song, W. (2014) Robust errors-in-variables linear regression by Laplace distribution. *Statistics & Probability Letters*, 84:113–120.
- [33] Chen*, K., Dong, H., and Chan, K.-S. (2013) Reduced rank regression via adaptive nuclear norm penalization. *Biometrika*, 100(4):901–920.
- [34] Chen, K., Chan, K.-S., and Stenseth, N. C. (2012) Reduced rank stochastic regression with a sparse singular value decomposition. *Journal of the Royal Statistical Society: Series B*, 74(2):203–221.
- [35] Chen, K. and Chan, K.-S. (2011) Subset ARMA selection via the adaptive lasso. *Statistics and Its Interface*, 4:197–205.
- [36] Chen, K., Jiang, W., and Tanner, M. A. (2010) A note on some algorithms for the Gibbs posterior. *Statistics & Probability Letters*, 80(15–16):1234–1241.

APPLICATIONS

- [37] Johnson, B. T., Sisti, A., Bernstein, M., Chen, K., Hennessy, E. A., Acabchuk, R. L., and Matos, M. (2021) Community-level factors and incidence of gun violence in the United States, 2014–2017. *Social Science & Medicine*. In press.
- [38] Aleman, J. O., Henderson, W. A., Walker, J. M., Ronning, A., Jones, D., Walter, P. J., Daniel, S. G., Bittinger, K., Vaughan, R., MacArthur, R., Chen, K., Breslow, J. L., and Holt, P. R. (2021) Excess dietary fructose does not alter gut microbiota or permeability in humans: a pilot randomized controlled study. *Journal of Clinical and Translational Science*. Accepted.

- [39] Wang, J., Tang, K., Feng, K., Lin, X., Lv, W., Chen, K., and Wang, F. (2021) Impact of temperature and relative humidity on the transmission of COVID-19: a modelling study in China and the United States. *BMJ Open*, 11:e043863.
- [40] Li, X., Tan, A., Chen, K., Pan, Y., Gentry, T., and Dou, F. (2021) Effect of cover crop type and application rate on soil nitrogen mineralization and availability in organic rice production. *Sustainability*, 13(5):2866.
- [41] Chen, K. and Wang, F. (2021) Discussion on “The timing and effectiveness of implementing mild interventions of COVID-19 in large industrial regions via a synthetic control method”. *Statistics and Its Interface*, 14(1):15–17.
- [42] Chang[†], S., Aseltine, R., Riddhi[†], D., Chen, K., Rogers, S., and Wang, F. (2020) Machine learning for suicide risk prediction in children and adolescents with electronic health records. *Translational Psychiatry*, 10, 413.
- [43] Doshi[†], R., Chen, K., Wang, F., Schwartz, H., Herzog, A., and Aseltine, R. (2020) Identifying risk factors for mortality among patients previously hospitalized for a suicide attempt. *Scientific Reports*, 10:15223.
- [44] Li, X., Dou, F., Guo, J., Velarca, M. V., Chen, K., Gentry, T., and McNear, D. (2020) Soil microbial community responses to nitrogen application in organic and conventional rice (*Oryza Sativa* L.) production. *Soil Science Society of America Journal*, 84(6):1885–1897.
- [45] Li, X., Dou, F., Watkins, K. B., Wang, S., Chen, K., Zhou, X., McClung, A., Storlien, J. O., and Hons, F. M. (2020) Seeding rate effects on organic rice growth, yield, and economic returns. *Agronomy Journal*, 112(5):4104–4119.
- [46] Liu, Y., Huang, J., Urbanowicz, R. J., Chen, K., Manduchi, E., Greene, C. S., Scheet, P., Moore, J. H., and Chen, Y. (2020) Embracing heterogeneity for finding genetic interactions in large-scale research consortia. *Genetic Epidemiology*, 44(1):52–66.
- [47] Zhou, C., Huang, Y., Jia, B., Wang, S., Dou, F., Samonte, S., Chen, K., and Wang, Y. (2019) Optimization of nitrogen rate and planting density for improving grain yield of different rice genotypes in northeast China. *Agronomy*, 9(9):555.
- [48] Doshi, R., Aseltine, R., Wang, F., Schwartz, H., Rogers, S., and Chen, K. (2018) Illustrating the role of health information exchange in a learning health system: Improving the identification and management of suicide risk. *Connecticut Medicine*, 82(6):327–333.
- [49] Dou, F., Ping, C.-L., Li, X., Jorgenson, T., Guo, L., Chen, K., and Michaelson, G. (2017) Soil organic carbon reactivity along the eroding coastline of northern Alaska. *Soil Science*, 182(6):227–232.
- [50] Mickelsen, L., Kolling, F., Chimileski, B., Fujita, A., Norris, C., Chen, K., Nelson, C., and Jackson, A. (2017) Neurochemical heterogeneity among lateral hypothalamic hypocretin/orexin and melanin-concentrating hormone neurons identified through single cell gene expression analysis. *eNeuro*, 4(5):13–17.
- [51] Chen, K. and Aseltine, R. (2017) Using hospitalization and mortality data to target suicide prevention activities. *Journal of Adolescent Health*, 61:192–197.

