

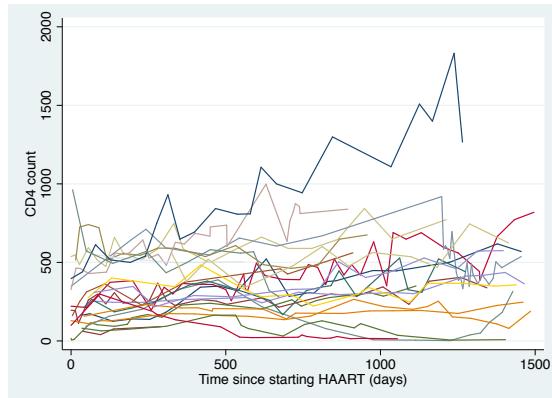


Longitudinal Data Analysis

Fall 2015

Chapter 1

Introduction



Instructor

Nick Jewell (jewell@berkeley.edu)

GSI

Robin Mejia (mejia@nasw.org)

Texts

- Nick, Alan Hubbard and Brianna Heggeseth are currently writing a book on the course material and drafts of chapters will be posted on *Bcourses*.
- The course will be managed through *Bcourses.berkeley.edu*

Course Description

- This course covers the statistical issues surrounding estimation of effects using data on subjects followed through time. The course emphasizes a regression model approach to disease incidence, continuous, binary and count outcome data.
- Background expected in statistical/mathematical material including regression (linear and logistic), basic understanding of statistical estimation and inference.

Assignments/Final Projects

- We will have approximately 6-7 assignments, and one final (group) project.
- Most assignments will involve computer analysis of data. Although the student can use any software they find convenient, STATA will be emphasized and taught.
- The final project should involve a data analysis of the student's choosing, to be presented at end of term (dates forthcoming).



What is longitudinal data?

Longitudinal Data

- Data collected (or inferred) at multiple different times on multiple units (e.g., person---compare with time series that often only involve one unit)—will need notation to keep track of this complication
- Focus will almost always focus on mean changes of the outcome variable and what variables “explains” such changes (but not always—see trajectory analyses).
- Outcome can be binary (disease yes/no), continuous (CD4 levels in an HIV-infected subject), or counts (number of diarrhea episodes in a time block).

Repeated Measures

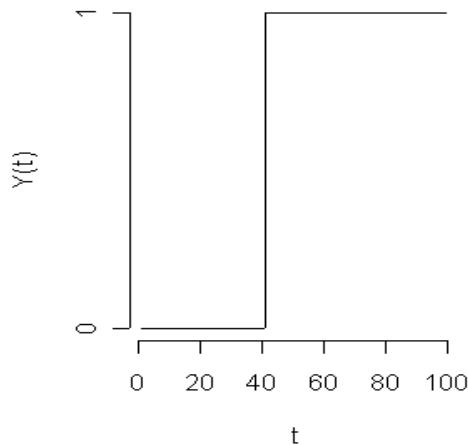
- Continuous outcome data: we will looking an models/methods that are extensions to simple linear regression models.
- Binary and count data using extensions of logistic and Poisson (log-linear) regression.

Longitudinal Data Reduced to Single Outcome

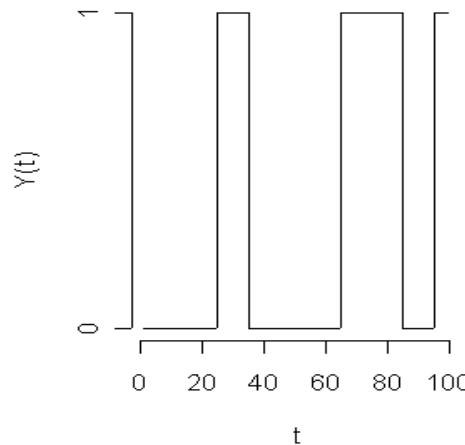
- Single event data using survival analysis.
 - Examples: time to death, time to tumor recurrence.
- Summaries of multiple event data.
 - Examples: the number of sex partners, number of seizures. Analysis techniques include Poisson and negative binomial regressions.
- These analyses can appear cross-sectional although they involve longitudinal data collection strategies and often introduce challenges to standard cross-sectional regression methods

Graphical examples of outcome types in longitudinal studies

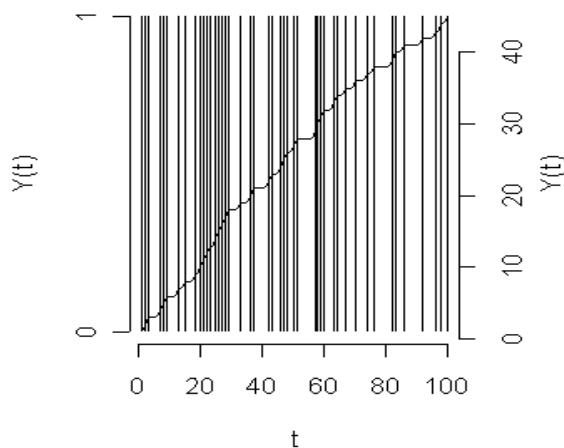
A Single Event



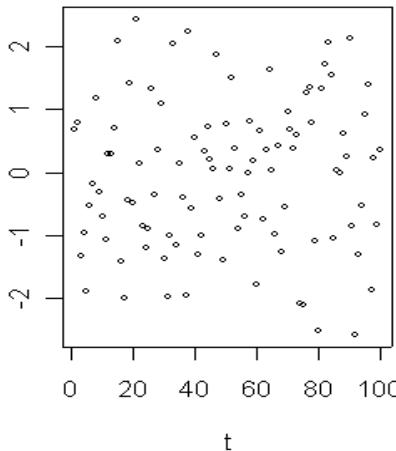
State transitions



Repeated Events



Repeated Cont. Outcomes



Possible hierarchies of units

- Time when the measurement if made (time-structured - longitudinal studies)
- Place (position, region) where the measurement is made (spatial data)
- Subunit on which the measurement is made (e.g., students within classrooms, “clustered” data)
- Combinations

