
Bibliography

- Agresti, A. (2013). *Categorical Data Analysis*. Wiley, 3rd edition.
- Aliferis, C. F., Statnikov, A., Tsamardinos, I., Mani, S., and Xenofon, X. D. (2010). Local Causal and Markov Blanket Induction for Causal Discovery and Feature Selection for Classification Part I: Algorithms and Empirical Evaluation. *Journal of Machine Learning Research*, 11:171–234.
- Andersen, S., Olesen, K., Jensen, F., and Jensen, F. (1989). HUGIN – a Shell for Building Bayesian Belief Universes for Expert Systems. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1080–1085. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley, 3rd edition.
- Andreassen, S., Jensen, F. V., Andersen, S. K., Falck, B., Kjærulff, U., Woldbye, M., Sørensen, A. R., Rosenfalck, A., and Jensen, F. (1989). MUNIN – An Expert EMG Assistant. In Desmedt, J. E., editor, *Computer-Aided Electromyography and Expert Systems*, pages 255–277. Elsevier.
- Ash, R. B. (2000). *Probability and Measure Theory*. Academic Press, 2nd edition.
- Balov, N. and Salzman, P. (2020). *catnet: Categorical Bayesian Network Inference*.
- Bang-Jensen, J. and Gutin, G. (2009). *Digraphs: Theory, Algorithms and Applications*. Springer, 2nd edition.
- Bøttcher, S. G. and Dethlefsen, C. (2003). deal: A Package for Learning Bayesian Networks. *Journal of Statistical Software*, 8(20):1–40.
- Bouckaert, R. R. (1995). *Bayesian Belief Networks: From Construction to Inference*. PhD thesis, Utrecht University, The Netherlands.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1):1–32.
- Carpenter, B., Hoffman, M. D., Brubaker, M., Lee, D., Li P., and Betancourt, M. (2015). *The Stan Math Library: Reverse-Mode Automatic Differentiation*.
- Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1997). *Expert Systems and Probabilistic Network Models*. Springer.
- Centers for Disease Control and Prevention (2004). *The 1999–2004 Dual Energy X-ray Absorptiometry (DXA) Multiple Imputation Data Files and Technical Documentation*. National Center for Health Statistics, Hyattsville, MD (USA).
- Centers for Disease Control and Prevention (2010). *National Health and Nutrition Examination Survey: Body Composition Procedures Manual*. National Center for Health Statistics, Hyattsville, MD (USA).

- Chao, Y.-S., Wu, H.-T., Scutari, M., Chen, T.-S., Wu, C.-J., Durand, M., and Boivin, A. (2017). A Network Perspective on Patient Experiences and Health Status: The Medical Expenditure Panel Survey 2004 to 2011. *BMC Health Services Research*, 17(579):1–12.
- Cheng, J. and Druzdzel, M. J. (2000). AIS-BN: An Adaptive Importance Sampling Algorithm for Evidential Reasoning in Large Bayesian Networks. *Journal of Artificial Intelligence Research*, 13:155–188.
- Cooper, G. F. and Yoo, C. (1999). Causal Discovery from a Mixture of Experimental and Observational Data. In *Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence*, pages 116–125. Morgan Kaufmann, San Francisco, CA, USA.
- Crawley, M. J. (2013). *The R Book*. Wiley, 2nd edition.
- Cussens, J. (2012). Bayesian Network Learning with Cutting Planes. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence*, pages 153–160.
- DeGroot, M. H. and Schervish, M. J. (2012). *Probability and Statistics*. Prentice Hall, 4th edition.
- Denis, J.-B. (2020). *rbmn: Handling Linear Gaussian Bayesian Networks*.
- Diestel, R. (2005). *Graph Theory*. Springer, 3rd edition.
- Edwards, D. I. (2000). *Introduction to Graphical Modelling*. Springer, 2nd edition.
- Eggeling, R., Viinikka, J., Vuoksenmaa, A., and Koivisto, M. (2019). On Structure Priors for Learning Bayesian Networks. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, pages 1687–1695.
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Wiley, 3rd edition.
