Specification and Estimation of Multinomial Choice Models

Reading: BBMW Chapter 5
(can skip: 5.3.2, 5.4.3, 5.4.4, 5.6.2-5.6.5, 5.9.1-5.9.3)

(Same drill regarding your prep for class on Tuesday)

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

Random Utility Model: Review

- Decision rule: Utility maximization
  - Individual n selects the alternative with the highest utility U_{in} among those in the choice set Cn

- Utility: $U_{in} = V_{in} + \varepsilon_{in}$
  - V_{in}: Systematic utility expressed as a function of observable variables, e.g.

$$V_{in} = \beta' X_{in} = \sum_{k=1}^{K} \beta_k X_{ink}$$

  - ε_{in}: Random utility component

Random Utility Model: Review (cont.)

Choice probability

$$
\begin{aligned}
P(i|C_n) &= P(U_{in} \geq U_{jn},\ \forall j \in C_n) \\
&= P(U_{in} - U_{jn} \geq 0,\ \forall j \in C_n) \\
&= P(U_{in} = max_j\ U_{jn}, \forall j \in C_n)
\end{aligned}
$$

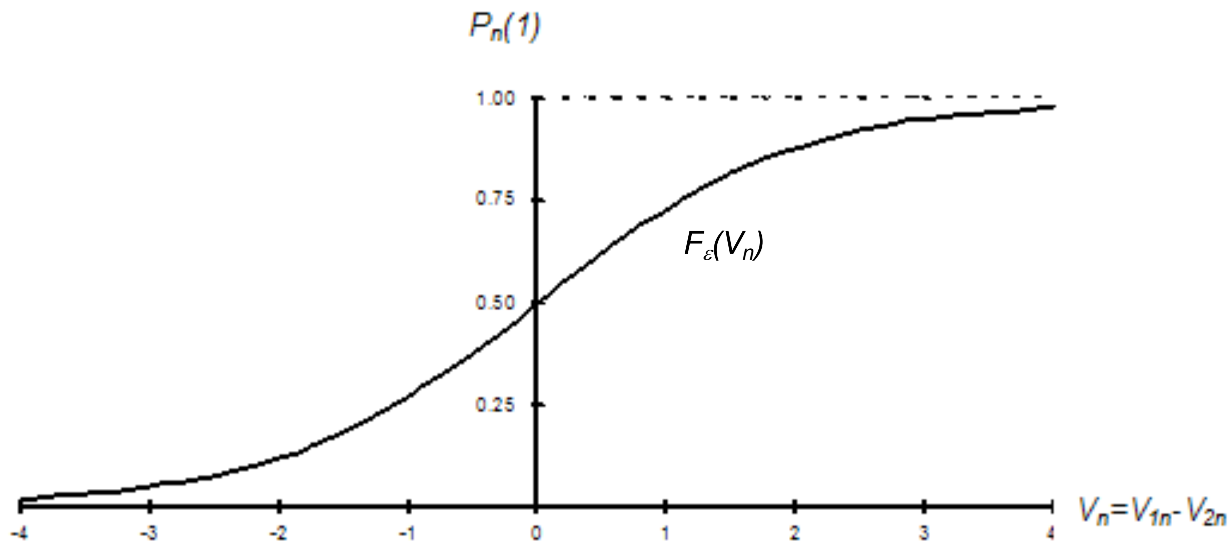
Random Utility Model: Review (cont.)

For binary choice,

$$P_n(1) = P(U_{1n} \geq U_{2n})$$

$$= P(U_{1n} - U_{2n} \geq 0)$$

$$= P(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n}) = F_{\varepsilon_2 - \varepsilon_1}(V_{1n} - V_{2n})$$ [univariate CDF of $\varepsilon_2 - \varepsilon_1$]



Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

Extension to More than Two Alternatives

Choice set C_n: J_n alternatives, $J_n \geq 2$

$$P(i \mid C_n) = P[V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n]$$
$$= P\left[\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n\right]$$