More complicated examples

- Measurements over time on various subunits (eyes within subjects)
- Measurements on subunits of subunits (hierarchical data structures), e.g., eyes within subjects within siblings
- Often the repeated measurements over time and/or place and/or subunit may also occur under different conditions (different treatments, covariate values, etc.) - cross-over trials.

Time Scales: How often do you have to follow subjects?

1.6. Time Frames for Observing Longitudinal Data 11

Measure	Time interval	Changing?
Human height	Minutes through days	No
	Weeks	Perhaps newborns?
	Month(s)	Infants
	Year(s)	For children/adolescents
	Decades	Not sure: maybe for adults/seniors?
Human weight	Minutes through hour(s)	No
	Days	Maybe not, depending on accuracy of scale
	Week(s)	Yes, in weight-loss program
	Months	Certainly (kids)
Wage/salary	< 1 hour	Not for salaried workers
	Day	If on commission or piece-work
	Month/year	If on monthly pay plan
Housing cost (rent or mortgage) income	Annual	Yes
	Hourly	For businesses
Gender	Any	No, only need once
Race	Any	No, only need once
DNA	Any	No, only need once with exceptions
Blood pressure	Minutes	Potentially, depending on study goals
	Years	Same
3-month recall of illicit drug use	Every 3 months	Yes for drug users
Depression score on psychological exam	Monthly	Yes in study of treatment for depression

Weiss (2005),
chapter 1

Table 1.1. Evaluation of some measurements and time intervals between measurements.

Time Scales

1.8. The Language of Longitudinal Data 21

Time interval	Measures
Milliseconds	Neuron firing
Seconds	REM sleep
Minutes	Heart rate
Hours	Pharmacokinetics: blood Concentration after drug dosage
Days	Blood pressure, stock prices
Weeks	Weight loss, store purchases
Months	Growth, 3-month recall
Fractions of 24 hours	Circadian rhythms
Fractions of 30 days	"
Fractions of 365 days	"
Years	Growth and aging Country gross domestic product
Quarters	Educational studies
Semesters	"
Grade in school	"

Weiss (2005),
chapter 1

Table 1.2. Examples of time spacings and research study areas where those spacings might be used.

Lags

- Might make more sense to regress the outcome against a history of the predictors (not just the current value).
- Sometimes, there is a priori reasons to pick a specific *lag* (time difference between a predictor and the outcome).
 - e.g., exposure to pathogen and infectious disease could be known to be bounded by some length of time (hours, days, years)
 - Exposure to carcinogens and development of cancer (years, decades)

Balanced vs. “Messy” Data

- We will have examples of very regimented experiments with highly controlled times of measurement on each unit (dental data, eye experiments, etc.).
- Observational studies of human health, though, often have more complicated, messier form:
 - Different numbers of observations on each person, measured at different times.
 - The number and time of measurements can be “informative”, so introduces potential bias
 - Can be considered a form of missing data

Predominantly a Regression Course

- Emphasize estimating the association explanatory variables with an outcome variable.
- Could potentially examine any aspect of the distribution of the outcome, Y , conditional on the covariates, X .
- Most of the course concentrates on estimating how X affects the mean of Y (although other models are considered):

transformation of the mean of Y in subgroups $X = x$

$$= h(E[Y | X = x]) = \beta_0 + \beta_1 x$$

Regression Models

- Continuous Outcome: Typically linear ($h(\mu) = \mu$):

$$E[Y | X = x] = \beta_0 + \beta_1 x$$

- Binary – Typically logistic, $h(\mu) = \log[\mu / (1 - \mu)]$:

$$\log\left(\frac{E[Y | X = x]}{1 - E[Y | X = x]}\right) = \beta_0 + \beta_1 x, \text{ leads to}$$

$$E[Y | X = x] = P(Y = 1 | X = x) = \frac{1}{1 + \exp\{-(\beta_0 + \beta_1 x)\}}$$

Regression Models, cont

- Counts – typically log-linear, $h(\mu) = \log[\mu]$:

$$\log(E[Y | X = x]) = \beta_0 + \beta_1 x, \text{ leads to}$$

$$E[Y | X = x] = e^{\beta_0 + \beta_1 x}$$

- Often goal is to estimate β_0, β_1 (coefficients), which represent the parameters of interest.

