Modelling Informative Dropout using NONMEM

Nick Holford
University of Auckland

General Principles (Abowd)

- The basic insight is that missing data should be modeled using the same probability and statistical tools that are the basis of all data analysis.
- Missing data are not an anomaly to be swept under the carpet.
- They are an integral part of every analysis.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Missing Data Mechanisms

Little & Rubin distinguish between:

- Missing Completely at Random (MCAR)

  $P(M|Y) = P(M)$ for all Y

- Missing at Random (MAR)

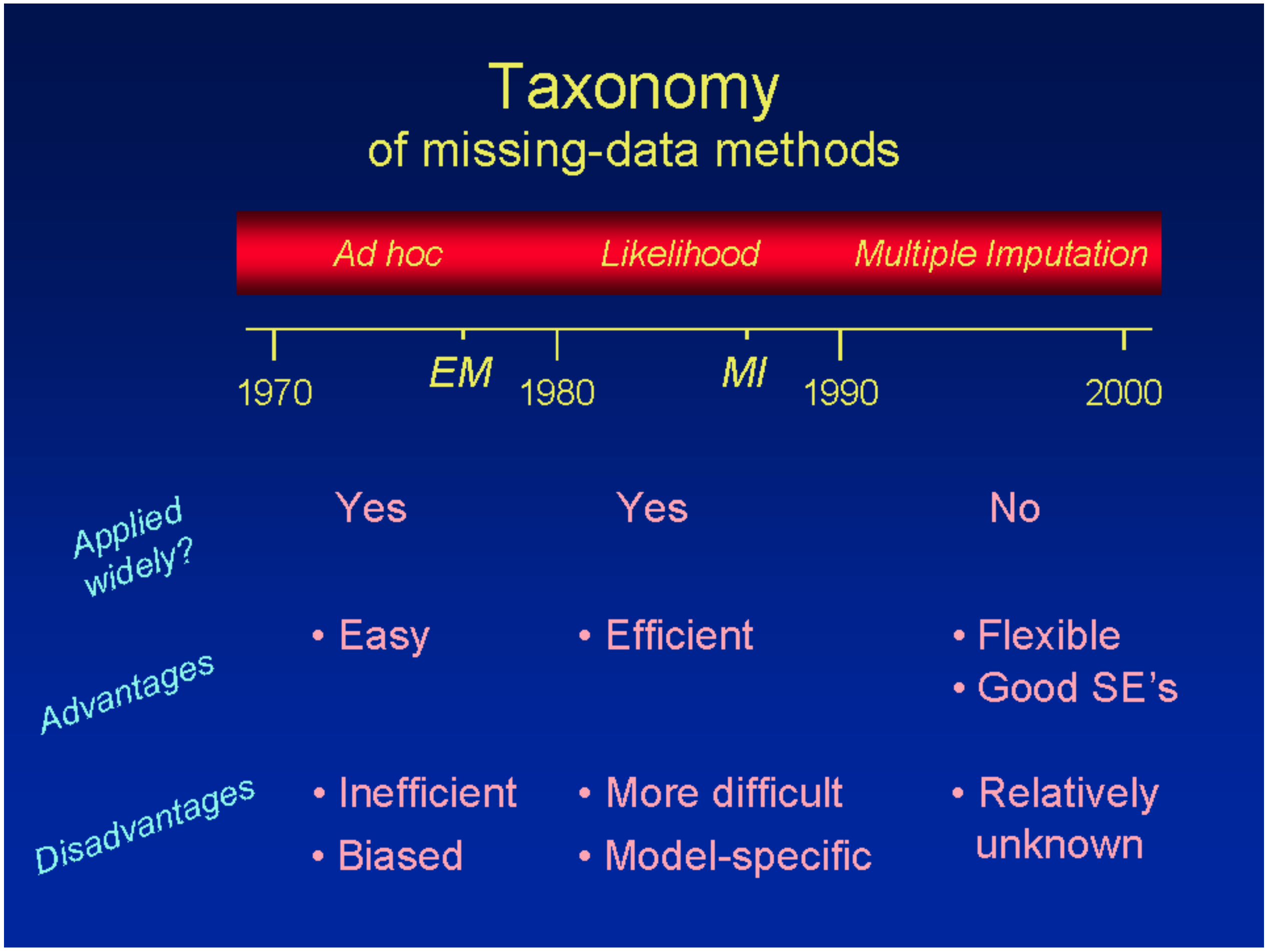
  $P(M|Y) = P(M|Y_{obs})$ for all Y_{miss}