Case of Three Alternatives

Choice set $Cn = \{1,2,3\}\ \forall n$

$$P_n(1) = P(1|C_n) = P(U_{1n} \geq U_{2n} \text{ and } U_{1n} \geq U_{3n})$$

$$= P(V_{1n} + \varepsilon_{1n} \geq V_{2n} + \varepsilon_{2n} \text{ and } V_{1n} + \varepsilon_{1n} \geq V_{3n} + \varepsilon_{3n})$$

$$= P(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n} \text{ and } \varepsilon_{3n} - \varepsilon_{1n} \leq V_{1n} - V_{3n})$$

$$= F_{\varepsilon_2-\varepsilon_1,\varepsilon_3-\varepsilon_1}(V_{1n} - V_{2n}, V_{1n} - V_{3n})$$

$$= \int_{-\infty}^{V_{1n}-V_{2n}} \int_{-\infty}^{V_{1n}-V_{3n}} f_{\varepsilon_2-\varepsilon_1,\varepsilon_3-\varepsilon_1}(q_1, q_2)\, dq_1 dq_2$$

$F_{\varepsilon_2-\varepsilon_1,\varepsilon_3-\varepsilon_1}$: Bivariate CDF of ε_2-ε_1 and ε_3-ε_1

Different assumptions are made on the (joint) distribution of $\varepsilon=(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J)'$, leading to different models.

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

3. Probit

- Probit: ε $\sim MVN(0, \Sigma)$, $f(\varepsilon) = (2\pi)^{-\frac{J}{2}} \left|\Sigma\right|^{-\frac{1}{2}} e^{-\frac{1}{2}(\varepsilon'\Sigma^{-1}\varepsilon)}$

  variance-covariance matrix

$$\Sigma = \begin{bmatrix} Var(\varepsilon_1) & Cov(\varepsilon_1,\varepsilon_2) & \cdots & Cov(\varepsilon_1,\varepsilon_J) \\ Cov(\varepsilon_2,\varepsilon_1) & Var(\varepsilon_2) & \cdots & Cov(\varepsilon_2,\varepsilon_J) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\varepsilon_J,\varepsilon_1) & Cov(\varepsilon_J,\varepsilon_2) & \cdots & Var(\varepsilon_J) \end{bmatrix}$$

- Properties of Probit
  - Flexible substitution patterns
  - Extends to normally distributed random coefficients (random heterogeneity)
  - Extends to panel data with serial correlation

Trinomial Probit

- Model
$$U_{1n} = V_{1n} + \varepsilon_{1n} \quad \textit{where} \quad \begin{pmatrix} \varepsilon_{1n} \\ \varepsilon_{2n} \\ \varepsilon_{3n} \end{pmatrix} \sim N\left(0, \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}\right)$$
$$U_{2n} = V_{2n} + \varepsilon_{2n}$$
$$U_{3n} = V_{3n} + \varepsilon_{3n}$$

- In deviation with respect to first alternative

$$U_{1n} - U_{2n} = V_{1n} - V_{2n} - (\varepsilon_{2n} - \varepsilon_{1n})$$

$$U_{1n} - U_{3n} = V_{1n} - V_{3n} - (\varepsilon_{3n} - \varepsilon_{1n})$$

- Probability

$$P_n(1) = P(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n} \ \ and \ \varepsilon_{3n} - \varepsilon_{1n} \leq V_{1n} - V_{3n})$$

$$\begin{bmatrix} \varepsilon_2 - \varepsilon_1 \\ \varepsilon_3 - \varepsilon_1 \end{bmatrix} \sim N\left(0, \Sigma_1 = \begin{bmatrix} \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} & \sigma_1^2 + \sigma_{23} - \sigma_{12} - \sigma_{13} \\ \sigma_1^2 + \sigma_{23} - \sigma_{12} - \sigma_{13} & \sigma_1^2 + \sigma_3^2 - 2\sigma_{13} \end{bmatrix}\right)$$

$$P_n(1) = \int_{-\infty}^{V_{1n} - V_{2n}} \int_{-\infty}^{V_{1n} - V_{3n}} n(q; 0, \Sigma_1)\, \mathrm{d}q_2 \mathrm{d}q_1$$

Probit (cont.)