Hazard Regression Models (Disease Incidence Data)

- In survival analysis of disease incidence data, the typical regression approach models the hazard
- The hazard $\lambda(t)$ is:

$$\lambda(t \mid X = x) = P(\text{fail at time } t \mid \text{still at risk just before time } t, X = x)$$

- Possible model is:
$$\lambda(t \mid X = x) = \lambda_0(t)e^{\beta x}$$
- or proportional hazards model

Experimental (Randomized) Multiple Event Data, Time Invariant Covariate (T_x) of Interest

A randomized, controlled trial of an in-home drinking water intervention among HIV+ persons

- Pilot study of 50 HIV+ subjects who were randomized either active water filter or placebo device.
- Followed longitudinally and the number of highly credible gastro-intestinal (HCGI) events were recorded in (on average) a 6 month period.
- Purpose is to estimate the amount of HCGI attributable to drinking water among this population.

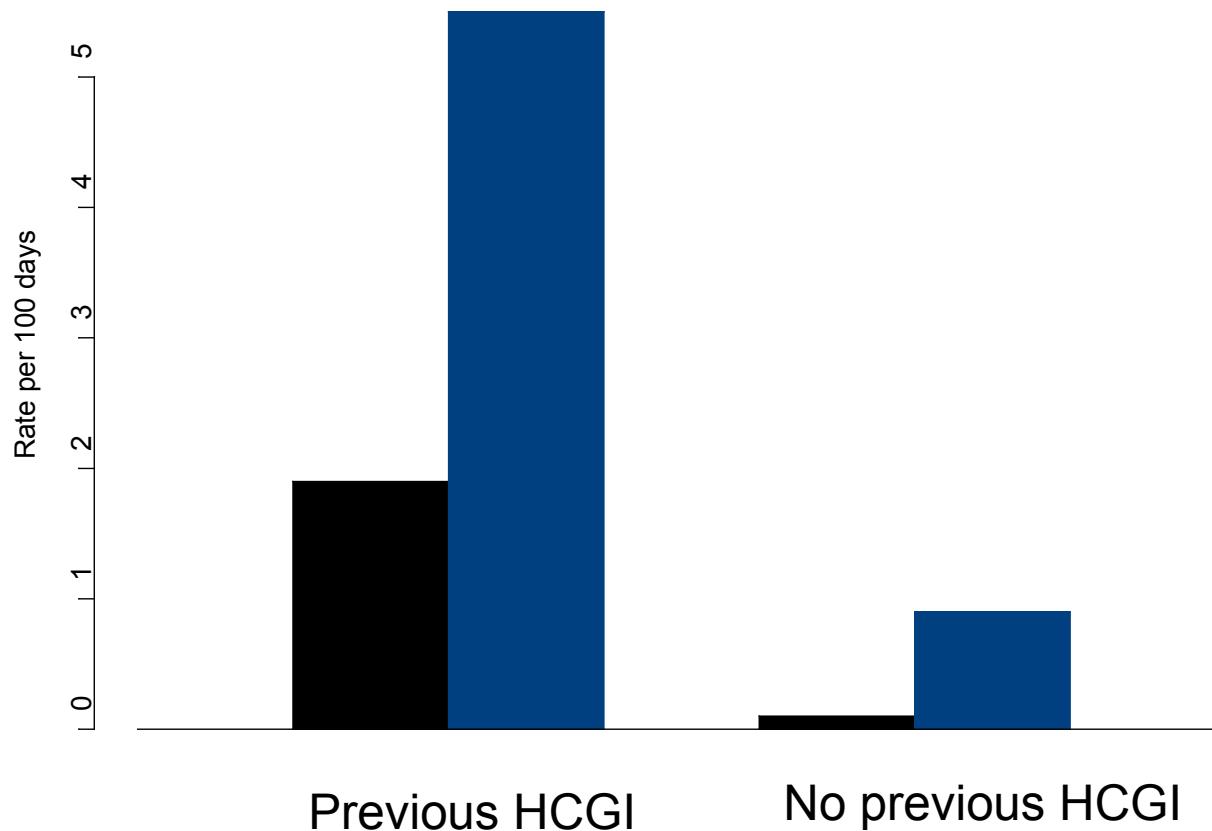
Drinking H₂O and HIV

- Smallest time interval of observation is the day
- Intervention is fixed in time (yes vs. no).
- Though planned to have equal number of measurements per person, many days missed, and the distribution of missing data not the same across subjects

Table 1.2: EXTRACT OF DATA FROM STUDY OF
DENCE OF GASTROINTESTINAL SYMPTOMS

Id. Number	Date	hegi	group
A7283	14780	-	6
A7283	14781	0	6
A7283	14782	0	6
A7283	14783	0	6
A7283	14784	0	6
A7283	14785	0	6
A7283	14786	0	6
A7283	14796	0	6
C1632	14738	-	7
C1632	14739	-	7
C1632	14740	-	7
C1632	14741	0	7
C1632	14742	0	7
C1632	14743	0	7
C1632	14744	1	7
C1632	14745	0	7
C1632	14746	0	7

Results of randomized water intervention among HIV+ persons



**Observational, Messy
Repeated Measures Data
Continuous Outcome with
Time Dependent Variables
(Predictors) of Interest**

Longitudinal Data on HIV+ patients

- Deeks, et al. (1999) report the results from a longitudinal study of HIV-infected adults undergoing Highly Active Anti-Retroviral Therapy (HAART) at San Francisco General Hospital (SFGH).
- Patients were included in this analysis if they received at least 16 weeks of continuous therapy with an anti-retroviral regimen
- The following data was obtained during the initial review: date of birth, sex and length of previous exposure to each individual anti-retroviral agent.