- Not Missing at Random (NMAR)

  $P(M|Y)$ depends on Y_{miss}

Statistical Analysis with Missing Data, 2nd edition, Roderick J. A. Little and Donald B. Rubin (New York: John Wiley & Sons, 2002).

Shafer's Taxonomy of Methods



Shafer JL http://www.stat.psu.edu/%7Ejls/asa97/slide7.html

Modelling Informative Missingness

Model for the Data (PKPD and Disease Progress)

+

Model for the Missingness Mechanism

= Maximum Likelihood Joint Model

Disease Progress Model

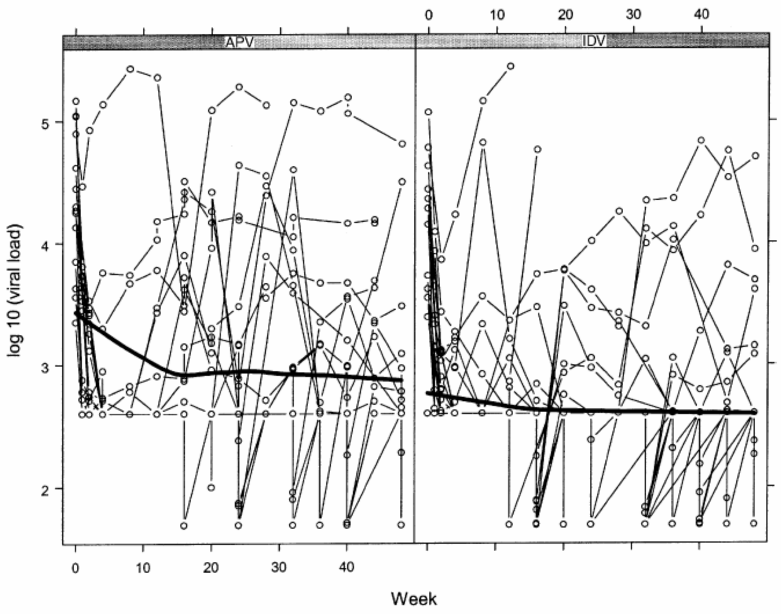


Fig. 1. Observed viral load over time for a random sample of 40 patients. The thick lines are loess smooth of the data.

$$Y_{ij} = a_{trt} + \eta_{1i} + (b_{trt} + \eta_{2i}) * t_j + \varepsilon_{ij}$$