- [52] Choi, S., Hoffman, E. A., Wenzel, S. E., Castro, M., Fain, S., Jarjour, N., Schiebler, M. L., Chen, K., and Lin, C.-L. (2017) Quantitative computed tomography imaging-based clustering differentiates asthmatic subgroups with distinctive clinical phenotypes. *Journal of Allergy and Clinical Immunology*, 140(3):690–700.
- [53] Chen[†], Y., Chen, K., and Kalichman, S. C. (2017) Barriers to HIV medication adherence in the context of regimen simplification. *Annals of Behavioral Medicine*, 51(1):67–78.
- [54] Dou, F., Lee, T., Chen, K., Wright, A., and Mohammad, A. (2016) Planting date and variety effects on rice main and ratoon crop production in south Texas. *Communications in Soil Science and Plant Analysis*, 47(21):2414–2440.
- [55] Dou, F., Soriano, J., Tabien, R., and Chen, K. (2016) Soil texture and cultivar effects on rice (*Oryza sativa*, L.) grain yield, yield components and water productivity in three water regimes. *PLoS ONE*, 11(3):e0150549.
- [56] Soriano, J., Dou, F., Tabien, R., Harper, C., and Chen, K. (2016) Growth, development, yield and harvest index of two diverse rice cultivars in different water regimes and soil textures. *International Journal of Agronomy and Agricultural Research*, 8(2):82–94.
- [57] Choi, S., Hoffmann, E. A., Wenzel, S. E., Castro, M., Fain, S. B., Jarjour, N. N., Schiebler, M. L., Chen, K., and Lin, C.-L. (2015) Quantitative assessment of multiscale structural and functional alterations in asthmatic populations. *Journal of Applied Physiology*, 118(10):1286–1298.
- [58] Chen, K., Ciannelli, L., Decker, M., Ladd, C., Cheng, W., Zhou, Z., and Chan, K.-S. (2014) Reconstructing source-sink dynamics in a population with a pelagic dispersal phase. *PLoS ONE*, 9(5):e95316.
- [59] Chen, K., Chan, K.-S., Bailey, K., Aydin, K., and Ciannelli, L. (2012) A probabilistic cellular automata approach for predator-prey interactions of arrowtooth flounder (*Atheresthes stomias*) and walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(2):259–272.
- [60] Hunsicker, M., Ciannelli, L., Bailey, K., Buckel, J., White, J., Link, J., Essington, T., Anderson, T., Brodeur, R., Chan, K.-S., Chen, K., Englund, G., and et. al. (2011) Functional responses and scaling in marine predator-prey interactions: contemporary issues and emerging concepts. *Ecology Letters*, 14(12):1288–1299.

BOOKS & BOOK CHAPTERS

- [61] Lin, J., Wang, B., Hu, X., Chen, K., and Liu, R., editors (2016) *Statistical Applications from Clinical Trials and Personalized Medicine to Finance and Business Analytics*. Springer, International.
- [62] Luo[†], C., Dey, D. K., and Chen*, K. (2016) Partially supervised sparse factor regression model for multi-class classification. In Lin, J., Wang, B., Hu, X., Chen, K., and Liu, R., editors, *Statistical Applications from Clinical Trials and Personalized Medicine to Finance and Business Analytics*, pages 323–335. Springer, International.

CONFERENCE PAPERS/PROCEEDINGS

- [63] Ghahramani, G., Brendel, M., Chen, Q., Keenan, T., Chen, K., Chew, E., Lu, Z., Peng, Y., and Wang, F. (2021) Deep learning survival analysis on the progression to late AMD in the Age-Related Eye Disease Study. *Annual Meeting of Association for Research in Vision and Ophthalmology (ARVO)*.
- [64] Li, F., Liu, W., Druhl, E., Tucker, K., Lingeman, J., Pogoda, T., Wang, F., Chen, K., Aseltine, R., Kerns, R., Becker, W., Berlowitz, D., Ney, J., Carlson, K., and Yu, H. (2020) A pilot study of extracting social determinants of health from clinical text. *The 13th ACM International WSDM Conference (Workshop Paper)*.
- [65] Rogers, S., deMayo, R., Chen, K., Wang, F., and Aseltine, R. (2018) EHR phenotyping & data-driven suicide prevention. *American Medical Informatics Association (AMIA) Informatics Summit*.
- [66] Chen, K., Wang, F., and Aseltine, R. (2016) Using hospitalization and suicide mortality data to identify subpopulation of high suicide risk via survival modeling. *American Medical Informatics Association (AMIA) Annual Symposium*.
- [67] Choi, S., Chen, K., Hoffman, E., Wenzel, S., Castro, M., Fain, S., Jarjour, N., Schiebler, M., and Lin, C.-L. (2015) Linking and clustering multiscale structural and functional variables in asthmatic populations. *American Journal of Respiratory and Critical Care Medicine*, 191:A2464.
- [68] Zhang, M., Chen, K., Sparrow, S., Bechtel, P., and Pantoja, A. (2008) Simulating CO₂ released from soil: a Bayesian approach. *Eos Trans. AGU*, 89(53):Fall Meet. Suppl., Abstract B11D-0404.