- Once patients were identified, their medical records were reviewed every 3-4 months until November 1998.
- Plasma HIV RNA assays were performed using a branched DNA (bDNA) assay.
- Repeated and irregular measurements of CD4 and viral load (time-structured repeated measures)
- Data not always matched in time.
- Goal is to find how CD4 varies with viral load and how this pattern varies in the population

VL/CD4 data on treated HIV+ patients

- Smallest time interval of observation is the day
- Observational Study (nothing randomized)
- Predictor of interest (VL) changes in time (natural experiment *WITHIN UNIT!!!*)
- Subjects measured when they come into clinic, so very messy (different numbers of measurements for each subject, measured at very different times, etc.).
- Expect the impact of VL to have a lag.

Sample of HIV+ Data

Table 1.1: EXTRACT OF DATA FROM SFGH/HAART STUDY

¹ Id.	Number	days	CD4 count	log(viral load)	gender	age
1		39	45	2.70	1	32.0
1		137	119	5.22	1	32.0
1		147	113	.	1	32.0
1		179	74	5.20	1	32.0
1		187	95	.	1	32.0
1		298	137	3.87	1	32.0
1		335	.	5.07	1	32.0
1		354	167	5.14	1	32.0
1		411	.	4.66	1	32.0
1		1684	427	.	1	32.0
2		0	196	5.68	1	44.0
2		7	369	3.93	1	44.0
2		13	353	4.11	1	44.0
2		27	474	3.55	1	44.0
2		55	425	3.10	1	44.0
2		111	493	2.70	1	44.0
2		139	464	2.70	1	44.0
2		167	448	2.70	1	44.0
2		195	427	2.70	1	44.0

CD4 Count vs. Viral Load, cont.

- Possible Model:

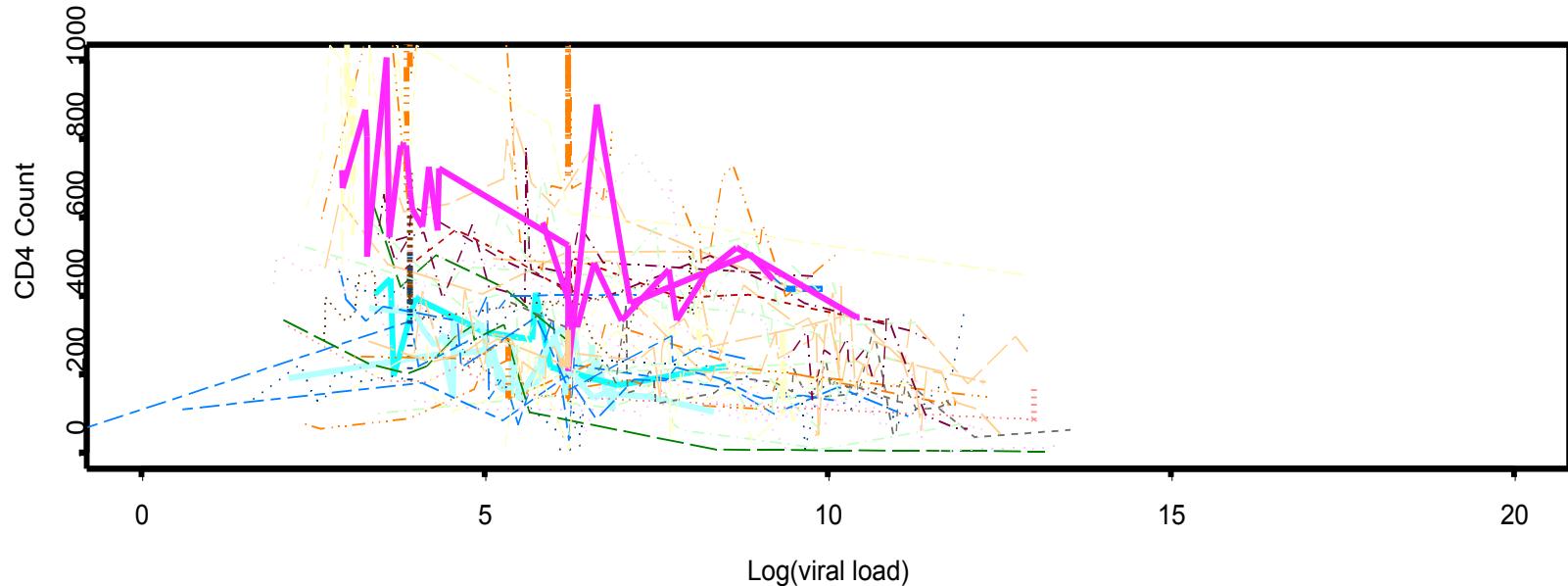
$$Y_{ij}(t) = \beta_0 + \beta_{0i} + (\beta_1 + \beta_{1i})X_{ij}(t - \delta) + e_{ij}$$