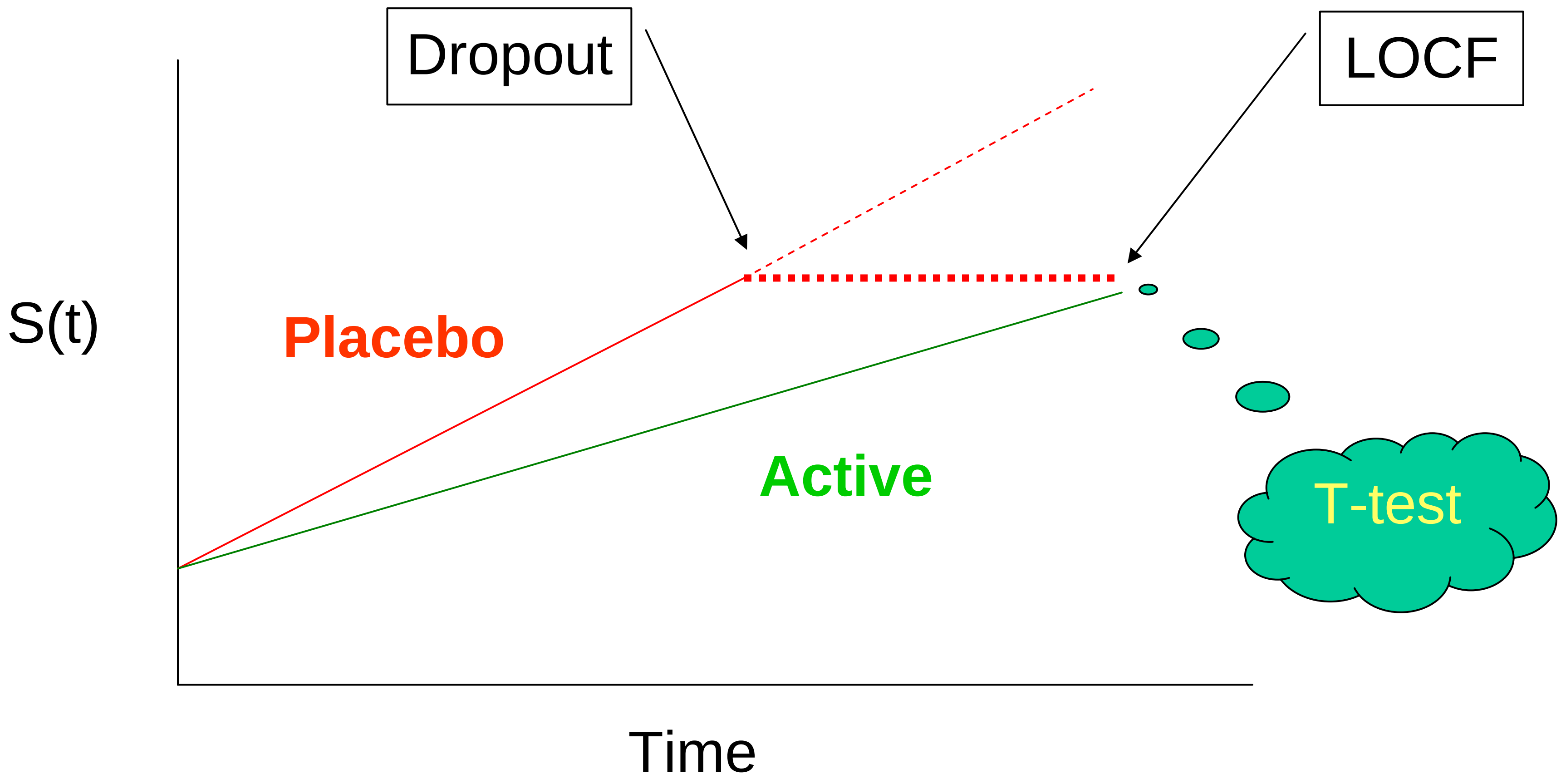
a_{trt} and b_{trt} are fixed effect parameters which may depend on treatment covariate (trt)

Model can describe symptomatic (trt affects a) and protective (trt affects b) disease progress actions of treatment

Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn 2003;30(1):83-103

Disease Progression and Last Observation Carried Forward

Statistical Madness



Better Models for Missing Data

CLINICAL PHARMACOLOGY & THERAPEUTICS
VOLUME 72 NUMBER 6 DECEMBER 2002

COMMENTARY

More efficient clinical trials through use of scientific model-based statistical tests

E. Niclas Jonsson, PhD, and Lewis B. Sheiner, MD *Uppsala, Sweden, and San Francisco, Calif*

Hu & Sale Terminology

Variables are dropout time *T*, observed (Y_O) and unobserved values (Y_U) of disease progress state (e.g. HIV viral load)

(a) **completely random** (CRD), if *Ti* is independent of η, and therefore (Y_O, Y_U);

(b) **random** (RD), if *Ti* (given Y_O) is independent of Y_U, but may depend on Y_O. In addition, any dependence of *Ti* on η is only through Y_O;

(c) **informative** (ID), if *Ti* (given Y_O) depends on Y_U, or explicitly depends on η other than through Y_O.

Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn 2003;30(1):83-103

Hazard, Risk, Survival, Dropout

$$hazard = \beta_0 \cdot e^{\beta \cdot X}$$

$$Risk_i = \int_0^{t_i} hazard$$

$$Survival_i = e^{-Risk_i}$$

$$Dropout_i = 1 - e^{-Risk_i}$$

Probability of Event

IF (DV.EQ.0) ; has not dropped out

$$\Pr_{ID} = e^{-\int_{0}^{t} \beta_0 \cdot e^{\beta_2 \cdot Disprg(t)}}$$

ELSE ; dropped out between $t_{lastobs}$ and t_{end}

$$\Pr_{ID} = e^{-\int_{0}^{t_{lastobs}} \beta_0 \cdot e^{\beta_2 \cdot Disprg(t)}} \cdot \left(1 - e^{-\int_{t_{lastobs}}^{t_{end}} \beta_0 \cdot e^{\beta_2 \cdot Disprg(t)}} \right)$$

β_2 is a parameter describing the informative dropout hazard

-2 Log Likelihood

Disease Progress

$$ELS = \sum_{i=1}^{NSUB} \sum_{j=1}^{NOBS} \left(\frac{\left(Y_{OBS,ij} - Y_{PRED,ij}\right)^2}{Var_{ij}} + \ln\left(Var_{ij}\right) \right)$$

$$CCONTR_{-2LL,ij} = \left(\frac{Y_{OBS,ij} - Y_{PRED,ij}}{SD_{ij}} \right)^2 + 2 \cdot \ln\left(SD_{ij}\right)$$

Dropout Probability

$$CCONTR_{-2LL} = -2 \cdot \ln(\Pr)$$

Hu & Sale Code

```
$MODEL COMP=CUMHAZ ; compartment for integration of hazard
COMP=(HZLAST, INITIALOFF) ; comp for LAST PERIOD hazard
$PK
INTERC=(THETA(1) - THETA(2)*(TRT-1))+ETA(1)
SLOPE=THETA(3)+ETA(2)
BSHZ=THETA(4)
BETA=THETA(5)
BET2=THETA(6)
$DES
VIRL=INTERC+SLOPE*(T-12)
TEMP=BETA*LOCF+BET2*VIRL
DADT(1)=EXP(TEMP)
DADT(2)=EXP(TEMP)
$ERROR
CMHZ=BSHZ*A(1)
HZLA=BSHZ*A(2)
IF (DVID.EQ.1) THEN ; DV=Viral Load
IPRE=INTERC+SLOPE*(TIME-12)
Y=2*LOG(THETA(7))+( (DV-IPRE)/THETA(7) )**2
ENDIF
IF (DVID.EQ.2 .AND. DV.EQ.0) THEN ; NO dropout
Y=-2*(-CMHZ)
ENDIF
IF (DVID.EQ.2 .AND. DV.EQ.1) THEN ; dropout
Y=-2*(-(CMHZ-HZLA)) - 2*LOG(1 - EXP(-HZLA))
ENDIF
```

Modifications of Hu & Sale Method

- User supplied $SUB CCONTR for joint ELS and LIKE contributions to objective function
- "Standard" code for disease progress prediction