MANUSCRIPTS

- [69] Cui, S., Liang, J., Pan, W., Chen, K., Zhang, C., and Wang, F. (2021) Learning to collaborate. Submitted to NeurIPS 2021.
- [70] Xu, W., Chang, S., Li, Y., Doshi, R., Chen, K., Wang, F., and Aseltine, R. (2021) Improving suicide risk prediction via targeted data fusion: proof of concept with statewide data. JAMIA.
- [71] Chen, K., Li, Y., and Li, G. (2021) Principal amalgamation analysis for compositional data. Technical Report.
- [72] Liu[†], X., Cong, X., Li, G., Mass, K., and Chen^{*}, K. (2021) Multivariate log-contrast regression with sub-compositional predictors: testing the associations between preterm infant’s gut microbiomes and neurobehavioral outcomes. *Statistics in Medicine*. In revision. [arXiv:2006.00487](https://arxiv.org/abs/2006.00487).
- [73] Pan, W., Chang, S., Chen, K., Henchcliffe, C., and Wang, F. (2021) Learning phenotypic associations for Parkinson’s disease with longitudinal clinical records. *npj Parkinson’s Disease*. DOI:10.1101/2020.03.15.20036657.
- [74] Wang[†], B., Song, Y., and Chen^{*}, K. (2021) Targeted integrative learning via a directional distance segmented regression.

- [75] Ghahramani, G., Brendel, M., Lin, M., Chen, Q., Keenan, T., Chen, K., Chew, E., Lu, Z., Peng, Y., and Wang, F. (2021) Deep learning-based survival analysis on the prognosis of late AMD using the longitudinal data in AREDS. *American Medical Informatics Association (AMIA) Annual Symposium (Regular Paper)*.
- [76] Zhou, Y., Cantoni, C., Lin, Q., Liu, Z., Pan, Y., Chen, K., Han, Y., Xiao, H., Gormley, M., Liu, Y., Bokoliya, S., Panier, H., Suther, C., Evans, E., Deng, L., Newland, P., Wu, Y., Tarr, P., Lovett-Racke, A., and Piccio, L. (2021) Alterations of host-gut microbiome interactions in multiple sclerosis. *Genome Medicine*. In revision.
- [77] Li[†], Y., Chen, K., Yan, J., and Zhang, X. (2021) Uncertainty in optimal fingerprinting is underestimated. *Environmental Research Letters*. In revision.
- [78] Luo, C., Chen, K., Doshi, R., Rickles, N., Chen, Y., Schwartz, H., and Aseltine, R. H. (2021) The association of prescription opioid use with suicide attempts: An analysis of statewide medical claims data. Submitted.
- [79] Wang, X., Li, X., Dou, F., Sun, W., Chen, K., Wen, Y., and Ma, X. (2021) Elucidating the impact of three metallic nanoagrichemicals and their bulk and ionic counterparts on the chemical properties of bulk and rhizosphere soils in rice paddies. *Environmental Pollution*. Submitted.
- [80] Li, Y., Chen, K., Wan, H., Yan, J., and Zhang, X. (2021) Detection and attribution analysis of changes in regional temperature with properly calibrated confidence intervals. Technical Report.
- [81] Li, G., Li, Y., and Chen, K. (2020) It's all relative: new regression paradigm for microbiome compositional data. [arXiv:2011.05951](https://arxiv.org/abs/2011.05951).
- [82] Wang[†], W., Luo, C., Aseltine, R., Wang, F., Yan, J., and Chen^{*}, K. (2020) Suicide risk modeling with uncertain diagnostic records. [arXiv:2009.02597](https://arxiv.org/abs/2009.02597).
- [83] Xu, T., Chen, K., and Li, G. (2020) The more data, the better? Demystifying deletion-based methods in linear regression with missing data. [arXiv:2010.13332](https://arxiv.org/abs/2010.13332).
- [84] Li, X., Guo, J., Velarca, M. V., McClung, A., Chen, K., and Dou, F. (2020) Effect of nitrogen application rate under organic and conventional systems on rice growth and productivity. *Submitted*.
- [85] Chen^{*}, K., Dong, R., Zheng, Z., and Xu[†], W. (2020) Statistically guided divide-and-conquer for sparse factorization of large matrix. [arXiv:2003.07898](https://arxiv.org/abs/2003.07898).
- [86] Pan, W., Cui, S., Wen, H., Chen, K., Zhang, C., and Wang, F. (2020) Correcting the user feedback-loop bias for recommendation systems. *KDD*.
- [87] Pan, W., Cui, S., Chen, K., Zhang, C., and Wang, F. (2020) Adaptive learning for clinical risk prediction models with electronic health records from multiple hospitals. *KDD*.
- [88] Halder, A., Mohammed, S., Chen, K., and Dey, D. (2020) Spatial risk estimation in Tweedie compound Poisson double generalized linear models. [arXiv:1912.12356](https://arxiv.org/abs/1912.12356).

- [89] Li[†], Y., Chen, K., Yan, J., and Zhang, X. (2019) Regularized fingerprinting in detection and attribution of climate change with weight matrix optimizing the efficiency in scaling factor estimation.