Lag

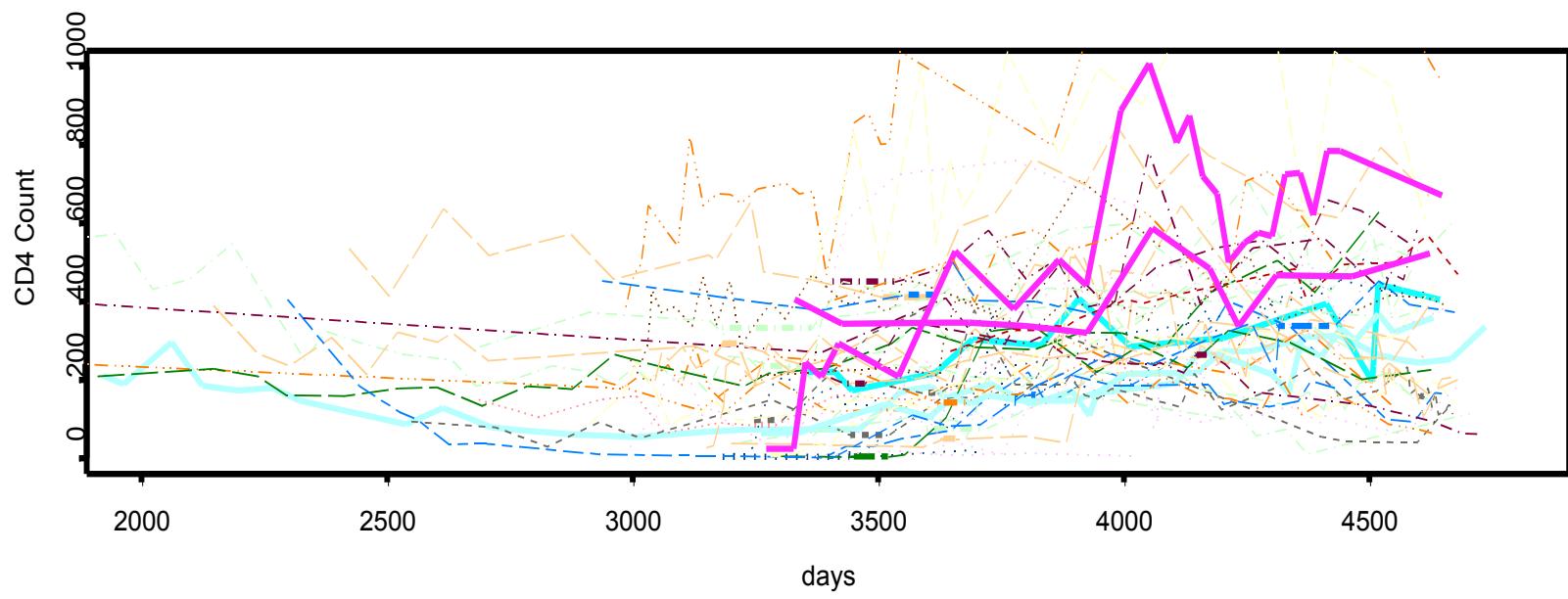


i th subject, j th measurement (at time t) time, $X_{ij(t-\delta)}$ is viral load at time $t-\delta$, $Y_{ij}(t)$ is the CD4 count at time t .

CD4 vs. log(Viral Load)



CD4 vs. time



**Observational, Messy
Repeated Measures Data
Binary Outcome with Time
Dependent Variables
(Predictors) of Interest**

The Effect of Drug and Alcohol Use on Teenage Sexual Activity

- Minnis & Padian (2001) conducted a longitudinal study of teenagers in San Rafael, California to investigate the association between drug and alcohol use and sexual activity on the same day.
- Participants were asked to keep track of their activities over approximately one month and binary indicator variables were created to show whether drug/alcohol use and/or sexual activity were reported for each 24 hour period.

The Effect of Drug and Alcohol Use on Teenage Sexual Activity

- Data is available for 109 teenagers for whom information on 1 to 33 different days are available.
- The average number of longitudinal observations is 16, with the total number of data points (that is, teenager-days) equal to 1,708.

Sex, Drugs, and Longitudinal Data

- Smallest time interval of observation is the day
- Observational Study (nothing randomized)
- Predictor of interest (drugs/alcohol) changes in time (natural within individual “experiment”)
- Subjects suppose to call in every day, but some missing data
- Not interested in lag in outcome versus predictor.

Extract of Teenage Drugs and Sex Data

Table 1.3: EXTRACT OF DATA FROM TEENAGE SURVEY ON DR SEXUAL ACTIVITY

Id. Number	Date	Drug/Alcohol Use	Sexual Activity
10122	03 Jun 98	yes	no
10123	04 Jun 98	no	no
10123	05 Jun 98	no	no
10123	06 Jun 98	yes	no
10123	07 Jun 98	no	no
10123	08 Jun 98	no	no
10123	09 Jun 98	no	no
10123	12 Jun 98	no	no
10123	14 Jun 98	yes	no
10123	16 Jun 98	no	no
10123	17 Jun 98	no	no
10123	18 Jun 98	no	yes
10123	19 Jun 98	no	no
10123	20 Jun 98	no	no
10123	21 Jun 98	no	no
10123	23 Jun 98	no	no
10123	25 Jun 98	no	yes
10123	28 Jun 98	no	no
10123	29 Jun 98	no	yes

