

TU-L0020 - Statistical Research Methods in Industrial Engineering and Management
Spring 2016
3-6 Credits
In English

1 COURSE STAFF AND CONTACT INFORMATION

Dr. Mikko Rönkkö Course instructor mikko.ronkko@aalto.fi

- All course exercises are available and returned through MyCourses.
- All articles and other reading material is available through Zotero.

2 OVERVIEW OF THE COURSE

The goal is to develop an understanding of how multivariate statistical methods are used in management research and how results are usually presented in journal articles. The course is designed for both those interested in just reading and understanding research done with statistical methods and for those who already use or plan to use statistical research methods in their own work.

During the course we will go through six empirical papers published in *Academy of Management Journal* and *Strategic Management Journal*, and analyze how these papers were done. The methods and research designs used in these papers cover a majority of methods and designs used in these journals. If you master all the methods and techniques presented during the course, you would be ideally capable to publish in these journals – if you have already a new theoretical insight and access to appropriate data.

The number of credits varies between 3-6 depending on which assignments students choose to complete. Assignments can be done independently or during the five optional computer exercise sessions. The content of each course component is explained later in the course brochure.

Credits	Content
3	Pre-exam, course lectures, readings and pre-assignments for lectures 1, 2, 4, and 6, and learning diary. You can miss one lecture, but must complete all readings and pre-lecture assignments on time unless otherwise agreed.
+0.5 each	Pre-assignments and readings for lectures 3 and 5.
+0.4 each	Data analysis exercises for classes 1-5
	Computer exercise sessions are organized so that students can work together on the exercises and receive guidance. These sessions will not be graded and participation will not be recorded.

The number of credits is rounded down.

3 WHO CAN PARTICIPATE

This course is targeted to industrial engineering and management doctoral students. A prerequisite for the course is TU-L0000 - Research Methods in Industrial Engineering and Management. Faculty or students from other departments or universities are admitted if space permits.

4 LEARNING OUTCOMES

The main goal of the course is to provide a foundation that enables independent self-study of quantitative research methods. Rather than addressing how the methods are used, we focus on why certain methods are used and why these methods work. Completing the 3 credit base module will introduce you to the logic of supporting causal claims with quantitative analyses, claiming construct validity, simple hypothesis testing with linear regression analysis and its extensions, and introduce you to more advanced techniques of structural equation modeling and multilevel modeling.

The optional pre-assignment and reading module for lectures 3 and 5 will introduce you to the logic of multilevel modeling and structural equation modeling with the purpose of providing a foundation for additional self-study of these topics.

The five optional data analysis exercises will introduce you to using statistical software, data management and structuring a data-analysis project, as well as to the workflows of statistical analysis using the methods discussed during the lectures.

5 WORKLOAD

The workload calculation below is for the full 6 credit version of the course¹.

Content	Units	Workload	Hours
Pre-exam			10
Lectures	6	3 hours each	18
Lecture related self-study, incl. learning diary	6	3 hours each	18
Empirical articles (easy reading)	141	7 h / 100 pages	10
Methodological literature (challenging reading)	668	10 h / 100 pages	67
Pre-lecture assignments	6	2 hours each	12
Data analysis exercises	5	8 hours each	40
Total hours			175
Credits			6.5

¹ Asko Karjalainen, Katariina Alha, and Suvi Jutila, *Anna Aikaa Ajatella: Suomalaisten Yliopisto-Opintojen Mitoitusjärjestelmä* (Oulun yliopisto, opetuksen kehittämissyksikkö, 2007).

6 COURSE CONTENT

The course consist of pre-exam, learning diary, six lectures with readings and pre-assignments, and optional analysis exercises. Computer exercises are voluntary sessions where you can come and work on the assignments and get support from the course instructor.

All assignments and exercises are distributed and returned through MyCourses. The enrolment key is **srm2016**.

<https://mycourses.aalto.fi/enrol/index.php?id=11288>

6.1 Pre-exam

The pre-exam is an oral exam, and you need to pass by the second lecture to pass the course. The exam takes about 30 minutes, and you can schedule your time by emailing the course instructor. You can retake the exam as many times as needed to pass it, but you cannot retake the exam to improve your grade.

The exam does not have a fixed set of materials that you need to study. Instead, you are provided a list of questions that you need to answer in the exam with references to articles and books that will provide answers to the questions. You can choose which material to study and which not, but please make sure that you can answer all the questions. You can have your notes with you in the exam. If you have difficulties in obtaining some material that you would like to study, email the course instructor.