- Friedman, N. and Koller, D. (2003). Being Bayesian about Bayesian Network Structure: A Bayesian Approach to Structure Discovery in Bayesian Networks. *Machine Learning*, 50(1–2):95–126.
- Friedman, N., Linial, M., Nachman, I., and Pe’er, D. (2000). Using Bayesian Network to Analyze Expression Data. *Journal of Computational Biology*, 7:601–620.
- Friedman, N., Pe’er, D., and Nachman, I. (1999). Learning Bayesian Network Structure from Massive Datasets: The “Sparse Candidate” Algorithm. In *Proceedings of 15th Conference on Uncertainty in Artificial Intelligence*, pages 206–221. Morgan Kaufmann, San Francisco, CA, USA.
- Gasse, M., Aussem, A., and Elghazel, H. (2014). A Hybrid Algorithm for Bayesian Network Structure Learning with Application to Multi-Label Learning. *Expert Systems with Applications*, 41(15):6755–6772.
- Goeman, J. J. (2018). *penalized R Package*.
- Hartemink, A. J. (2001). *Principled Computational Methods for the Validation and Discovery of Genetic Regulatory Networks*. PhD thesis, School of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.
- Hartigan, J. A. (1975). *Clustering Algorithms*. Wiley.
- Hartley, S. W. and Sebastiani, P. (2013). PleioGRiP: Genetic Risk Prediction with Pleiotropy. *Bioinformatics*, 29(8):1086–1088.

- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition.
- Hausser, J. and Strimmer, K. (2009). Entropy Inference and the James-Stein Estimator, with Application to Nonlinear Gene Association Networks. *Journal of Machine Learning Research*, 10:1469–1484.
- Heckerman, D., Geiger, D., and Chickering, D. M. (1995). Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *Machine Learning*, 20(3):197–243. Available as Technical Report MSR-TR-94-09.
- Højsgaard, S. (2020). *gRain: Graphical Independence Networks*.
- Højsgaard, S., Dethlefsen, C., and Bowsher, C. (2020). *gRbase: A Package for Graphical Modelling in R*.
- Højsgaard, S., Edwards, D., and Lauritzen, S. (2012). *Graphical Models with R*. Use R! series. Springer.
- Holmes, D. E. and Jaim, L. C. (2008). *Innovations in Bayesian Networks: Theory and Applications*. Springer.
- Ide, J. S. and Cozman, F. G. (2002). Random Generation of Bayesian Networks. In *Proceedings of the 16th Brazilian Symposium on Artificial Intelligence (SBIA)*, pages 366–375. Springer-Verlag.
- Johnson, N. L., Kemp, A. W., and Kotz, S. (2005). *Univariate Discrete Distributions*. Wiley, 3rd edition.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, volume 1. Wiley, 2nd edition.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. Wiley, 2nd edition.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. Wiley.
- Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., and Bühlmann, P. (2012). Causal Inference Using Graphical Models with the R Package pcalg. *Journal of Statistical Software*, 47(11):1–26.
- Kenett, R. S., Perruca, G., and Salini, S. (2012). *Modern Analysis of Customer Surveys: With Applications Using R*, chapter 11. Wiley.
- Kjærulf, U. B. and Madsen, A. L. (2013). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Springer, 2nd edition.
- Koivisto, M. and Sood, K. (2004). Exact Bayesian Structure Discovery in Bayesian Networks. *Journal of Machine Learning Research*, 5:549–573.
- Koller, D. and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.
- Korb, K. B. and Nicholson, A. E. (2011). *Bayesian Artificial Intelligence*. Chapman and Hall, 2nd edition.
- Koski, T. and Noble, J. M. (2009). *Bayesian Networks: An Introduction*. Wiley.
- Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous Multivariate Distributions*. Wiley, 2nd edition.
- Kratzer, G. and Furrer, R. (2019). *mcmcabn: a Structural MCMC Sampler for DAGs Learned from Observed Systemic Datasets*.

- Kratzer, G., Pittavino, M., Lewis, F. I., and Furrer, R. (2019). *abn: an R Package for Modelling Multivariate Data Using Additive Bayesian Networks*.
- Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.
- Kucukelbir, A., Ranganath, R., Gelman, A., and Blei, D. (2015). Automatic Variational Inference in Stan. In *Advances in Neural Information Processing Systems 28*, pages 568–576.
- Kuipers, J. and Moffa, G. (2017). Partition MCMC for Inference on Acyclic Digraphs. *Journal of the American Statistical Association*, 112(517):282–299.
- Kuipers, J., Moffa, G., and Heckerman, D. (2014). Addendum on the Scoring of Gaussian Directed Acyclic Graphical Models. *The Annals of Statistics*, 42(4):1689–1691.
- Kulinskaya, E., Morgenthaler, S., and Staudte, R. G. (2008). *Meta Analysis: A Guide to Calibrating and Combining Statistical Evidence*. Wiley.
- Kullback, S. (1968). *Information Theory and Statistics*. Dover Publications.
- Larrañaga, P., Sierra, B., Gallego, M. J., Michelena, M. J., and Picaza, J. M. (1997). Learning Bayesian Networks by Genetic Algorithms: A Case Study in the Prediction of Survival in Malignant Skin Melanoma. In *Proceedings of the 6th Conference on Artificial Intelligence in Medicine in Europe (AIME'97)*, pages 261–272. Springer, Berlin.
- Lauritzen, S. (1996). *Graphical Models*. Oxford University Press.
- Loève, M. (1977). *Probability Theory*. Springer-Verlag, 4th edition.
- Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS Project: Evolution, Critique and Future Directions (with discussion). *Statistics in Medicine*, 28(5):3049–3067.
- Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS—A Bayesian Modelling Framework: Concepts, Structure, and Extensibility. *Statistics and Computing*, 10(4):325–337.
- Mahalanobis, P. C. (1936). On the Generalised Distance in Statistics. *Proceedings of the National Institute of Sciences of India*, 2(1):49–55.
- Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press.
- Margaritis, D. (2003). *Learning Bayesian Network Model Structure from Data*. PhD thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA. Available as Technical Report CMU-CS-03-153.
- McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall.
- Melançon, G., Dutour, I., and Bousquet-Mélou, M. (2001). Random Generation of Directed Acyclic Graphs. *Electronic Notes in Discrete Mathematics*, 10:202–207.
- Morota, G., Valente, B. D., Rosa, G. J. M., Weigel, K. A., and Gianola, D. (2012). An Assessment of Linkage Disequilibrium in Holstein Cattle Using a Bayesian Network. *Journal of Animal Breeding and Genetics*, 129(6):474–487.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.
- Nagarajan, R., Scutari, M., and Lèbre, S. (2013). *Bayesian Networks in R with Applications in Systems Biology*. Use R! series. Springer.

- Neapolitan, R. E. (2003). *Learning Bayesian Networks*. Prentice Hall.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2nd edition.
- Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.
- Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC)*, pages 1–10.
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, 6(1):7–11.
- Pourret, O., Naim, P., and Marcot, B. (2008). *Bayesian Networks: A Practical Guide to Applications*. Wiley.
- Robert, C. P. and Casella, G. (2009). *Introducing Monte Carlo Methods with R*. Springer.
- Ross, S. (2018). *A First Course in Probability*. Prentice Hall, 10th edition.
- Russell, S. J. and Norvig, P. (2009). *Artificial Intelligence: A Modern Approach*. Prentice Hall, 3rd edition.
- Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A., and Nolan, G. P. (2005). Causal Protein-Signaling Networks Derived from Multiparameter Single-Cell Data. *Science*, 308(5721):523–529.
- Schäfer, J., Opgen-Rhein, R., Zuber, V., Ahdesmäki, M., Silva, A. P. D., and Strimmer, K. (2017). *corpcor: Efficient Estimation of Covariance and (Partial) Correlation*.
- Scutari, M. (2010). Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software*, 35(3):1–22.
- Scutari, M. (2016). An Empirical-Bayes Score for Discrete Bayesian Networks. *Journal of Machine Learning Research (Proceedings Track, PGM 2016)*, 52:438–448.