Observational, Time to Event (one time only) Data, Time Invariant Predictors

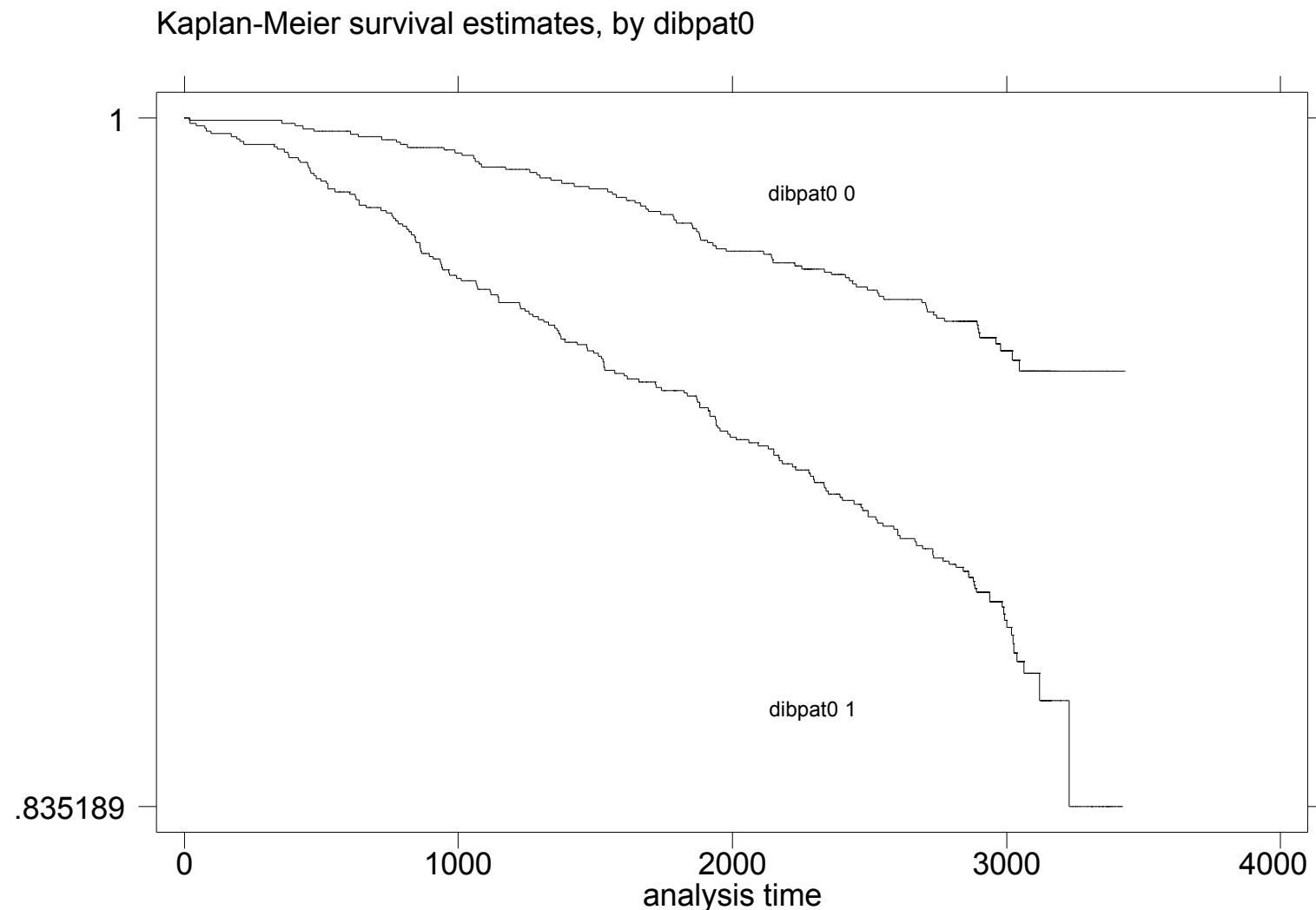
Western Collaborative Group Study

- Collected follow-up data on 3,154 employed men from 10 Californian companies (1960-61).
- Aged 39-59 years old at baseline
- Looked for onset of CHD for about 9 years
- Risk factors measured: smoking, blood pressure, cholesterol, weight, behavior type
- 257 CHD “events”

Western Collaborative Group Study

id	age0	height0	weight0	chol0	behpat0	ncigs0	chd69	time169
2001	49	73	150	225	2	25	0	1664
2002	42	70	160	177	2	20	0	3071
2003	42	69	160	181	3	0	0	3071
2004	41	68	152	132	4	20	0	3064
2005	59	70	150	255	3	20	1	1885
2006	44	72	204	182	4	0	0	3102
2007	44	72	164	155	4	0	0	3074
2008	40	71	150	140	2	0	0	3071
2009	43	72	190	149	3	25	0	3064
2010	42	70	175	325	2	0	0	1032
2011	53	69	167	223	2	25	0	3091
2013	41	67	156	271	2	20	0	3081
2014	50	72	173	238	1	50	1	1528
2017	43	72	180	189	3	30	0	3072

Western Collaborative Group Study



**Ecological Time Series –
Sample Size of 1, with
repeated (daily)
measurements, where both
outcome and predictor vary
over time, and lags are
relevant**

Leptospirosis and Climate

- Leptospirosis is a bacterial disease that affects humans and animals.
- In humans it causes a wide range of symptoms with around 5–10% of infected individuals suffering severe forms of the disease, and, on rare occasions, death.
- Outbreaks of leptospirosis are usually caused by exposure to water contaminated with the urine of infected animals, typically following heavy rainfall with subsequent sewer flooding.
- Urban outbreaks in large Latin American city slums are assumed to result from poor sanitation infrastructure and proliferation of rodent populations.