The pre-exam questions are available on MyCourses.

6.2 Learning diary

The learning diary is for you to check that you have learned the key concepts and principles covered in the course. The list of questions is provided to you on the first lecture and you will work on the answers independently or in small groups if you want. You should work on the learning diary document soon after each class to write down how you understood the content of the classes and lecture materials. The final version of the learning diary is returned one week after the last lecture.

6.3 Lectures and pre-lecture assignments

The core of the course consists of six lectures and pre-lecture assignments every other week. You can miss one lecture, but must complete all pre-lecture assignments on time.

We will have a lunch break in the middle of each 4 hour lecture. The course has reserved the meeting room 1593 for lunch on the lecture days so those who want to have lunch together and discuss the course content or other research related matters together can do so.

The pre-lecture assignments consist of reading methodological literature and empirical papers and then doing an assignment where you apply the methodological ideas that you just studied to analyze the empirical papers. You will be provided individual feedback on the pre-lecture assignments and these are discussed in class.

The pre-lecture assignments for lectures 3 and 5 are optional and completing these will result in one additional credit.

6.3.1 Lecture 1: Course introduction, causal inference, and basics of linear regression model

This lecture introduces the course content, the principles of causal inference, and basics of linear regression models.

Required readings:

Singleton, R. A. J., & Straits, B. C. (2009). *Approaches to Social Research* (5 edition.). New York: Oxford University Press. (Chapter 4: Elements of Research Design)

Aguinis, H., & Vandenberg, R. J. (2014). An Ounce of Prevention Is Worth a Pound of Cure: Improving Research Quality Before Data Collection. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 569–595. doi:10.1146/annurev-orgpsych-031413-091231

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. doi:10.1016/j.leaqua.2010.10.010

Hekman, D. R., Aquino, K., Owens, B. P., Mitchell, T. R., Schilpzand, P., & Leavitt, K. (2010). An Examination of Whether and How Racial and Gender Biases Influence Customer Satisfaction. *Academy of Management Journal*, 53(2), 238–264. doi:10.5465/AMJ.2010.49388763

Deephouse, D. L. (1999). To be different, or to be the same? It's a question (and theory) of strategic balance. *Strategic Management Journal*, 20(2), 147–166. doi:10.1002/(SICI)1097-0266(199902)20:2<147::AID-SMJ11>3.0.CO;2-Q

6.3.2 Lecture 2: Further applications and extensions of linear regression model

This lecture continues from the previous lecture with additional issues in linear regression models. We discuss the use of linear regression to estimate mediation and moderation models. The lecture concludes with generalized linear model, which is an extension to linear regression covering most commonly used single dependent variable models as special cases (e.g. logistic regression, poisson regression, tobit regression, etc.).

Required readings:

Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach* (4th ed.). Mason, OH: South Western, Cengage Learning. (Chapters 2-4, 6.1-6.3)

Hekman, D. R., Aquino, K., Owens, B. P., Mitchell, T. R., Schilpzand, P., & Leavitt, K. (2010). An Examination of Whether and How Racial and Gender Biases Influence Customer Satisfaction. *Academy of Management Journal*, 53(2), 238–264. doi:10.5465/AMJ.2010.49388763

Deephouse, D. L. (1999). To be different, or to be the same? It's a question (and theory) of strategic balance. *Strategic Management Journal*, 20(2), 147–166. doi:10.1002/(SICI)1097-0266(199902)20:2<147::AID-SMJ11>3.0.CO;2-Q

Kraimer, M. L., Shaffer, M. A., Harrison, D. A., & Ren, H. (2012). No Place Like Home? An Identity Strain Perspective on Repatriate Turnover. *Academy of Management Journal*, 55(2), 399–420. doi:10.5465/amj.2009.0644

6.3.3 Lecture 3: Clustered and longitudinal data

Independence of observations is one of the key assumptions of linear regression. However, this assumption is violated when working with longitudinal or clustered data (i.e. members in teams). In this lecture we introduce random and fixed effects estimators from the econometrics tradition for analyzing these types of data and mixed/random/multilevel models from the modeling tradition.

Required readings:

Hausknecht, J. P., Hiller, N. J., & Vance, R. J. (2008). Work-Unit Absenteeism: Effects of Satisfaction, Commitment, Labor Market Conditions, and Time. *Academy of Management Journal*, 51(6), 1223–1245. doi:10.5465/AMJ.2008.35733022

The following readings and pre-lecture assignment are optional for extra credit:

Luke, D. A. (2004). *Multilevel Modeling*. Sage. (Chapters 1, 2)

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. doi:10.1016/j.leaqua.2010.10.010

Deephouse, D. L. (1999). To be different, or to be the same? It's a question (and theory) of strategic balance. *Strategic Management Journal*, 20(2), 147–166.