SOFTWARE

- [90] Zhe, S. and Chen, K. *Compact: Regression with Compositional Covariates* (2020). R package version 0.1.0.
- [91] Wang, W., Chen, K., and Yan, J. *intsurv: Integrative Survival Modeling* (2019). R package version 0.2.0.
- [92] Li, Y., Chen, K., and Yan, J. *tls: Tools of Total Least Squares in Error-in-Variables Models* (2018). R package version 0.1.0.
- [93] Chen, K. *rrpack: Reduced-Rank Regression* (2017). R package version 0.1-5.
- [94] Luo, C. and Chen, K. *CVR: Canonical Variate Regression* (2017). R package version 0.1.
- [95] Vaughan, G., Chen, K., and Yan, J. *sgee: Stagewise Generalized Estimating Equations* (2017). R package version 0.2.
- [96] Mishra, A. and Chen, K. *secure: Sequential Co-Sparse Factor Regression* (2017). R package version 0.5.

(Other computational packages are available at kun-chen.uconn.edu/code/.)

GRANTS

FUNDED

- Developing suicide risk algorithms for diverse clinical settings using data fusion. National Institutes of Health ([R01-MH124740](#)), PI (with Dr. Robert Aseltine and Dr. Fei Wang); 09/16/2020–06/30/2024. Active.
- Improving suicide prediction using NLP-derived social determinants of health. National Institutes of Health ([R01-MH125027](#)). PI on sub-award (PI: Dr. Hong Yu at UMass & VA); 09/01/2020–06/30/2024. Active.
- Improving the identification and management of suicide risk among patients using prescription opioids (HEAL Supplement). National Institutes of Health ([R01-MH112148-03S1](#)). PI on sub-award (PI: Dr. Robert Aseltine); 09/18/2020–06/30/2021. Active.
- Reciprocal modulation of the microbiome and cellular senescence in metabolic dysfunction. National Institutes of Health ([R01-AG068860](#)), PI on sub-award (PI: Dr. Yanjiao Zhou at UCHC); 09/10/2020–05/31/2025. Active.
- Understanding community- and individual-level factors underlying firearm violence in America: Focus on the State of Connecticut. The Interdisciplinary CLAS Pilot Grant Program. Co-PI (with Dr. Mary Bernstein from Sociology and Dr. Blair Johnson from Psychological Sciences); 12/2019–12/2020. Active.
- Comprehensive heterogeneous response regression from complex data. National Science Foundation ([IIS-1718798](#)), PI; \$250,000; 09/01/2017–08/31/2021. Active.

- Improving the identification of patients at risk of suicide. National Institutes of Health (R01-MH112148). PI on sub-award (PI: Dr. Robert Aseltine); \$432,886; 07/01/2017–06/30/2020. Active.
- Integrative multivariate analysis with multi-view data. National Science Foundation (DMS-1613295). PI; \$150,000; 09/01/2016–08/31/2020. Active.
- Modeling and analysis of large insurance claim and occurrence data: a partnership between UConn & Travelers. Travelers Insurance. Co-PI (PI: Dr. Dipak Dey); 08/01/2016–07/31/2020. Active.
- Data Science Lab: Real-world data science problems meet future data scientists. UConn CLAS Fund for Innovative Education in Science. PI (with Elizabeth Schifano and Jun Yan); \$65,000; 01/01/2017–12/31/2019. Completed.
- An integrative statistics-guided image-based multi-scale lung model. National Institutes of Health (U01-HL114494). PI on sub-award; \$173,521; 08/01/2013–05/31/2018. Completed.
- Structured low rank modeling for multivariate statistical learning. Simons Foundation (#359404). PI; \$35,000; 09/2015–08/2020. Completed in 2016 (due to other support).
- Garrett Lee Smith Suicide Prevention. U.S. Substance Abuse and Mental Health Services Administration. PI on sub-award; \$20,995; 07/01/2016–12/31/2016. Completed.
- New England eConsult Network. PI on sub-award; \$23,207; 01/19/2016–05/30/2016. Completed.
- Integrative multivariate analysis with multiple sets of variables of high dimensionality. UConn Faculty Large Grant. PI; \$19,399; 01/01/2014–12/31/2014. Completed.
- Source-sink reconstruction through regularized multi-component regression analysis. K-State Faculty Development Award. PI; \$1,000; 06/01/2013–12/31/2013. Completed.
- High-dimensional multivariate modeling via matrix decomposition and regularization. K-State Faculty Enhancement Award. PI; \$10,000; 01/01/2012–12/31/2012. Completed.