- In the general case, the choice probability of alternative 1 is:

$$P_n(1) = \int_{-\infty}^{V_{1n}-V_{2n}} \int_{-\infty}^{V_{1n}-V_{3n}} \cdots \int_{-\infty}^{V_{1n}-V_{Jn}} n(q;0,\Sigma_1)\ \mathrm{d}q$$

- $n(\varepsilon;0,\Sigma)$ denotes the multivariate normal density with mean vector 0 and variance-covariance matrix Σ.
- A J-1 dimensional integral

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

Logit

- ε_{jn} independently and identically distributed (i.i.d.)

$$f(\varepsilon_1, \ldots, \varepsilon_J) = \prod_{j=1}^{J} f(\varepsilon_j)$$

- $\varepsilon_{jn} \sim$ *Extreme Value* $(0,\mu) \quad \forall j$

$$F(\varepsilon) = \exp\left[-e^{-\mu\varepsilon}\right], \ \mu > 0$$

$$f(\varepsilon) = \mu e^{-\mu\varepsilon} \exp\left[-e^{-\mu\varepsilon}\right]$$

- Variance: $\pi^2/6\mu^2$

$$\mathrm{P}\left(i \mid C_n\right) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}}$$

Derivation of Logit

$$P_n(1) = P(V_{1n} + \varepsilon_{1n} \geq \underbrace{\max_{j=2,\dots,J_n} (V_{jn} + \varepsilon_{jn}))}_{}$$

$$U_n^* \sim EV(\frac{1}{\mu}\ln\sum_{j=2}^{J_n} e^{\mu V_{jn}}, \mu)$$

[Remember: If $\varepsilon_i \sim EV(\eta_i, \mu)$ then

$$\varepsilon = \max_{i=1,\dots,J} \varepsilon_i \sim EV(\frac{1}{\mu}\ln\sum_{i=1}^{J} e^{\mu\eta_i}, \mu)]$$

$$U_n^* = V_n^* + \varepsilon_n^* \text{ where } V_n^* = \frac{1}{\mu}\ln\sum_{j=2}^{J_n} e^{\mu V_{jn}} \text{ and } \varepsilon_n^* \sim EV(0, \mu)$$

Derivation of Logit (cont.)

$$
\begin{aligned}
P_n(1) &= P(V_{1n} + \varepsilon_{1n} \geq V_n^* + \varepsilon_n^*) \\
&= P((\varepsilon_n^* - \varepsilon_{1n}) \leq (V_{1n} - V_n^*)) \\
&= F_\varepsilon(V_{1n} - V_n^*) \qquad \text{[Remember:} F_\varepsilon(V_n) = 1/(1 + e^{-\mu V_n})] \\
&= \frac{1}{1 + e^{\mu(V_n^* - V_{1n})}} = \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\mu V_n^*}} \\
&= \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\ln \sum_{j=2}^{J_n} e^{\mu V_{jn}}}} = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^{J_n} e^{\mu V_{jn}}}
\end{aligned}
$$

Example: Logit for 3 Alternatives

Choice set $C_n = \{1,2,3\} \;\forall n$

$$P_n(1) = F_{\varepsilon_2-\varepsilon_1,\varepsilon_3-\varepsilon_1}(V_{1n} - V_{2n}, V_{1n} - V_{3n})$$

$$= \frac{1}{1 + e^{-\mu(V_{1n}-V_{2n})} + e^{-\mu(V_{1n}-V_{3n})}}$$

$$= \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\mu V_{2n}} + e^{\mu V_{3n}}}$$