Leptospirosis and Climate

- The data used here arose from surveillance data in an infectious disease hospital in Salvador, Brazil, an institution that accounts for 95% of case notifications in the city (Flannery et al., 2001).
- In addition, meteorological information on daily rainfall, temperature (maximum, minimum, and average), and relative humidity for the same period were also collected.
- One goal for data analysis is estimation of the lag time between high rainfall days and days of high case counts, providing insight into the disease's incubation period in addition to suggesting appropriate time periods for possible intervention after periods of heavy rain.

LDA—what is it all about and why care?

- Longitudinal data refers to multiple observations on outcomes (and possibly covariates) that change over time
- What are some of the advantages of longitudinal data
 - More observations = more power (?)
 - Better handle on causality (e.g. paired/before & after experiments)
 - Better picture of dynamic effects (e.g. time-dependent covariates—BMI may be cross-sectionally associated with serum cholesterol but this does not answer the question of whether changing your BMI changes your cholesterol)

LDA—what is it all about and why care?

- What are some of the disadvantages (challenges?) of longitudinal data
 - Repeated observations on the same individual are more alike than observations on different individuals (so maybe less additional power than we first thought)
 - Relationship (regression) models may not directly apply (at least inference methods (tests and CIs) will surely be affected by the within-person correlation noted above)
 - Do we want regression models that describe within-person affects of covariates or models that describe population effects of covariates? Does it make a difference?

LDA—what is it all about and why care?

- What are some of the disadvantages (challenges?) of longitudinal data
 - Time-dependent confounding if you have time-dependent exposures (i.e. today's exposure may depend on yesterday's outcome values)

List of Topics

- Introduction to course, examples of data, notation some background material.
- Major themes in course including advantages and complications of longitudinal data (why does longitudinal data matter).
- Graphical representation of longitudinal data.
- Time-dependent (repeated) outcomes, baseline covariates (predictors).
 - Summarizing a function of repeated outcome measures.
 - Counts → Poisson and Neg-Binomial regression.

■ Possible time-dependent covariates (repeated measures regressions)

- Ordinary linear regression with repeated measures, longitudinal data.
- More general approaches:
 - Marginal (GEE) and transitional models.
 - Mixed (hierarchical, multi-level) models.
 - Contrasting the two approaches (To GEE or not to GEE).

Other Possible Topics (Time-Permitting)

- Survival Analysis: Right-censored data, K-M curves, Cox regression.
- Trajectory analysis
- Semi-parametric, causal inference methods
- Ecological *time series*.

Notation—why do we need new stuff?

- So many things (random variables, random vectors, regression coefficients, variance and co-variance parameters,...) so little time.
- If one could explain a statistical model, estimation procedure, etc. simply and efficiently in English every time, we would not need notation.
- Every little detail of notation has meaning (e.g., communicating the units, subunits, parameter vs random variable, etc.).

Notation, cont.

- However, we have complicated data and models and we need a shared language that efficiently translates what we're talking about.
- One goal of this course is to translate a scientific hypothesis regarding longitudinal data into a specific statistical model that yields parameters of interest. This starts with notation.

Typical Symbols Commonly Used for Different types of Objects

- Parameters
- Random Variables, Random Vectors,
Random Matrices
 - Measured Variables
 - Latent Variables
- Constants
- Operations (expectations, sums, logs, etc.)

Notation

- Say one collects an outcome, Y , and a covariate, X on a randomly sampled set of n subjects out of a much larger target population.
- We could represent this experiment as random draws of $O=(Y,X)$.
- We could index each observation, e.g., for each individual, i , $O_i=(Y_i,X_i)$, $i=1,\dots,n$.
- One might be interested in a parameter of the distribution of the O_i , such as the mean of the outcome in certain sub-groups defined by a specific value of the covariate, $X=x$, or $E(Y|X=x)$.

Latent Variables, Parameters, Model

- Consider normal linear model:
 - $O=(Y, X)$, i.i.d. (independent and identically distributed)
 - Statistical Model for Y is:
$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad \varepsilon \text{ i.i.d } N(0, \sigma^2)$$
 - ε is a “latent” variable (error term).
- What are the random variables?
- What are the parameters?
- What does the equation imply about the distribution of Y given X ?

Vectors (of Random Variables, Parameters)

- It's convenient to represent longitudinal data not just as single numbers, but also vectors and matrices.
- A random vector is a set of random variables (e.g., the set of cholesterol measurements made on a subject).
- Another random vector could be the covariates (explanatory variables) measured on the subject at a single time (e.g., age, race, weight, ...).
- A parameter vector is a set of parameters (e.g., the set of predicted mean cholesterol for each of the measurement times).

Matrices

- For this course, matrices are often a convenient way to display both data and parameters.
- An example of representing data is the matrix being simply a set of vectors (e.g., each row of the matrix is a different observation).
- Set of measurements of explanatory variables made on a subject longitudinally can be represented as a matrix where every row contains all the measurements made on that subject at a specific time point– much like spreadsheet.
- Also, matrices are a convenient way to display certain sets of parameters, such as the set of all correlations of outcomes measured on the same subject.