TEACHING

@DEPARTMENT OF STATISTICS, UNIVERSITY OF CONNECTICUT

- STAT:5725 Linear Statistical Models I (3 cr., graduate level; *Fall 2016.*)
- STAT:6694 Linear Statistical Models II (3 cr., graduate level; *Spring 2019.*)
- STAT:3375 Introduction to Mathematical Statistics (3 cr., undergraduate level; *Fall 2013, Fall 2014.*)
- STAT:5665 Applied Multivariate Statistics (3 cr., graduate level; *Fall 2015, Fall 2018, Fall 2019*)
- STAT: 3515/5515 Design of Experiments (3 cr., undergraduate and graduate levels; *Spring 2015, Spring 2016, Spring 2017, Spring 2018, Spring 2021.*)
- STAT:5361 Statistical Computing (3 cr., graduate level; *Spring 2014, Spring 2015.*)
- STAT:3115/5315 Analysis of Experiments (3 cr., undergraduate and graduate levels; *Spring 2014, Spring 2017.*)
- STAT:6494 Data Science in Action (3 cr., graduate level; *Spring 2018.*)

@DEPARTMENT OF STATISTICS, KANSAS STATE UNIVERSITY

- STAT:905 High Dimensional Data and Statistical Learning (3 cr., graduate level; *Fall 2012*)
- STAT:730 Multivariate Statistical Methods (3 cr., graduate level; *Spring 2012, Spring 2013*)
- STAT:510 Introductory Probability and Statistics (3 cr., undergraduate level; *Fall 2011, Spring 2012, Spring 2013*)

@DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE, UNIVERSITY OF IOWA

- 22S:101 Biostatistics (3 cr., undergraduate level; *Fall 2010*)
- * 22S:002 Statistics and Society (3 cr., undergraduate level; *Fall 2007, Spring 2008, Fall 2008*)

@DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF ALASKA FAIRBANKS

- * STAT:F401 Regression and Analysis of Variance (4 cr., undergraduate level; *Fall 2005*)
(★ Taught as Teaching Assistant.)

SHORT COURSES & TUTORIALS

- “Integrative Multivariate Statistical Learning in Healthcare Research with Real-World Data”. Half-day short course, with Dingfeng Jiang. 2017 ICSA Applied Statistics Symposium, Chicago, IL. June 2017.
- “Practical Integrative Statistical Learning: Recent Developments and Case Studies”. Full-day short course, with Robert Aseltine. The 31th New England Statistics Symposium, Storrs, CT. April 2017.
- “Modern Multivariate Statistical Learning: Methods and Applications”. Full-day short course, with Jun Yan. The 29th New England Statistics Symposium, Storrs, CT. April 2015.

PRESENTATIONS

INVITED TALKS

1. Southwest University of Finance and Economics, Chengdu, China. September 2020.
2. Syracuse University. November 2019.
3. Mathematical Sciences, Worcester Polytechnic Institute. October 2019.
4. Department of Computer Science, University of Massachusetts Lowell. September 2019.
5. “Targeted integrative learning via distance segmented regression”. Joint Statistical Meeting, Denver, CO. August 2019.
6. “Boosted sparse and low-rank tensor regression”. 2019 Symposium on Data Science and Statistics, Seattle, WA. May 2019.
7. “Integrative Survival Analysis with Uncertain Event Times—Towards a Data Driven Suicide Prevention Framework”. Department of Biostatistics, Yale University, New Haven, CT. April, 2019.
8. “Integrative Survival Analysis with Uncertain Event Times—Towards a Data Driven Suicide Prevention Framework”. Mental Health Data Science, Department of Biostatistics, Columbia University, New York. December 2018.
9. “Integrative statistical learning with real world healthcare data: towards a data driven suicide prevention framework”. 2018 Joint Statistical Meeting, Vancouver, Canada. August 2018.

10. "Stagewise Co-Sparse and Low-Rank Matrix Factorization". 2018 ICSA China Conference, Qingdao, China. July 2018.
11. "Sparse Log-Contrast Regression with Functional Compositional Predictors". The 8th International Forum on Statistics, Renmin University of China, Beijing, China. July 2018.
12. "Stagewise Co-Sparse Low-Rank Matrix Decomposition". 2018 ICSA Applied Statistics Symposium. New Brunswick, New Jersey. June 2018.
13. "Sparse Log-Contrast Regression with Functional Compositional Covariates". 2018 Modern Modeling Methods Conference. Storrs, CT. May 2018.
14. "Stagewise Co-Sparse and Low-Rank Matrix Factorization". Baruch College. New York. April 2018.
15. "Integrate, Divide, and Conquer: On Sparse and Low-Rank Multivariate Statistical Learning". School of Statistics, Beijing Normal University. Beijing, China. December 2017.
16. "Integrate, Divide, and Conquer: On Sparse and Low-Rank Multivariate Statistical Learning". School of Statistics, Renmin University of China. Beijing, China. December 2017.
17. "Integrate, Divide, and Conquer: On Sparse and Low-Rank Multivariate Statistical Learning". Health Informatics PIC Distinguished Speaker Series, IBM Watson Research Center, New York. November 2017.
18. "On Sparse and Low-Rank Models for Integrative Multivariate Statistical Learning". Biostatistics and Epidemiology, Weill Cornell Medical College, New York. November 2017.
19. "Dealing with uncertain suicidal deaths due to imperfect data integration: a first step towards a data-driven suicide prevention framework". Department of Mathematics and Statistics, Boston University, Boston, MA. October 2017.
20. "Integrative Cox regression for modeling uncertain survival records due to imperfect data integration". 2017 ICSA Applied Statistical Symposium, Chicago, IL. June 2017.
21. "Leveraging mixed and incomplete outcomes via reduced-rank regression". Modern Modeling Methods Conference, Storrs, CT. May 2017.
22. "On integrative learning of mixed and incomplete data". IMS/ASA Spring Research Conference, Rutgers University, New Brunswick, NJ. May 2017.
23. "Regularized mixture regression with mixed and incomplete outcomes". The 31st New England Statistics Symposium, Storrs, CT. April 2017.
24. "Using hospitalization and suicide mortality data to identify subpopulation of high suicide risk via survival modeling". AMIA 2016 Annual Symposium, Chicago, IL. November 2016.
25. "On large-scale predictive modeling of mixed and incomplete outcomes". Department of Mathematics & Statistics, University of Massachusetts, Amherst, MA. October 2016.
26. "Canonical variate regression for integrative analysis of genomics data".
 - Joint Statistical Meeting, Chicago. Topic-contributed session. August 2016.
 - Department of Biostatistics, Columbia University. Guest lecture. May 2016.
28. "Model diagnostics in reduced rank estimation". ICSA Applied Statistics Symposium, Atlanta, GA. June 2016.
29. "Robust multivariate mixture model via mean-shift penalization". Modern Modeling Methods Conference, Storrs, CT. May 2016.
30. "A tailored robust multivariate clustering approach via mean-shift penalization".
 - Center for Statistical Science, Peking University, Beijing, China. May 2016.