Kraimer, M. L., Shaffer, M. A., Harrison, D. A., & Ren, H. (2012). No Place Like Home? An Identity Strain Perspective on Repatriate Turnover. *Academy of Management Journal*, 55(2), 399–420. doi:10.5465/amj.2009.0644

6.3.4 Lecture 4: Measurement and introduction to factor analysis

This lecture discusses the concept of measurement, which arises from the efforts to quantify abstract concepts such as innovativeness. We discuss reliability and validity and present two measurement theories. The lecture introduces factor analysis, which can be used to provide support that different indicators may measure the same construct, and reliability statistics that can be used once unidimensionality has been established with factor analysis.

Required readings:

Singleton, R.A.J. & Straits, B.C., 2005. *Approaches to Social Research* 4th ed., New York: Oxford University Press. (Chapter 5: Measurement. Other editions will do as well)

DeVellis, R. F. (2003). *Scale development theory and applications*. Thousand Oaks: Sage. (Chapters 2-4 and 6, third edition of the book will also do)

Yli-Renko, H., Autio, E., & Sapienza, H. J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, 22(6-7), 587–613. doi:10.1002/smj.183

Mesquita, L. F., & Lazzarini, S. G. (2008). Horizontal and Vertical Relationships in Developing Economies: Implications for Smes' Access to Global Markets. *Academy of Management Journal*, 51(2), 359–380

6.3.5 Lecture 5: Latent variable models and simultaneous equations models

This lecture continues from the previous lecture by introducing confirmatory factor analysis. From there, we generalize to structural equation modeling with latent variables. Structural equation models with observed variables are discussed in the context of mediation models and instrumental variable models.

Readings and pre-lecture assignment are optional for extra credit:

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press. (Chapters 1, 5-8, 153 pages)

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. doi:10.1016/j.leaqua.2010.10.010

Mesquita, L. F., & Lazzarini, S. G. (2008). Horizontal and Vertical Relationships in Developing Economies: Implications for Smes' Access to Global Markets. *Academy of Management Journal*, 51(2), 359–380.

Yli-Renko, H., Autio, E., & Sapienza, H. J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, 22(6-7), 587–613. doi:10.1002/smj.183

6.3.6 Lecture 6: Research design and course conclusion

In this lecture we will address research design, reporting, as well as a miscellaneous topics related to data and analysis. We end with a summary of the course lectures.

Required readings:

Colquitt, J. A., & George, G. (2011). Publishing in AMJ—Part 1: Topic Choice. *Academy of Management Journal*, 54(3), 432–435. doi:10.5465/AMJ.2011.61965960

Bono, J. E., & McNamara, G. (2011). Publishing in AMJ—Part 2: Research Design. *Academy of Management Journal*, 54(4), 657–660. doi:10.5465/AMJ.2011.64869103

Grant, A. M., & Pollock, T. G. (2011). Publishing in AMJ—Part 3: Setting the Hook. *Academy of Management Journal*, 54(5), 873–879. doi:10.5465/amj.2011.4000

Sparrowe, R. T., & Mayer, K. J. (2011). Publishing in AMJ—Part 4: Grounding Hypotheses. *Academy of Management Journal*, 54(6), 1098–1102. doi:10.5465/amj.2011.4001

Zhang, Y. (Anthea), & Shaw, J. D. (2012). Publishing in AMJ—Part 5: Crafting the Methods and Results. *Academy of Management Journal*, 55(1), 8–12. doi:10.5465/amj.2012.4001

Geletkanycz, M., & Tepper, B. J. (2012). Publishing in AMJ—Part 6: Discussing the Implications. *Academy of Management Journal*, 55(2), 256–260. doi:10.5465/amj.2012.4002

6.4 Data-analysis exercises (optional)

Each of the lectures 1 to 5 has a data analysis exercise that can be completed after the lecture. In each of the exercises, you will conduct a small data analysis project using the techniques presented in the lecture with either a dataset provided by the course instructor or using your own dataset. You should submit a free form document explaining the analyses you did, the thought process that lead to the analyses, and how you interpreted the results.

All returned data analysis exercises will receive individual feedback.

6.5 Computer exercise sessions (optional)

The computer exercise sessions are optional and not graded. These sessions take place in the computer class on every other week. The sessions typically start with small demos of the topics of the previous lecture.