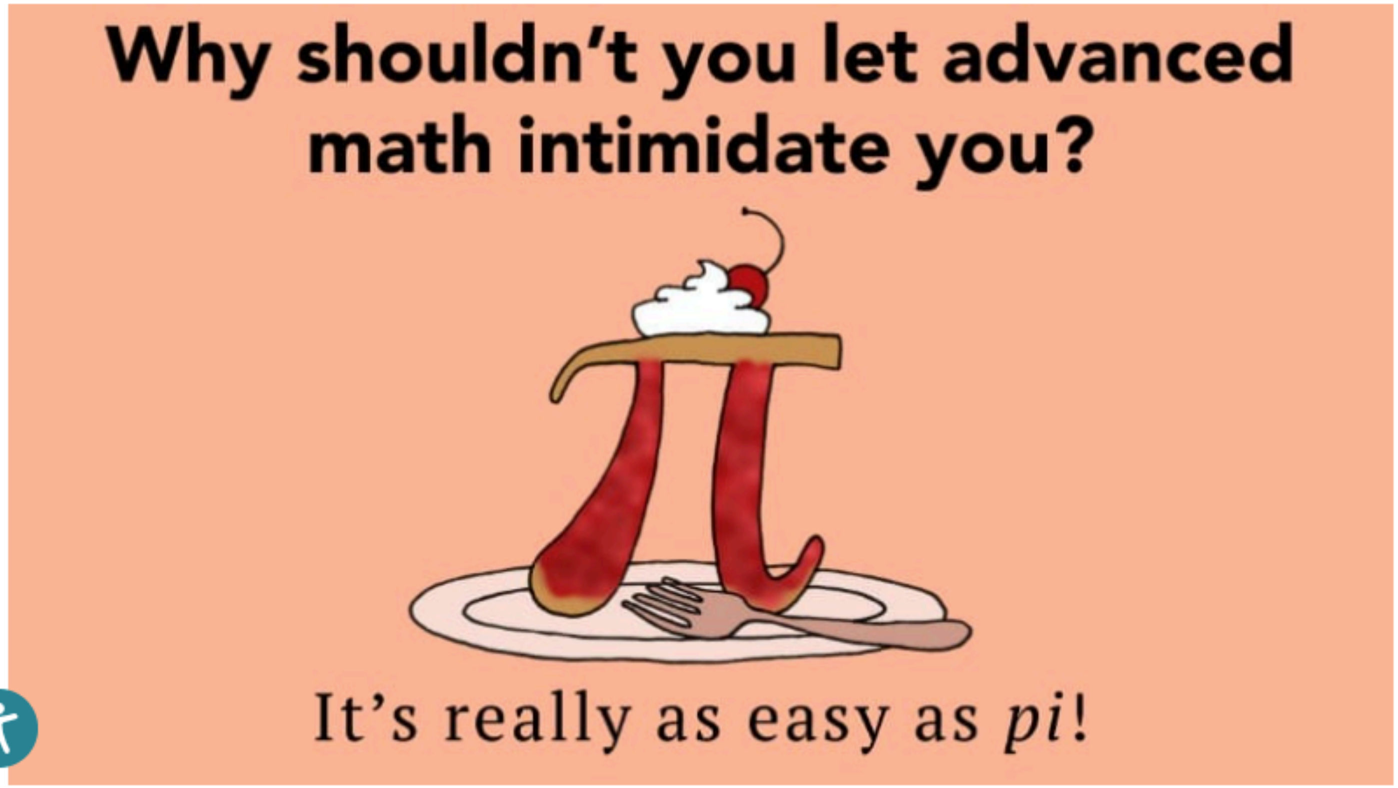
Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

40 Cheesy Math Jokes That'll Make "Sum" of Your Students LOL

Why was six afraid of seven? Because seven eight nine!

Elizabeth Mulvahill on January 8, 2020



Properties of Logit

Independence from Irrelevant Alternatives (IIA)