Modified Code

```
$INPUT ID TRT TIME CMT LOCF
DV MDV DVID EVID

$ESTIM MAX=9990 SIG=4 NOABORT
METHOD=CONDITIONAL LAPLACE

$CONTR DATA=(DVID)
$SUBR ADVAN=6 TOL=6
CONTR=contr.for
CCONTR=ccontr_like.for

$MODEL
COMP=(CUMHAZ)
COMP=(HZLAST,INITIALOFF)

$PK
BSHZ=THETA(1) ; Baseline hazard
BETA=THETA(2) ; RD hazard
BET2=THETA(3) ; ID hazard
EFFECT=TRT*THETA(4)
INTRI=(THETA(5)+EFFECT)*EXP(ETA(1))
SLOPI=THETA(6)*EXP(ETA(2))
```

```
$DES
DISPRG=INTRI + SLOPI*T
EXPHAZ=EXP(BETA*LOCF + BET2*DISPRG)
DADT(1)=EXPHAZ
DADT(2)=EXPHAZ

$ERROR
CMHZ=BSHZ*A(1) ; Cum hazard overall
HZLA=BSHZ*A(2) ; Cum hazard from last obs
IF (HZLA.LE.0) HZLA=1.0D-10

IF (DVID.EQ.1) THEN
   Y=INTRI + SLOPI*TIME + ERR(1); Status
ENDIF
IF (DVID.EQ.2.AND.DV.EQ.0) THEN
   PD0=EXP(-CMHZ)        ; Pr no dropout
   Y=PD0
ENDIF
IF (DVID.EQ.2.AND.DV.EQ.1) THEN
   PL0=EXP(-(CMHZ-HZLA)) ; Pr no drop last
   PU1=1-EXP(-HZLA)      ; Pr drop unknown
   Y=PL0 * PU1           ; Pr dropout
ENDIF
```

CCONTR

```
      SUBROUTINE CCONTR (ICALL,CNT,P1,P2,IER1,IER2
      SAVE
C LVR and NO should match values in NSIZES
      PARAMETER(LVR=30,NO=50)
      COMMON /ROCM4/ Y(NO),DATA(NO,3)
      DOUBLE PRECISION CNT,P1,P2,Y
      DIMENSION P1(*),P2(LVR,*)
      TYPE=DATA(1,1)
C Value of TYPE is provided as a user defined data item
      IF (TYPE.EQ.1)THEN
C CELS is used for continuous type data
         CALL CELS(CNT,P1,P2,IER1,IER2)
      ELSE
C CLIK is used for LIKE or -2LL
C first argument is 1 for LIKE and 2 for -2LL
         CALL CLIK(1,CNT,P1,P2,IER1,IER2)
      ENDIF
      RETURN
      END
```

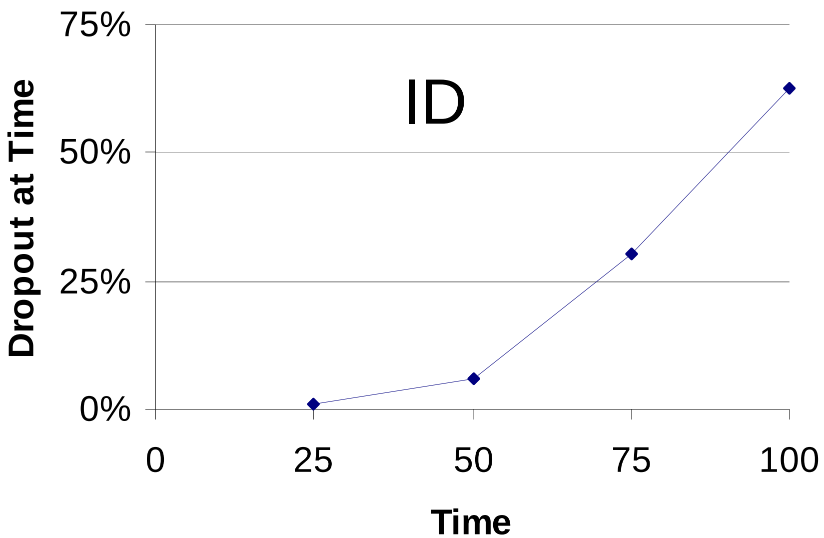
Dropout Data Format

#ID	TRT	TIME	CMT	LOCF	DV	MDV	DVID	EVID
1	1	0	1	0	-0.6	0	1	0
1	1	25	1	-0.6	28.1	0	1	0
1	1	50	1	28.1	53.2	0	1	0
1	1	75	1	53.2	81.8	0	1	0
1	1	100	1	81.8	108.7	0	1	0
1	**1**	**100**	**1**	**108.7**	**0**	**0**	**2**	**0**
2	0	0	1	0	0.1	0	1	0
2	0	25	1	0.1	28.8	0	1	0
2	0	25	2	28.8	0	1	0	2
2	**0**	**50**	**1**	**28.8**	**1**	**0**	**2**	**0**

ID 2: Dropout is between time 25 and 50.
Dropout risk compartment is turned on at 25 (CMT=2).