In the data-analysis demos we will go through the analyses required for the data-analysis assignment for the week. You will implement the analyses with the computers in the class (or their own laptops). The results files and analysis logs will be sufficient to complete the data-analysis exercise for the week. After the class you should not need to do additional analyses, but simply write a document explaining the analyses and interpret the results.

In addition to working with data, we will do “hand calculations” and simulations. “hand calculation” refers to calculating statistical analyses without statistical software. Most statistical estimation involve minimizing or maximizing an estimation function. For example, least squares estimator minimizes the sum of squares of prediction errors and maximum likelihood estimator maximizes the likelihood of the data given a hypothesized model. During these exercises we specifying the estimation functions in Excel and estimating the model by minimizing or maximizing this function with the Solver tool in Excel. The purpose of these assignments is to make you understand how the analysis tools work in practice. While you are unlikely to encounter problems with linear regression, more advanced modeling techniques may not always work well or you could get non-sense results. In these scenarios understanding what the analysis software actually does is very important when troubleshooting the analysis.

Monte Carlo simulation refers to an analysis where you generate many samples of data (e.g. 1000 samples) from a known model and then apply a statistical technique to each sample separately to see how an analysis method performs over repeated samples. This type of simulations can be very valuable for a researcher just applying the methods. (See the talk by John Rauser.) First, by simulating datasets you open up an entirely new way to teach yourself statistical analyses. For example, if you do not understand how heteroskedasticity affects linear regression, you can just simulate a number of datasets from a model with and without heteroskedasticity and compare the results yourself. Second, you can verify that you have understood an analysis technique correctly by applying the technique to simulated datasets to see that you indeed get the correct results.

The recommended software to use are Stata and R, but it is also possible (although not recommended) to do most of the exercises in SPSS. If you plan to do the course assignments using R, it is highly recommended that you do a tutorial (e.g. <https://www.datacamp.com>) on R before the course starts.

7 SCHEDULE

Date	Week	Day	Time	Location	Topic
7 Jan 16	1	Thu	10:00- 14:00	1021- 1022	Lecture 1: Course introduction, causal inference, and basics of linear regression model
14 Jan 16	2	Thu	10:00- 14:00	1174- 1176	Computer exercise 1 (optional)
21 Jan 16	3	Thu	10:00- 14:00	1021- 1022	Lecture 2: Further applications and extensions of linear regression model
28 Jan 16	4	Thu	12:00- 16:00	1174- 1176	Computer exercise 2 (optional)
4 Feb 16	5	Thu	12:00- 16:00	1021- 1022	Lecture 3: Clustered and longitudinal data
11 Feb 16	6	Thu	10:00- 14:00	1174- 1176	Computer exercise 3 (optional)
18 Feb 16	6	Thu	10:00- 14:00	1021- 1022	Lecture 4: Measurement and introduction to factor analysis
25 Feb 16	8	Wed	10:00- 14:00	1174- 1176	Computer exercise 4 (optional)
3 Mar 16	9	Thu	10:00- 14:00	1021- 1022	Lecture 5: Latent variable models and simultaneous equations models
10 Mar 16	10	Thu	10:00- 14:00	1174- 1176	Computer exercise 5 (optional)
	11				No lecture
24 Mar 16	12	Thu	10:00- 14:00	1021- 1022	Lecture 6: Research design and course conclusion

8 GRADING

All submitted work will be graded between 1-5 and your grade will be a weighted average of the parts of the course that you completed.

Course part	Weight	Notes
Pre-exam	1	
Lecture participation	0.5 each (total 3)	By default, you will receive 2 for being present and your grade will increase based on your lecture participation
Learning diary	3	
Pre-lecture exercises	1 each (total 6)	
Data-analysis exercises	1 each (total 5)	
Computer exercise sessions	N/A	Not graded

9 COURSE MATERIAL

The reading materials for the course are distributed through the Zotero reference management system. To get access to the materials:

- 1 Create an user account at [Zotero.org](https://www.zotero.org)
- 2 Email your username to the course instructor
- 3 The course instructor will send you an invitation to a group library, which you need to accept.

After you have accepted the invitation, you can access the material either [online](#) with a web browser or by installing the Zotero software on your computer.

Slides will be available before each lecture.