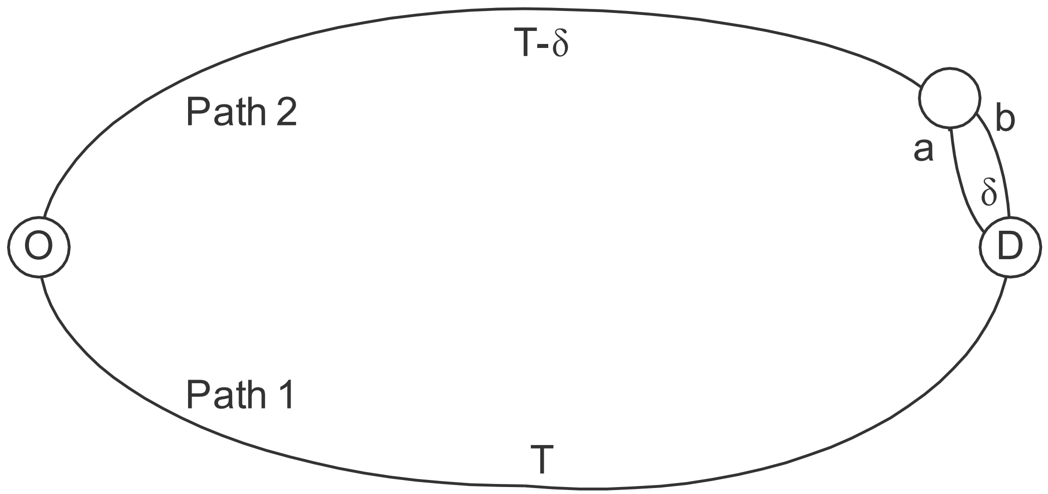
The model $$P_n\left(i \mid C_n\right) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}}$$

Odds ratio $$\frac{P\left(i \mid C_{1n}\right)}{P\left(j \mid C_{1n}\right)} = \frac{P\left(i \mid C_{2n}\right)}{P\left(j \mid C_{2n}\right)}$$

$$i, j \in C_{1n}, i, j \in C_{2n}, C_{1n} \subseteq C_n \text{ and } C_{2n} \subseteq C_n$$

Examples of IIA

Route choice with an overlapping segment



$$P(1|\{1,2a,2b\}) = P(2a|\{1,2a,2b\}) = P(2b|\{1,2a,2b\}) = \frac{e^{\mu T}}{\sum_{j \in \{1,2a,2b\}} e^{\mu T}} = \frac{1}{3}$$

Red Bus / Blue Bus Paradox

- Consider auto and bus with the same utility
  - $C = \{auto, bus\}$ and $V_{auto} = V_{bus} = V$
  - $P(auto) = P(bus) = 1/2$
- Suppose that a new bus service is introduced that is identical to the existing bus service, except the buses are painted differently (red vs. blue).
  - $C = \{auto, red\ bus, blue\ bus\}$; $V_{red\ bus} = V_{blue\ bus} = V$
  - Logit now predicts:
    $P(auto) = P(red\ bus) = P(blue\ bus) = 1/3$
  - We'd expect
    $P(auto) = 1/2$, $P(red\ bus) = P(blue\ bus) = 1/4$

IIA and Aggregation

- Divide the population into two equally-sized groups:
  those who prefer autos, and those who prefer public transport

- Mode shares before introducing blue bus

Population	Auto Share	Red Bus Share	
Auto people	90%	10%	P(auto)/P(red bus) = 9
Public Transport people	10%	90%	P(auto)/P(red bus) = 1/9
Total	**50%**	**50%**	

- Auto and red bus share ratios remain constant for each group after introducing blue bus

Population	Auto Share	Red Bus Share	Blue Bus Share
Auto people	81.8%	9.1%	9.1%
Public Transport people	5.2%	47.4%	47.4%
Total	**43.5%**	**28.25%**	**28.25%**

Motivation for Nested Logit

- Overcome the IIA Problem of Logit when
  - Alternatives are correlated
    (e.g., red bus and blue bus, overlapping paths)

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

Specification of the V's

- For all i in C_n, $U_{in} = V_{in} + \varepsilon_{in}$
  - What is C_n?
  - What is V_{in}?

Choice Set

- Universal choice set
  - All potential alternatives for the population
  - Restricted to relevant alternatives
- Individual choice set
  - Deterministic availabilities constraints
  - (e.g., car driver unavailable if no driver's license)

Functional Form

- $V_{in} = V(Z_{in}, S_n) = V(X_{in})$

  Z_{in} – attributes

  S_n – characteristics
- Linear-in-parameters utility functions:

$$V_{in} = \beta' X_{in} = \sum_{k=1}^{K} \beta_k X_{ink}$$

  - Not as restrictive as it may seem
- Interaction of attributes and characteristics
  - e.g., $V_{1n} = \beta_1 + \beta_2 * cost_{1n} / income_n + \ldots$

Explanatory variables: alternatives attributes

- Numerical and continuous
- $(z_{in})_p \in R, \ \forall i, n, p$
- Associated with a specific unit p
- Examples:
  - Auto in-vehicle time (in min.)
  - Transit in-vehicle time (in min.)
  - Auto out-of-pocket cost (in cents)
  - Transit fare (in cents)
  - Walking time to the bus stop (in min.)