LIKE vs -2LL Methods
ID Bias and Imprecision

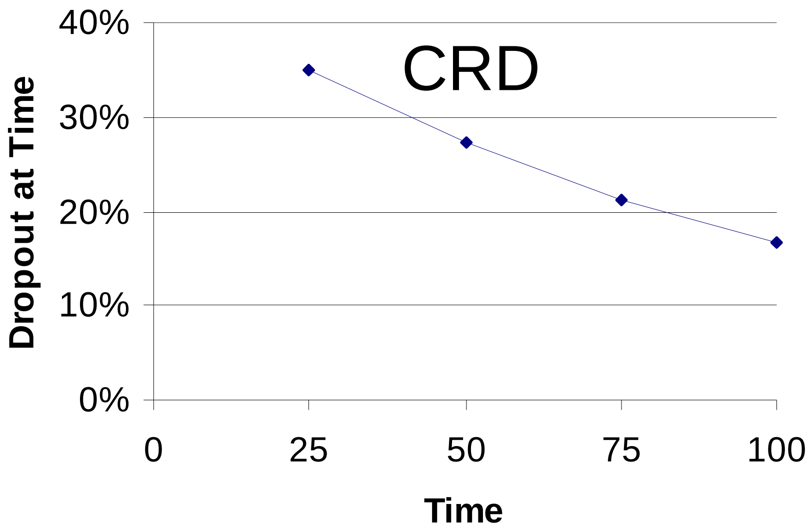
1000 subjects observed at t_i = 0, 25, 50, 75 and 100
100 replications
Slope 1 u/time SD 1 u
Baseline 0.0001
ID hazard 0.065
Average Dropout 53% (95 percentile 50-56%)



	LIKE	Success	79%	7 h 4 min	-2LL	Success	44%	7h 53 min
	Bias	**loCI**	**hiCI**	**RMSE**	**Bias**	**loCI**	**hiCI**	**RMSE**
Slope	0.01%	-0.14%	0.17%	0.7%	0.1%	-0.11%	0.3%	0.67%
PPVslope	-0.15%	-0.77%	0.46%	2.7%	-0.23	-0.77	0.32	2.64
Baseline	2.3%	-1.3%	5.8%	15.8%	15%	11%	17%	10%
ID hazard	-0.16%	-0.8%	0.5%	2.9%	-2.2%	-2.7%	-1.7%	1.7%

CRD (null) Randomization Test

1000 subjects observed at t_i = 0, 25, 50, 75 and 100
1000 replications; RD and ID: One extra parameter
Bootstrap mean and 95% confidence interval
Slope 1 u/time SD 1 u
Baseline 0.01
ID hazard 0.0
Average Dropout 50% (95 percentile 47-53%)
Type I error rate 5%



Model Comparison	**CritOBJ**	**Low**	**High**	**F success**
Null: CRD Alternate: RD	3.75	3.27	4.25	95%
Null: CRD Alternate: ID	3.67	3.33	4.17	95%

Questions

- Is Missingness Informative?
  – Only one bit of information per subject
- Can NONMEM get the right answer?
  – LIKE method with CCONTR is OK
  – Direct coding of -2LL is biased
- Can we distinguish CR, RD and ID?
  – Randomization test shows ΔOBJ is approximately χ^2 distributed

More Material

From

John Abowd

Missing Data Mechanisms

- The complete data are defined as the matrix *Y* ($n \times K$).
- The **pattern** of missing data is summarized by a matrix of indicator variables *M* ($n \times K$).
- The data generating **mechanism** is summarized by the joint distribution of *Y* and *M*.