- ICSA Conference on Data Science, Yunnan, China. July 2016.
- 32. “Sequential estimation in sparse factor regression”.
 - Conference on Statistical Learning and Data Science, University of North Carolina at Chapel Hill. June, 2016.
 - Big Statistics & Data Science Joint Conference, Renmin University, China. May 2016.
 - Department of Statistics, Florida State University. April 2016.
 - Department of Mathematics and Statistics, Boston University. April 2016.
 - Department of Statistics, University of South Carolina. March 2016.
- 37. “Canonical variate regression”.
 - Department of Statistics, University of Missouri Columbia. September 2015.
 - 60th World Statistics Congress (ISI 2015). August 2015.
- 39. “Linking lung airway structure to pulmonary function via hierarchical feature selection.” ICSA Applied Statistics Symposium and 13th Graybill Conference, Fort Collins, CO. June 2015.
- 40. “Some recent developments on reduced rank modeling”. Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA. November 2014.
- 41. “On sparse and low-rank estimation in high dimensions”. Department of Statistics, Kansas State University, Manhattan, KS. October 2014.
- 42. “On some low-rank models in multivariate time series analysis”. International Conference on Advances in Interdisciplinary Statistics and Combinatorics, Greensboro, NC. October 2014.
- 43. “Sparse and orthogonal factor regression”. 16th Meeting of New Researchers in Statistics and Probability. Harvard University, Boston, MA. August 2014.
- 44. “Sparse orthogonal factor regression in high dimensions with an application to eQTL data analysis”.
 - WNAR Annual Meeting, Honolulu, HI. June 2014.
 - The 6th International Statistics Forum at Renmin University, Beijing, China. May 2014.
- 46. “Some recent developments on multivariate modeling”. University of Science and Technology of China, Hefei, Anhui, China. May 2014.
- 47. “Sparse and low-rank regression in high dimensions. BIRS Workshop on Statistical and Computational Theory and Methodology for Big Data Analysis, Banff International Research Station, Banff, Alberta, Canada. February 2014.
- 48. “Source-sink reconstruction through regularized multi-component regression”.
 - Department of Biostatistics, Brown University, Providence, RI. September, 2013.
 - IMS-China International Conference on Statistics and Probability, Chengdu, China. July 2013.
 - Department of Statistics, University of Connecticut, Storrs, CT. February 2013.
- 51. “Reduced rank estimation and its extensions”. College of Mathematics and Software, Sichuan Normal University, Chengdu, China. June 2013.
- 52. “Adaptive reduced-rank estimation and its complexity”. Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA. November 2012.
- 53. “Statistical source-sink reconstruction”. Department of Statistics, Kansas State University, Manhattan, KS. October 2012.

54. "Regularized multivariate regression for rank reduction and variable selection". ICSA Applied Statistics Symposium, Boston, MA. June 2012.
55. "Some dimension reduction methods in high-dimensional multivariate regression". Department of Mathematics, Kansas State University, Manhattan, KS. April 2012.
56. "Reduced-rank stochastic regression with sparse singular value decomposition".
 - Distinguished Student Paper Award, ENAR Spring Meeting, Miami, FL. March 2011.
 - Department of Statistics, Kansas State University, Manhattan, KS. February 2011.
 - Department of Statistics, Oregon State University, Corvallis, OR. February 2011.
 - Department of Mathematics, Wake Forest University, Winston-Salem, NC. February 2011.
 - Department of Statistics, University of Wyoming, Laramie, WY. January 2011.
 - Department of Statistics, Western Michigan University, Kalamazoo, MI. January 2011.
 - Department of Statistics & Actuarial Science, University of Iowa, Iowa City, IA. September 2010.
63. "Statistical methods, cellular automata models and wavelets". International Workshop on Predator-prey Interactions in Marine Ecosystems, Oregon State University, Corvallis, OR. March 2010.