- Scutari, M. (2020). Bayesian Network Models for Incomplete and Dynamic Data. *Statistica Neerlandica*. In print.
- Scutari, M. and Brogini, A. (2012). Bayesian Network Structure Learning with Permutation Tests. *Communications in Statistics – Theory and Methods*, 41(16–17):3233–3243.
- Scutari, M., Graafland, C. E., and Gutiérrez, J. M. (2019). Who Learns Better Bayesian Network Structures: Accuracy and Speed of Structure Learning Algorithms. *International Journal of Approximate Reasoning*, 115:235–253.
- Scutari, M. and Nagarajan, R. (2013). On Identifying Significant Edges in Graphical Models of Molecular Networks. *Artificial Intelligence in Medicine*, 57(3):207–217.
- Shäfer, J. and Strimmer, K. (2005). A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology*, 4:32.
- Sinoquet, C. and Mourad, R. (2013). *Probabilistic Graphical Models for Genetics, Genomics, and Postgenomics*. Oxford University Press.

- Spector, P. (2009). *Data Manipulation with R*. Springer-Verlag.
- Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. MIT Press, 2nd edition.
- Stan Development Team (2020a). *PyStan: the Python interface to Stan*.
- Stan Development Team (2020b). *RStan: the R interface to Stan*.
- Sucar, L. E. (2015). *Probabilistic Graphical Models: Principles and Applications*. Springer.
- Suzuki, J. (2017). An Efficient Bayesian Network Structure Learning Strategy. *New Generation Computing*, 35(1):105–124.
- Tsamardinos, I., Aliferis, C. F., and Statnikov, A. (2003). Algorithms for Large Scale Markov Blanket Discovery. In *Proceedings of the 16th International Florida Artificial Intelligence Research Society Conference*, pages 376–381. AAAI Press.
- Tsamardinos, I. and Borboudakis, G. (2010). Permutation Testing Improves Bayesian Network Learning. In Balcázar, J., Bonchi, F., Gionis, A., and Sebag, M., editors, *Machine Learning and Knowledge Discovery in Databases*, pages 322–337. Springer.
- Tsamardinos, I., Brown, L. E., and Aliferis, C. F. (2006). The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. *Machine Learning*, 65(1):31–78.
- Ueno, M. (2011). Robust Learning of Bayesian Networks for Prior Belief. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence*, pages 698–707.
- Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, 4th edition.
- Verma, T. S. and Pearl, J. (1991). Equivalence and Synthesis of Causal Models. *Uncertainty in Artificial Intelligence*, 6:255–268.
- Villa-Vialaneix, N., Liaubet, L., Laurent, T., Cherel, P., Gamot, A., and SanCristobal, M. (2013). The Structure of a Gene Co-Expression Network Reveals Biological Functions Underlying eQTLs. *PLOS ONE*, 8(4):e60045.
- Wang, K.-J., Chen, J.-L., and Wang, K.-M. (2019). Medical Expenditure Estimation by Bayesian Network for Lung Cancer Patients at Different Severity Stages. *Computers in Biology and Medicine*, 106:97–105.
- Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley.
- Yaramakala, S. and Margaritis, D. (2005). Speculative Markov Blanket Discovery for Optimal Feature Selection. In *Proceedings of the 5th IEEE International Conference on Data Mining*, pages 809–812. IEEE Computer Society, Los Alamitos, CA.
- Yu, J., Smith, V. A., Wang, P. P., Hartemink, A. J., and Jarvis, E. D. (2004). Advances to Bayesian Network Inference for Generating Causal Networks from Observational Biological Data. *Bioinformatics*, 20(18):3594–3603.
- Yuan, C. and Druzdzel, M. J. (2003). An Importance Sampling Algorithm Based on Evidence Pre-Propagation. In *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence*, pages 624–631.
- Zou, M. and Conzen, S. D. (2005). A New Dynamic Bayesian Network (DBN) Approach for Identifying Gene Regulatory Networks from Time Course Microarray Data. *Bioinformatics*, 21(1):71–79.