# Midterm 2

**Instructions:**

1. The examination is due at 11:00 am at the beginning of lecture on **Thursday, April 19**.

2. For your solution for each problem, please use only one side of each piece of paper. You may write neatly or use a word processor. Please turn in **separate solutions** to the two problems. Include a cover page with your name for each solution. Please **staple together all pages for each solution**, but do not staple the two solutions together. Each separate solution should contain a **maximum of five pages, including computer output**, not including the cover page. Many excellent solutions will use **fewer than the maximum allotment of pages**.

3. Begin each solution with a short paragraph that summarizes the main results and refers to the statistical evidence in support of your conclusions. This summary paragraph should be a well written, self-contained, description of your results that makes sense to a reader who has read the problem. This summary paragraph should not include computer output, tables, or too many details of the analysis. Follow the summary paragraph with a concise description of the data analysis and your thinking and reasoning to justify the analysis. Any computer output should be incorporated into your discussion. Only include computer output that is directly relevent to support your main conclusions.

4. You may use any books and notes that you desire, but the work that you turn in **must be completely your own**. In particular, **you may not discuss the examination with anyone other than the instructor**. Any questions you have can only be addressed to me.

5. Strive to provide a solution that addresses the main scientific questions of interest, and use the simplest and most straightforward statistical methods that are sufficient to this task. It is much more preferable to provide a simple sensible statistical analysis than to attempt to demonstrate the full range of your mastery of the most complex statistical methods.

6. The total exam score is 100 points. Each problem is equally weighted.

7. A separate e-mail to the class will indicate my extra office hours for questions regarding the examination.

# Problem 1: River Nitrogen Abundance

The rise in the abundance of algae in coastal waters is thought to be due to increases in nutrients such as nitrates and other forms of nitrogen. One hypothesis is that excessive amounts of nitrate are due to human influence. Researchers collected the data set found in the file `river.txt` to assess the evidence that nitrate in the discharges of rivers throughout the world are associated with human population density. Human populations can affect nitrogen input into rivers through industrial and automobile emissions to the atmosphere followed by rainfall, through fertilizer runoff, through sewage discharge, and through watershed disturbance. The watershed is the region of land that drains into a river.

The variables in this data set for rivers identified by name and country/world region are:

1. `Discharge`: the estimated annual average discharge of the river into an ocean (in $m^3$/sec)
2. `Runoff`: the estimated annual average runoff from the watershed (in liters/(sec $\times$ $km^2$))
3. `Precipitation`: precipitation (in cm/year)
4. `Area`: the area of the watershed (in $km^2$)
5. `Density`: population density (in people/$km^2$)
6. `NO3`: river nitrate concentration (in $\mu$M/liter)
7. `NitrateConc`: nitrate concentration in wet precipitation at sites located near watersheds (in $\mu$mol $NO_3$/(sec $\times$ $km^2$)
8. `Deposition`: the product of precipitation and nitrate concentration

The response variable of interest is river nitrate concentration (`NO3`). `Discharge`, `Runoff`, `Area`, and `Precipitation` are characteristics of the rivers and surrounding area. `Deposition`, `NitrateConc`, and `Density` are associated with human activities.

(a) Without regarding human density (`Density`) directly, are there associations between deposition (`Deposition`) and wet precipitation nitrate concentration (`NitrateConc`) with river nitrate concentration (`NO3`), after accounting for discharge, runoff, precipitation, and watershed area? Build an appropriate model using some of these variables (possibly transformed), and interpret the results. Justify the model that you select.

(b) Build a model to estimate the effect that human density has directly on river nitrate concentration in addition its indirect effect through influence on other explanatory variables, considering as well the possible contributions of all other explanatory variables.

(c) Predict the change in river nitrate concentration in the Columbia river in the USA if the population density in the area were to increase 20% from 10 to 12, assuming that all other relevant explanatory variables remain unchanged.

# Problem 2: Pig Weight Gain

Researchers performed an experiment to evaluate how average daily weight gain (adg) of pigs is affected by five treatments. Each treatment was assigned at random to two pens, for a total of ten pens. Each pen contains four pigs, so there are forty individual pigs in total in the experiment. The five treatments are combinations the concentration of zinc (Zn) supplement added to the daily ration of food and floor space in the pen. The three levels of zinc are 30, 80, and 130 ppm (low, medium, and high). The two levels of floor space are 0.28 and 0.14 square meters per pig (large and small). The level 0.28 is considered to be adequate space for the pigs whereas the level 0.14 is anticipated to restrict growth. It is hoped that adding zinc to the daily food ration might counteract the effects of restricted space.

The five treatments are:

1. low Zn / small space
2. medium Zn / small space
3. high Zn / small space
4. medium Zn / large space
5. high Zn / large space

The data are shown in the table below.

| | | Average Daily Gain (grams) | | | |
|---|---|---|---|---|---|
| Treatment | Pen | Pig 1 | Pig 2 | Pig 3 | Pig 4 |
| 1 | 1 | 452 | 443 | 440 | 449 |
| 1 | 2 | 460 | 459 | 463 | 455 |
| 2 | 1 | 473 | 460 | 467 | 464 |
| 2 | 2 | 433 | 448 | 453 | 447 |
| 3 | 1 | 466 | 468 | 464 | 470 |
| 3 | 2 | 489 | 485 | 484 | 495 |
| 4 | 1 | 562 | 548 | 555 | 552 |
| 4 | 2 | 550 | 565 | 565 | 552 |
| 5 | 1 | 573 | 581 | 567 | 574 |
| 5 | 2 | 552 | 552 | 549 | 552 |

Analyze these data and address the following issues.

(a) Which treatment combination results in the largest average daily weight gain?

(b) Estimate the relationship between adg and Zn in small pens.

(c) Is there evidence of an interaction between the levels of Zn and the levels of floor space?

(d) In a future experiment with 80 pigs available using the same treatment combinations, would it be preferable in assessing the effect of Zn on adg to use eight pigs per pen (enlarging the pens so that the space per pig were the same) or twice as many pens with four pigs per pen? Justify your response.