CONTRIBUTED TALKS

64. "Consistent bi-level variable selection via composite bridge regression". ENAR Spring Meeting, Baltimore, MD. February 2014.
65. "High dimensional multivariate regression via singular value decomposition". Frontiers in Biostatistical Methods Symposium, Kansas City, KS. May 2012.
66. "Simultaneous rank determination and variable selection in multivariate regression". ENAR Spring Meeting, Washington, DC. March 2012.
67. "Subset ARMA selection via regularization". Joint Statistical Meeting, Miami, FL. August 2011.
68. "Reduced-rank stochastic regression with a sparse singular value decomposition – Modeling larval drift effects on cod population dynamics as a case study". Joint Statistical Meeting, Vancouver, BC. August 2010.
69. "Scale-specific modeling using wavelets". Comparative Analysis of Marine Ecosystem Organization (CAMEO) Project Meeting, National Oceanic and Atmospheric Administration (NOAA), Seattle, WA. June 2009.

POSTERS

70. "Canonical variate regression". With Chongliang Luo, Jin Liu and Dipak Dey. The 29th New England Statistics Symposium, Storrs, CT. April 2015.
71. "Significant lung airway feature extraction via composite bridge regression: how airway remodeling impairs asthma quality of life". With Chan K. S., Hoffman E. and Lin C-L. Multiscale Modeling Consortium Meeting, Rockville, MD. September 2014.
72. "Sparse orthogonal factor regression". 16th Meeting of New Researchers in Statistics and Probability, Harvard University, Boston, MA. August 2014.

73. "Research highlights in statistical ecology". GLISEN (Joint University meeting between UConn and Stony Brook University), Groton, CT. October 2013.
74. "Lung-airway data interrogation via cluster analysis". With Chan K. S., Hoffman E. and Lin C-L. Multiscale Modeling Consortium Meeting, Rockville, MD. October 2011.
75. "Some MCMC algorithms for the Gibbs posterior". Third Midwest Statistics Research Colloquium, University of Chicago. March 2010.

STUDENTS

PH.D.

GRADUATED

- Zhe Sun, UConn Statistics (defended April 2021). On statistical modeling of longitudinal compositional data with applications in a preterm infant study. Post-doc at Yale.
- Xiaokang Liu, UConn Statistics (defended April 2020). Integrative multivariate learning via composite low-rank decompositions. Post-doc at UPenn.
- Wenjie Wang, UConn Statistics (defended August 2019). Integrative survival analysis. (Joint with Jun Yan). Eli Lilly and Company.
- Aditya Mishra, UConn Statistics (defended August 2017). On sequential estimation of multivariate associations. (Joint with Dipak Dey). Flatiron Research Fellow, Flatiron Institute, Simons Foundation.
- Chongliang Luo, UConn Statistics (defended July 2017). On integrative reduced-rank models and applications. (Joint with Dipak Dey). Postdoc Fellow, Department of Biostatistics, University of Pennsylvania.
- Gregory Vaughan, UConn Statistics (defended July 2017). Stagewise generalized estimating equations. (Joint with Jun Yan). Assistant Professor, Bentley University, Waltham, MA.
- Xiuqin Bai, K-State Statistics (defended June 2014) Robust mixture linear mixed effect model with repeated measure. (Joint with Weixin Yao). Assistant Professor, Eastern Washington University, Cheney, WA.
- Chun Yu, K-State Statistics (defended May 2014). Robust mixture regression model. (Joint with Weixin Yao). Assistant Professor, Jiangxi University of Finance and Economics, Jiangxi, China.

CURRENT

- Jianmin Chen, UConn Statistics (2021-).
- Boyang Tang, UConn Statistics (2021-).
- Jiadong Fang, UConn Statistics (2020-).
- Zhongmao Liu, UConn Statistics (2019-).
- Yeming Pan, UConn Statistics (2019-).
- Biju Wang, UConn Statistics (2018-).
- Wanwan Xu, UConn Statistics (2017-).
- Yan Li, UConn Statistics (2016-). Joint with Jun Yan.

SERVING/SERVED AS ASSOCIATE ADVISER

- Paul McLaughlin (PhD in Statistics, UConn, 2019).
- Yiyun Chen (PhD in Psychology, UConn, 2018).
- Yaohua Zhang (PhD in Statistics, UConn, 2017).
- Brian Bader (PhD in Statistics, UConn, 2016).

M.S.

- Indu Seetharaman, K-State Statistics (defended May 2013). Composite bridge regression for bi-level selection.
- Rohan Khatavkar, K-State Statistics (defended June 2013). Sparse and orthogonal singular value decomposition.

VISITING SCHOLAR

- Jian Liang, Tsinghua University, Beijing, China. 03/2016–09/2016.