9.1 Books

DeVellis, R. F. (2003). *Scale development theory and applications*. Thousand Oaks: Sage. (Chapters 2-4, and 6, third edition of the book will also do: 80 pages)

Luke, D. A. (2004). *Multilevel Modeling*. Sage. (Chapters 1-2: 52 pages)

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press. (Chapters 1, 5-8, 153 pages)

Singleton, R. A. J., & Straits, B. C. (2009). *Approaches to Social Research* (5 edition.). New York: Oxford University Press. (Chapter 4: Elements of Research Design, Chapter 5: Measurement: 72 pages)

Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach* (4th ed.). Mason, OH: South Western, Cengage Learning. (Chapters 2-4, 6.1-6.3, 13-14: 226 pages)

577 pages total

9.2 Articles

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. doi:10.1016/j.leaqua.2010.10.010

Video: <https://www.youtube.com/watch?v=dLuTjoYmfXs> (32:19)

Aguinis, H., & Vandenberg, R. J. (2014). An Ounce of Prevention Is Worth a Pound of Cure: Improving Research Quality Before Data Collection. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 569–595. doi:10.1146/annurev-orgpsych-031413-091231

Video: <https://www.youtube.com/watch?v=y1tNSEXh9Gk> (18:49)

Colquitt, J. A., & George, G. (2011). Publishing in AMJ—Part 1: Topic Choice. *Academy of Management Journal*, 54(3), 432–435. doi:10.5465/AMJ.2011.61965960

Bono, J. E., & McNamara, G. (2011). Publishing in AMJ—Part 2: Research Design. *Academy of Management Journal*, 54(4), 657–660. doi:10.5465/AMJ.2011.64869103

Grant, A. M., & Pollock, T. G. (2011). Publishing in AMJ—Part 3: Setting the Hook. *Academy of Management Journal*, 54(5), 873–879. doi:10.5465/amj.2011.4000

Sparrowe, R. T., & Mayer, K. J. (2011). Publishing in AMJ—Part 4: Grounding Hypotheses. *Academy of Management Journal*, 54(6), 1098–1102. doi:10.5465/amj.2011.4001

Zhang, Y. (Anthea), & Shaw, J. D. (2012). Publishing in AMJ—Part 5: Crafting the Methods and Results. *Academy of Management Journal*, 55(1), 8–12. doi:10.5465/amj.2012.4001

Geletkanycz, M., & Tepper, B. J. (2012). Publishing in AMJ—Part 6: Discussing the Implications. *Academy of Management Journal*, 55(2), 256–260. doi:10.5465/amj.2012.4002

91 pages total

9.3 Empirical articles used as examples

Deephouse, D. L. (1999). To be different, or to be the same? It's a question (and theory) of strategic balance. *Strategic Management Journal*, 20(2), 147–166. doi:10.1002/(SICI)1097-0266(199902)20:2<147::AID-SMJ11>3.0.CO;2-Q

Hausknecht, J. P., Hiller, N. J., & Vance, R. J. (2008). Work-Unit Absenteeism: Effects of Satisfaction, Commitment, Labor Market Conditions, and Time. *Academy of Management Journal*, 51(6), 1223–1245. doi:10.5465/AMJ.2008.35733022

Hekman, D. R., Aquino, K., Owens, B. P., Mitchell, T. R., Schilpzand, P., & Leavitt, K. (2010). An Examination of Whether and How Racial and Gender Biases Influence Customer

Satisfaction. *Academy of Management Journal*, 53(2), 238–264.
doi:10.5465/AMJ.2010.49388763 (AMJ best paper winner for 2010)

Video: <https://www.youtube.com/watch?v=NanlRhlHR8> (1:40)

Mesquita, L. F., & Lazzarini, S. G. (2008). Horizontal and Vertical Relationships in Developing Economies: Implications for Smes' Access to Global Markets. *Academy of Management Journal*, 51(2), 359–380.

Kraimer, M. L., Shaffer, M. A., Harrison, D. A., & Ren, H. (2012). No Place Like Home? An Identity Strain Perspective on Repatriate Turnover. *Academy of Management Journal*, 55(2), 399–420. doi:10.5465/amj.2009.0644

Yli-Renko, H., Autio, E., & Sapienza, H. J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, 22(6-7), 587–613. doi:10.1002/smj.183

141 pages total

9.4 Other material

Rauser, J. (2014, October 15). *Statistics Without the Agonizing Pain*. Presented at the Big Data Conference - Strata + Hadoop World, New York, NY. Retrieved from <http://strataconf.com/stratany2014/public/schedule/detail/37554>

Video: <https://www.youtube.com/watch?v=5Dnw46eC-0o> (11:47)