

# STAT 358 T

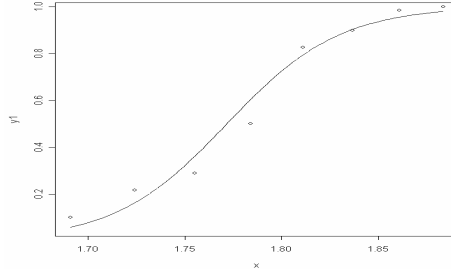
## Introduction to Biostatistics (Tutorial)

Fall Semester 2005

---

### INSTRUCTOR

Prof. Bernhard Klingenberg  
Bronfman Science Center 201  
e-mail: bklingen@williams.edu  
Phone: (413) 597 2467



---

### COURSE DESCRIPTION:

This tutorial covers statistical methods useful in the biological (and many other) sciences. We first review (and apply) methods of inference for continuous and categorical data and then cover topics such as multiple linear and logistic regression, analysis of time-to-event or survival data and analyses of repeated measures data.

---

### TEXTBOOKS:

**Regression Methods in Biostatistics.** Linear, Logistic, Survival and Repeated Measures Models. Vittinghoff, Glidden, Shiboski, McCulloch, Springer, 2005. (Required)

Supplemental Reading:

**An Introduction to Categorical Data Analysis,** Alan Agresti, Wiley, 1996. (On reserve in Schow library)

**Stats: Data and Models,** DeVeaux, Velleman, Bock, Addison Wesley, 2005. (On reserve)

Additionally, I will sometimes hand out copies of articles related to the topic of the week.

---

### WEEKLY DISCUSSIONS AND ASSIGNMENTS:

Each week every student will read the assigned pages, handouts or papers. Working in pairs, students will solve and write up the assigned examples. Then, each pair will get another pair's solution and critique it. Taking turns, one student from each pair will present a short summary of the weekly topic and the solutions to some examples in a weekly meeting of all students. At each meeting, I will also introduce and briefly discuss the next weekly topic.

---

### STATISTICAL COMPUTING

We are going to use the statistical software package R for our statistical analyses (although you can use any other package you are familiar with, i.e., you don't have to use R). R is a completely free software environment for statistical computing and graphics. To download R to your computer, go to <http://www.r-project.org>. The code with which above picture was created is on the other side. The best way to learn any statistical software package is to play around with it, so I encourage you to fool around! There is nothing you can break, and a few frozen screens are not bad (remember, though, to always save your data before trying something new).

## OUTLINE OF TOPICS

---

This is a very tentative schedule and we will add/delete as we go along:

Topic	Discussed where?	Week
Review of exploratory and descriptive methods for one, two and multivariable continuous and categorical outcomes. Getting familiar with the R package.	Pages 7 – 27. (e.g., chapt. 3 and 4 in DeVeaux).	1
Review of one and two sample methods for means and ANOVA. (Confidence intervals and hypotheses tests). Correlation Coefficients. Implementation in R	Pages 29 – 36 (e.g., chapt. 7, 23-25, 28 in DeVeaux)	2
Review of one and two sample methods for proportions. (Confidence intervals and hypotheses tests). Implementation in R	Chapters 19, 20, 22 in DeVeaux	3
Analysis of two-way contingency tables. Risk difference, Relative Risk and Odds Ratio. Chi-square tests of independence. Implementation in R	Pages 44 - 54 Agresti, pages 16-32	4
Linear Regression , Implementation in R	Pages 36 – 43 and 69 - 131	5/6
Predictor Selection in Regression, Implement. in R	Pages 132 - 156	7
Logistic Regression, Implementation in R	Pages 157 - 209	8/9
Survival Analysis, Implementation in R	Pages 211 - 251	10/11
Repeated Measures Analysis, Implemt. in R	Pages 253 - 288	12

## SAMPLE R-CODE

---

```
Beetles = read.table("C:\\Documents and Settings\\Williams user\\My Documents\\Stat  
358\\beetles.txt", header = TRUE)
```

```
beetles
```

```
attach(beetles)
```

```
p = Dead/Exposed
```

```
plot(Dose, p)
```

```
plot(Dose,log(p/(1-p)))          #on logit scale: should be roughly linear
```

```
fit1 = glm(Dead/Exposed ~ Dose, binomial, weights = Exposed, data = beetles)
```

```
summary(fit1)
```

```
x = seq(min(Dose), max(Dose), length=100)
```

```
y1 = predict(fit1, data.frame(Dose=x), type="response")
```

```
plot(x,y1,type="l")
```

```
points(Dose,p)
```