STUDENT ACHIEVEMENTS & AWARDS

- Jackson Lautier, Graduate Research Fellowship Award, National Science Foundation, 2020.
- Zhe Sun, John van Ryzin Award for the Best Paper & ENAR Distinguished Student Paper Award, 2020.
- Yan Li, ENAR Distinguished Student Paper Award, 2020.
- Yan Li, IBM Student Paper Award, from the 33rd New England Statistics Symposium in 2019.
- Xiaokang Liu, IBM Student Paper Award, from the 33rd New England Statistics Symposium in 2019.
- Xiaokang Liu, Honorable Mention in Student Poster Competition, from the Fourth International Workshop on the Statistical Analyses of Multi-outcome Data (SAM) in 2018.
- Wanwan Xu, Honorable Mention in Student Poster Competition, from the Fourth International Workshop on the Statistical Analyses of Multi-outcome Data (SAM) in 2018.
- Wenjie Wang, Student Paper Award, the 31st New England Statistics Symposium in 2017.
- Gregory Vaughan, Student Paper Award, from the Mental Health Section of the American Statistical Association in 2017.
- Chongliang Luo, Student Paper Award, from the International Chinese Statistical Association in 2015.
- Gyuhyeong Goh, Student Paper Award, from the Section on Bayesian Statistical Science of the American Statistical Association in 2015.

PROFESSIONAL ASSOCIATIONS

- Member, New England Statistical Society (NESS) 2017–
- Elected Member, International Statistical Institute (ISI) 2015–
- Member, International Society for Business and Industrial Statistics (ISBIS) 2015–
- Member, International Biometrics Society, ENAR 2010–
- Member, International Chinese Statistical Association (ICSA) 2010–

- Member, Institute of Mathematical Statistics (IMS) 2009–
- Member, American Statistical Association (ASA) 2006–

PROFESSIONAL ACTIVITIES

POSITIONS

- Program Chair-Elect, Section on Statistical Computing, American Statistical Association (ASA) 2022–2023
- Executive Secretary, New England Statistical Society (NESS) 2017–
- Member, Connecticut All-Payer Claims Database (APCD) Data Release Committee (DRC), Office of Health Strategy, State of Connecticut 2017–

EDITORIAL WORK

- Editorial Board member, *Journal of Machine Learning Research* 2020–
- Associate Editor, *Sankhya Series B* 2016–
- Co-Editor, 2015 ICASA Symposium Proceeding Book 2015

REFEREE SERVICE

I've reviewed more than 50 papers for top statistical and machine learning journals. Below is a selected list.

- *Annals of Applied Statistics*
- *Annals of Statistics*
- *Biometrics*
- *Biometrika*
- *Journal of Business and Economic Statistics*
- *Journal of Computational and Graphical Statistics*
- *Journal of Machine Learning Research*
- *Journal of Multivariate Analysis*
- *Journal of the American Statistical Association*
- *Journal of the Royal Statistical Society: Series B*
- *Statistics and Computing*
- *Statistics in Medicine*
- *Statistica Sinica*
- *Technometrics*

CONFERENCE SERVICE

- Associate Program Chair, 2022 Joint Statistical Meeting.
- Organizing Committee. New England Rare Disease Statistics (NERDS) Workshop. 2019.
- Organizing Committee. UConn Sports Analytics Conference. 2019.
- Organizing Committee. ASA-BI-NESS Webinar Series. 2019–.
- Program Committee. The 33rd New England Statistics Symposium, Hartford, CT. May 2019.
- Organizing Committee. The 3rd Stat4Onc Annual Symposium, Hartford, CT. April 2019.

- Organizing Committee (Co-Chair). Conference on Bayesian Modeling, Computation, and Applications, Storrs, CT. May 2018.
- Organizing Committee; full-day short course instructor. The 31th New England Statistics Symposium, Storrs, CT. April 2017.
- Panelist on career development; judge on poster competition. The 30th New England Statistics Symposium, New Haven, CT. April 2016.
- Invited Session Organizer; session chair. ICSA Applied Statistics Symposium and 13th Graybill Conference, Fort Collins, CO. June 2015.
- Full-day short course instructor; invited session organizer; judge on student paper competition; judge on poster competition. The 29th New England Statistical Symposium, Storrs, CT. April 2015.
- Invited Session organizer, The 6th International Statistics Forum, Renmin University of China, May 2014, Beijing, China.
- Session chair, BIRS Big Data Workshop, February 2014, Banff, Calgary, Canada.
- Session chair, 2012 ENAR Meeting in Washington, DC

UNIVERSITY & DEPARTMENT SERVICE

AT UCONN

- Co-Chair, Q-Subcommittee of UConn General Education Oversight Committee 2018–2021
- Member, Committee of UConn General Education Oversight Committee 2018–
- Member, Q-Subcommittee of UConn General Education Oversight Committee 2014–
- Chair, Department Undergraduate Curriculum Committee 2016–
- Member, CLAS Curriculum & Course Committee 2016–2018
- Chair, Committee on New England Statistical Society 2016–2017
- Member, Department Colloquium Committee 2014–2018
- Member, Department 3 + 1 Admission Committee 2014–
- Member, Department Graduate Admissions Committee 2013–

AT K-STATE

- Chair, Ph.D. Exam Committee on Linear Models 2012–2013
- Chair, Department Seminar Spring 2013
- Member, Student Assessment Committee 2011–2013
- Member, Departmental Scholarships and Awards Committee 2011–2012
- Member, Graduate Student Progress Committee 2011–2012

Last updated: June 28, 2021