$$m_{ij} = \begin{cases} 0, \text{ if } y_{ij} \text{ is observed} \\ 1, \text{ if } y_{ij} \text{ is missing} \end{cases}$$

$$p\left(Y, M \middle| \theta, \phi\right)$$

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Missing Completely at Random

- In this case the missing data mechanism does not depend upon the data *Y*.
- This case is called MCAR.

$$p(M \mid Y, \theta, \phi) = p(M \mid \phi)$$

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Missing at Random

- Partition Y into observed and unobserved parts.
- Missing at random means that the distribution of M depends only on the observed parts of Y.
- Called MAR.

$$Y = \left(Y_{\text{obs}}, Y_{\text{mis}}\right)$$

$$p(M \mid Y, \theta, \phi) = p(M \mid Y_{\text{obs}}, \phi)$$

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Not Missing at Random

- If the condition for MAR fails, then we say that the data are not missing at random, NMAR.
- E.g. dropout because of adverse effects or failure of drug to be effective

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

The Rubin and Little Taxonomy for Dealing with Missing Values

- Analysis of the complete records only
- Weighting procedures
- Imputation-based procedures
- Model-based procedures

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Analysis of Complete Records Only

- Assumes that the data are MCAR.
- Only appropriate for small amounts of missing data.
- Used to be common in economics, less so in sociology.
- Now very rare.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Weighting Procedures

- Modify the design weights to correct for missing records.
- Provide an item weight (e.g., earnings and income weights in the CPS) that corrects for missing data on that variable.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Imputation-based Procedures

- Missing values are filled-in and the resulting “Completed” data are analyzed
  - Mean imputation
  - Regression imputation
- Some imputation procedures (e.g., Rubin’s multiple imputation) are really model-based procedures.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Model-based Procedures

- A probability model based on $p(Y, M)$ forms the basis for the analysis.
- This probability model is used as the basis for estimation of parameters or effects of interest.
- Some general-purpose model-based procedures are designed to be combined with likelihood functions that are not specified in advance.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Little and Rubin's Principles

- Imputations should be
  - Conditioned on observed variables
  - Multivariate
  - Draws from a predictive distribution
- Single imputation methods do not provide a means to correct standard errors for estimation error.

From John Abowd http://instruct1.cit.cornell.edu/courses/cis440/8

Hazard Models for Dropout

CRD: $h(t;\, Y_O,\, Y_U,\, \eta,\, \beta) = \beta_0$

RD: $h(t;\, Y_O,\, Y_U,\, \eta,\, \beta) = \beta_0 \exp(\beta_1 Y_{i\text{-}1})$, for $t_{i\text{-}1} \le t < t_i$

ID: $h(t;\, Y_O,\, Y_U,\, \eta,\, \beta) = \beta_0 \exp(\beta_2 Y_U)$, for $t_{i\text{-}1} \le t < t_i$

$Y_{i\text{-}1}$ means LOCF
Y_U means disease progress model prediction

Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn 2003;30(1):83-103

A Special Case

```
$ERROR
CMHZ=BSHZ*A(1) ; Cum hazard overall
HZLA=BSHZ*A(2) ; Cum hazard from last obs
IF (HZLA.LE.0) HZLA=1.0D-10

IF (DVID.EQ.1) THEN
   Y=INTRI + SLOPI*TIME + ERR(1); Status
ENDIF
IF (DVID.EQ.2.AND.DV.EQ.0) THEN
   PD0=EXP(-CMHZ)          ; Pr no dropout
   Y=PD0
ENDIF
IF (DVID.EQ.2.AND.DV.EQ.1) THEN
   PL0=EXP(-(CMHZ-HZLA)) ; Pr no drop last
   PU1=1-EXP(-HZLA)      ; Pr drop unknown
   Y=PL0 * PU1           ; Pr dropout
ENDIF
; Dropout at time of visit
IF (DVID.EQ.2.AND.DV.EQ.2) THEN
   PL0=EXP(-(CMHZ-HZLA)) ; Pr no drop last
   PU1=THETA(PVISIT)     ; Pr due to visit
   Y=PL0 * PU1           ; Pr dropout
ENDIF
```