Explanatory variables: alternatives attributes (cont.)

- V_{in} is unitless
- Therefore, β depends on the unit of the associated attribute
- Example: consider two specifications

$$V_{in} = \beta_1 TT_{in} + \cdots \quad \& \qquad V_{in} = \beta'_1 \, TT'_{in} + \cdots$$

- If TT_{in} is a number of minutes, the unit of β_1 is $1/min$
- If TT'_{in} is a number of hours, the unit of β'_1 is $1/hour$
- Both models are equivalent, but the estimated value of the coefficient will be different

$$\beta_1 TT_{in} = \beta'_1 \, TT'_{in} \Rightarrow \; TT_{in}/TT'_{in} = \beta'_1 / \beta_1 = 60$$

Explanatory variables: alternatives attributes (cont.)

- Generic and alternative specific parameters

$$V_{auto} = \beta_1 TT_{auto}$$
$$V_{bus} = \beta_1 TT_{bus}$$

or

$$V_{auto} = \beta_1 TT_{auto}$$
$$V_{bus} = \beta_2 TT_{bus}$$

- Modeling assumption: a minute has/has not the same marginal utility whether it is incurred on the auto or bus mode

Explanatory variables: socio-economic characteristics

- Numerical and continuous
- $(S_n)_p \in R, \ \forall n, p$
- Associated with a specific unit p
- Note: S_n do not depend on i
- Examples:
  - Annual income (in thousand \$)
  - Age (in years)

Explanatory variables: socio-economic characteristics (cont.)

- Socio-economic variables can appear in $(J\text{-}1)$ utility functions, where J is the number of alternatives
- In general: alternative specific characteristics

$$V_1 = \beta_1 x_{11} + \beta_2 income + \beta_4 age$$
$$V_2 = \beta_1 x_{21} + \beta_3 income + \beta_5 age$$
$$V_3 = \beta_1 x_{31}$$

Interactions

- All individuals in a population are not alike
- Socio-economic characteristics define segments in the population
- Interactions of characteristics and attributes
  - Cost/income
  - Fare/disposable income
  - Out-of-vehicle time/distance

Correlation of attributes may produce degeneracy in the model.
E.g. speed and time if distance is constant