Typical rules regarding random variables vectors, etc.

- Because we have more flexibility in written documents than on the blackboard, we need different rules.
- In documents
 - Random Variables capitalized, realizations small, e.g., $P(Y=y)$.
 - Vectors in bold: $P(\mathbf{Y}=\mathbf{y})=P(Y_1=y_1, Y_2=y_2, \dots)$
 - Matrices in capital or bold or ...
- Hard to avoid that presenter (including us) will sometimes rely on the reader recognizing what's a scalar, vector, matrix, ..., by the context.

But, we'll do our best to be consistent in notation

■ On Board

- Random Variables capitalized, realizations small, e.g., $P(Y=y)$ (same)
- Vectors with arrow over top:

$$P(\vec{Y} = \vec{y}) = P(Y_1 = y_1, Y_2 = y_2, \dots)$$

- Matrices underlined or more likely understood “in context”.

Outcomes and Explanatory variables

- Y_{ij} will represent a response variable, the j th measurement of unit i .
- $\mathbf{X}_{ij}^T = (1, X_{ij1}, X_{ij2}, \dots, X_{ijp})$ or:
$$\mathbf{X}_{ij} = \begin{pmatrix} 1 \\ X_{ij1} \\ X_{ij2} \\ X_{ij3} \\ \vdots \end{pmatrix}$$

Can be (for instance) a vector of length $p+1$ of explanatory variables observed at the j th measurement (note, the 1 is included to allow for an intercept) – the superscript T means transpose (so untransposed is a column vector).

Numbers of observations on one individual and number of individuals

- $j = 1, n_i$. $i=1, m$ - so the number of longitudinal observations for person i is n_i , number of subjects is m .

$$\mathbf{X}_{ij} = \begin{pmatrix} 1 \\ X_{ij1} \\ X_{ij2} \\ \dots \\ X_{ijp} \end{pmatrix}$$

Parameters of Interest

- We will discuss estimates of parameters related to means (like regression coefficients) and those related to variances and covariances.
- For example,
 - $E(Y_{ij})$ or $E(Y_{ij}|\mathbf{X}_{ij}) = \mu_{ij}$ (or $= \mu_{ij}(\mathbf{X}_{ij})$)
 - $\text{Var}(Y_{ij})$ or $\text{Var}(Y_{ij}|\mathbf{X}_{ij}) = v_{ijj}$ (or $= v_{ijj}(\mathbf{X}_{ij})$)

Nesting Observations (measurement within individual)

- Set of repeated measures for unit i are collected into a n_i -vector $\mathbf{Y}_i^T = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})$.
- \mathbf{Y}_i has mean, $E(\mathbf{Y}_i) = \boldsymbol{\mu}_i$ and $n_i \times n_i$ covariance matrix $Var(\mathbf{Y}_i) = \mathbf{V}_i$.
- The jk element of \mathbf{V}_i is the covariance between Y_{ij} and Y_{ik} , that is $cov(Y_{ij}, Y_{ik}) = v_{ijk}$.

Nesting Observations (measurement within individual)

- Note, that $\text{cov}(Y_{ij}, Y_{ij}) = \text{var}(Y_{ij}) = v_{ijj}$
- To represent how observations co-vary on a subject, we will sometimes use correlation: R_i will be the $n_i \times n_i$ correlation matrix of Y_i .

Combining all observations into a big data set.

- We will lump the responses of all units into one big vector $\mathbf{Y}^T = (\mathbf{Y}_1, \dots, \mathbf{Y}_m)$ which is an N -vector (total number of observations):

$$N = \sum_{i=1}^m n_i$$

- Most of the course will focus on regression models of the sort:

$$\begin{aligned} Y_{ij} &= B_0 + B_1 X_{ij1} + \dots + B_p X_{ijp} + e_{ij} \\ &= \mathbf{X}_{ij}^T \boldsymbol{\beta} + e_{ij} \end{aligned}$$

Combining, cont.

- We can write the model for the data on the i th person as

$$\begin{matrix} \mathbf{Y}_i \\ n_i x 1 \end{matrix} = \begin{matrix} X_i \\ n_i x(p+1) \\ (p+1)x1 \end{matrix} \boldsymbol{\beta} + \begin{matrix} \mathbf{e}_i \\ n_i + 1 \end{matrix}$$

- and for the entire data as:

$$\begin{matrix} \mathbf{Y} \\ Nx1 \end{matrix} = \begin{matrix} X \\ Nx(p+1) \end{matrix} \boldsymbol{\beta} + \begin{matrix} \mathbf{e} \\ Nx1 \end{matrix}$$

Example: Sex and drug/alcohol use

i	X_{ij1}	X_{ij2}	Y_{ij}
ID	date	Sex	Drug/Alch
10123	3-Nov-98	no	no
10123	4-Nov-98	no	no
10123	5-Nov-98	no	no
10123	6-Nov-98	no	no
10123	7-Nov-98	no	no
10123	8-Nov-98	no	no
10123	9-Nov-98	no	no
10123	10-Nov-98	no	no
10123	11-Nov-98	no	no
10123	12-Nov-98	no	no
10125	7-Oct-98	no	no
10125	8-Oct-98	yes	no
10125	9-Oct-98	no	yes
10125	10-Oct-98	yes	no
10125	11-Oct-98	yes	no
10125	12-Oct-98	no	no