IE 433 Regression and Design of Experiments
Penn State University, WEB course

Instructor: Professor Enrique del Castillo, IME Department, Penn State, e-mail: exd13@psu.edu. **If you need to contact me, please use the internal CANVAS e-mail, as any other email may get lost**. Personal web page: http://www2.ie.psu.edu/Castillo/Personal Web Page/Castillo.htm

Course homepage: this course will use the CANVAS system. All video lessons and other additional class material will be posted there.

Course objective: the aim of this course is to provide an introduction to regression and experimental design techniques. Emphasis will be given to the application of these techniques to model and improving production processes in industry. Material learned in this class will prepare successful students beyond what is called a "**six sigma green belt**" in industry, although the treatment of topics is at a deeper level. The course is hence required for the IE minor in "Six-sigma".

Class meetings: The class will run through all 12 weeks of summer. Weekly video lectures and class handouts will be posted in CANVAS. Some of the class videos were recorded prior to 2017 when the university use to have a web-site system called "ANGEL", which was substituted in 2017 by CANVAS. Hence, if in any of the videos you hear me referring to "the ANGEL web site", it means "the CANVAS web site".

Prerequisite: IE 323 (Statistical Methods in IE). Without exceptions, you cannot take IE 323 and IE 433 at the same time. Equivalent coursework at PSU and elsewhere will be considered as appropriate equivalences on a case-by-case basis (you must consult with me for approval). I will also assume you know basic Matrix computations (i.e., multiplication and inversion of matrices). See below (page 3) for a more detailed list of concepts you need to know.

Textbook: I have arranged with John Wiley & Sons Inc. to setup a custom made book (**available through the PSU bookstore**) consisting of selected chapters from each of the following two books:

- D.C. Montgomery, E. Peck, and G. Vining, *Introduction to linear regression analysis*, **5th ed.**, Wiley 2012, and
- D.C. Montgomery *Design and Analysis of Experiments,* **8th ed.**, Wiley, 2012.

The easiest way to get this custom book is through the PSU bookstore, although it can also be found in some places over the internet using the ISBN number 9781118546994 under the title "Regression and Design of Experiments for PSU-corrected" (author: Montgomery). Note it should be the "**corrected**" version). The book is also **on reserve at the PSU library.**

Grading System

- Homework (4 sets) optional and count as **extra credit**, 5% of your final grade per each complete turned in homework, up to 20% (see more details below). Due when indicated in the CANVAS web site (see also below). Homework will be graded only as completed/not completed, and will not be returned. NO resubmissions.
- Midterm exam: 50%. Due 24 hr. after posted (see below for dates). NO resubmissions.

- Final Exam (comprehensive): 50%. Due 24 hr. after posted (see below for dates). NO resubmissions.

Exams: the midterm and the final exam will be posted on CANVAS. You will have 24 hours to complete it and return your answers via CANVAS. You will be responsible to turn in clear scanned answers for both of your exams. Unclear or non-readable handwriting will receive no credit. No make-ups. See due dates below or on CANVAS.

Exams dates (all times ET—see CANVAS FOR ANY UPDATES)

Exam	Posted on:	Due on:
MIDTERM	JUNE 29 at 1pm	JUNE 30 at 1pm
FINAL	AUGUST 12 at 1pm	AUGUST 13 at 1pm

Homework: four homework sets consisting mainly (but not exclusively) of end-of chapter problems from the textbook will be assigned. I will not grade these in detail and these are **optional, for extra credit only**. If you turn them in, for each complete homework solutions you turn in I will award +5% of the final grade. "Complete" means you answer all exercises and your homework is clean and clear enough I can see what was done (partially answered homework will receive partial credit). Thus, if you turn in the 4 complete homework solutions, you will have assured a +20% of your final grade. In each homework you will have at least one week to complete them. You will be responsible to turn in clear documents, where it is easy to read your solutions. Clearly indicate the problem numbers. The homework does not need to be typed, but must be clearly scanned in a PDF file. Selected solutions will be posted.

Homework dates (all times ET—see CANVAS FOR ANY UPDATES)

HWK	Assigned on:	Due on:
Homework 1	Week 1 (May 18-24)	June 5 at 5pm
Homework 2	Week 4 (June 8-14)	June 23 at 5pm
Homework 3	Week 7 (June 29-July 5)	July 17 at 5pm
Homework 4	Week 10 (July 20-26)	August 7 at 5pm

Electronic submissions of homework and exams: all homework and exam work must be submitted via CANVAS. Note that **only one submission per student per exam/homework** is allowed (i.e., no re-submissions allowed). Be careful with your uploaded files, as I will **not** consider any resubmissions thereafter. **Late submissions**: penalized 20% per hour late, until 5 hours after. All posted due times will be EST.

Grading scale: 90-100% → A; 85-90% → A-; 80-85% → B+; 75-80% → B; 70-75% → B-; 65-70% → C+; 60-65% → C; 55-60% → D; 0-55% → F.

Tentative schedule and reading assignments

Week #	Topics	Reading Assignment
	PART I: Regression Analysis (readings from **Regression book**)	
1	Course overview, Simple linear regression, Matrix Algebra review;	Ch. 2, Matrix algebra handout
2	Multiple linear regression	3.1, 3.2 (except 3.2.6)
3	Hypothesis testing and interval estimation in regression	3.3 (except 3.3.4) to 3.5
4	Common modeling mistakes and problems in regression	3.8 to 3.11
5	Model diagnostics and transformations	4.1 to 4.3, 6.1 to 6.3, 5.1 to 5.3;
6	Indicator variables	8.1, 8.2
7	Variable selection and model building	10.1, 10.2
	PART II: Design of experiments (readings from **DOE book**)	
8	Experiments with one factor	3.1 to 3.6, 3.8
9	Factorial Designs	5.1 to 5.4
10	2^k designs: 2^2, 2^3, general 2^k	6.1 to 6.4
11	Unreplicated 2^k designs, blocking in 2^k's	6.5, 7.1 to 7.6
12	Fractional factorials part I: half and quarter fractions, resolution	8.1 to 8.3
12	Fractional factorial part II: general 2^{k-p}, blocking in 2^{k-p}'s	8.4, 8.5

Academic Integrity Policy

You must work in all the exams and homework *strictly individually*. Any student in the class involved in violations to this policy will get a 0 (zero) in the corresponding homework/exam and may be subject to further disciplinary action through the College Academic Integrity and Judicial Affairs Committee at the discretion of the instructor, in accordance to PSU Policy G-9.

Prerequisite concepts you need to know

- Random variables: definition, expectation and variance operators
- Distribution functions: Normal (and standard normal), t, F, χ^2, sampling distributions of the sample mean, the sample variance, and of the ratio of two sample variances.
- Test of hypothesis and confidence intervals for Normal populations: on a mean, a variance, differences between two means and between two variances
- Matrices: basic notation, multiplication, inversion