STUDENT’S NAME (please print):

Take-home MATH 282A Fall 2015

- Take-home exam—open book but please work alone. Prove all your statements; all sub-problems have equal weight.
- Part I amounts to 70%. Part II amounts to 30% (plus 10% extra credit).
- The exam is due Monday Dec 7 by noon (sharp!). Please leave it in a sealed envelope at the mailbox of Jie Chen (Ashley) at the 7th floor of APM. Keep a copy of your exam for your records just in case it is misplaced.

Part I: Work out the following problems from Ch. 4 of the book.
Set 4a: ex. 1; set 4b: ex. 1, 2, 3; set 4c: ex. 1, 3; misc. set: ex. 4.
[Note a mistake in the book. In set 4b, ex. 2 part (a), the n in the denominator should be inside the square root.]

Part II: Let \( \epsilon_i \sim \text{i.i.d. } N(0, \sigma^2) \), and consider the two models:

\[
Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \text{for } i = 1, \ldots, n \tag{1}
\]

and

\[
Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \quad \text{for } i = 1, \ldots, n \tag{2}
\]

Let \( \hat{\beta} \) be the LS estimator of the vector \( \beta = (\beta_0, \beta_1)' \) under model (1).

(a) Focus on estimating \( \beta_1 \). Is \( \hat{\beta}_1 \) unbiased when you fit model (1) while model (2) is true? Vice versa, what is the problem of estimating \( \beta_1 \) by fitting model (2) using LS when model (1) is true?

(b) Assume the design points are such that \( \bar{x} = n^{-1} \sum_{i=1}^{n} x_i = 0 \) so that the design of model (1) is orthogonal. Re-parametrize the problem of model (2) to an equivalent quadratic regression with orthogonal design, i.e., let

\[
Y_i = \beta_0 + \beta_1 x_i + \beta_2 p(x_i) + \epsilon_i, \quad \text{for } i = 1, \ldots, n \tag{3}
\]

where \( p(x) = x^2 + c_1 x + c_0 \) is a quadratic function with coefficients \( c_1, c_0 \) chosen such that the regression of model (3) has orthogonal design. Is there still a problem when estimating \( \beta_1 \) by fitting model (3) while model (1) is true?

(c) Assume straight line model (1), and suppose that you are actually designing the experiment, i.e., picking the \( x_i \) for \( i = 1, \ldots, n \), where the responses \( Y_i \) will be measured. Suppose that the \( x_i \) must lie in the interval \([-1, 1]\); assume \( \bar{x} = 0 \) and that \( n \) is an even number. How should you pick \( x_1, \ldots, x_n \) in order to minimize the variance of \( \hat{\beta}_1 \)?

(d) (For extra credit). Is the optimal design found in part (c) still a good design if the true model is instead the quadratic regression (2)? Explain.