Interactions: discrete segmentation

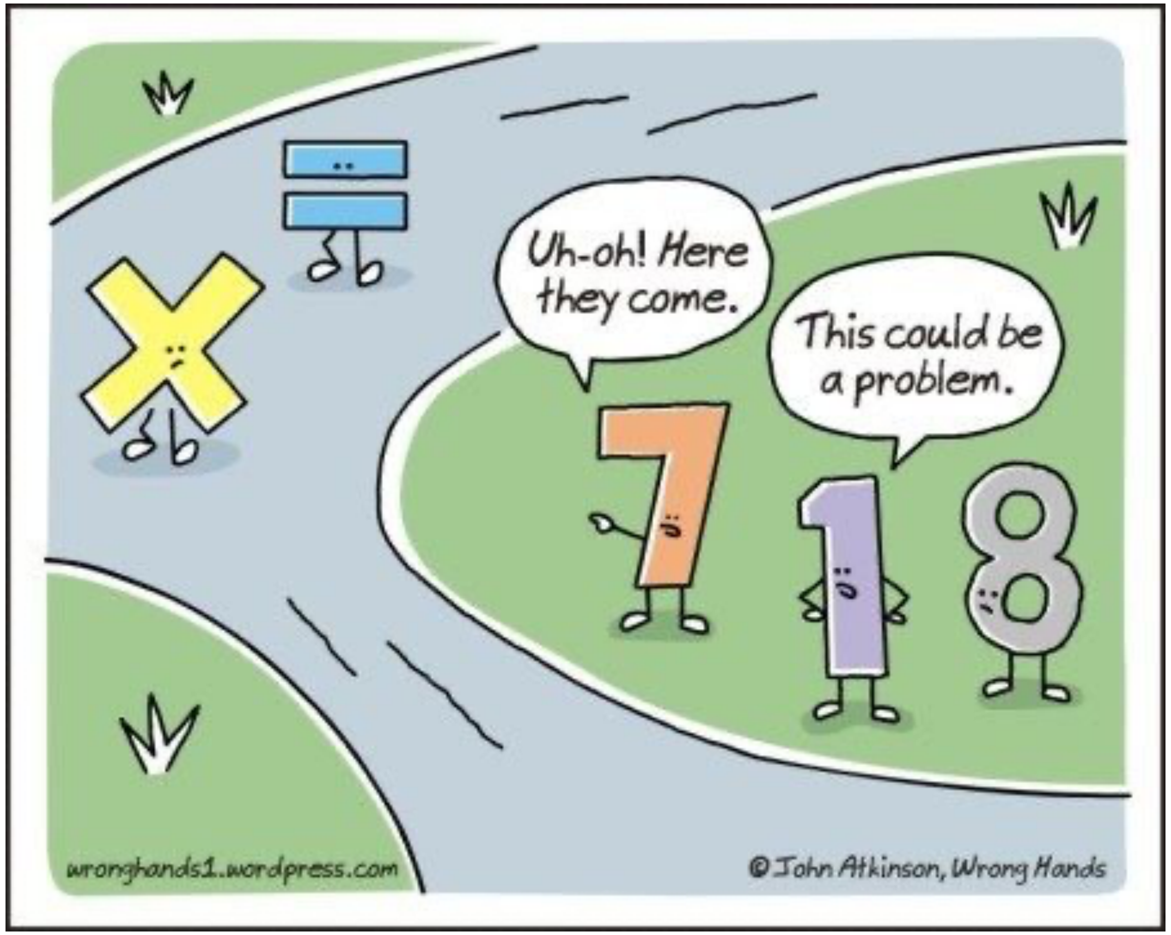
- The population is divided into a finite number of segments
- Each individual belongs to exactly one segment
- Example:

gender (M,F) and house location (metro, suburb, perimeter areas)

→ 6 segments

$$\beta_{M,m} TT_{M,m} + \beta_{M,s} TT_{M,s} + \beta_{M,p} TT_{M,p} +$$
$$\beta_{F,m} TT_{F,m} + \beta_{F,s} TT_{F,s} + \beta_{F,p} TT_{F,p}$$

$TT_i = TT$ if individual belongs to segment i, and 0 otherwise


Uh-oh! Here they come.
This could be a problem.
wronghands1.wordpress.com
© John Atkinson, Wrong Hands

Nonlinear specification

- Nonlinear transformations of the independent variables
- Dummy variables for discrete and qualitative variables
- Continuous variables
  - Categories
  - Splines
  - Box-Cox
  - Power series

Continuous variables: categories

- Assumption: sensitivity to travel time varies with travel time
- Categories are defined: travel time in minutes

  [0 - 90), [90 - 180), [180 - 270) , [270 -)
- Approaches:
  - Dummy variable
    - Not great because utility “jumps” at boundary points
  - Piecewise linear specification (spline)

Piecewise linear specification

- Specification:

$$V = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \ldots$$

where

$$x_1 = \begin{cases} x & \text{if } x < b_1 \\ b_1 & \text{otherwise} \end{cases}$$

$$x_2 = \begin{cases} 0 & \text{if } x < b_1 \\ x - b_1 & \text{if } b_1 \le x < b_2 \\ b_2 - b_1 & \text{otherwise} \end{cases}$$

$$x_3 = \begin{cases} 0 & \text{if } x < b_2 \\ x - b_2 & \text{if } b_2 \le x < b_3 \\ b_3 - b_2 & \text{otherwise} \end{cases}$$

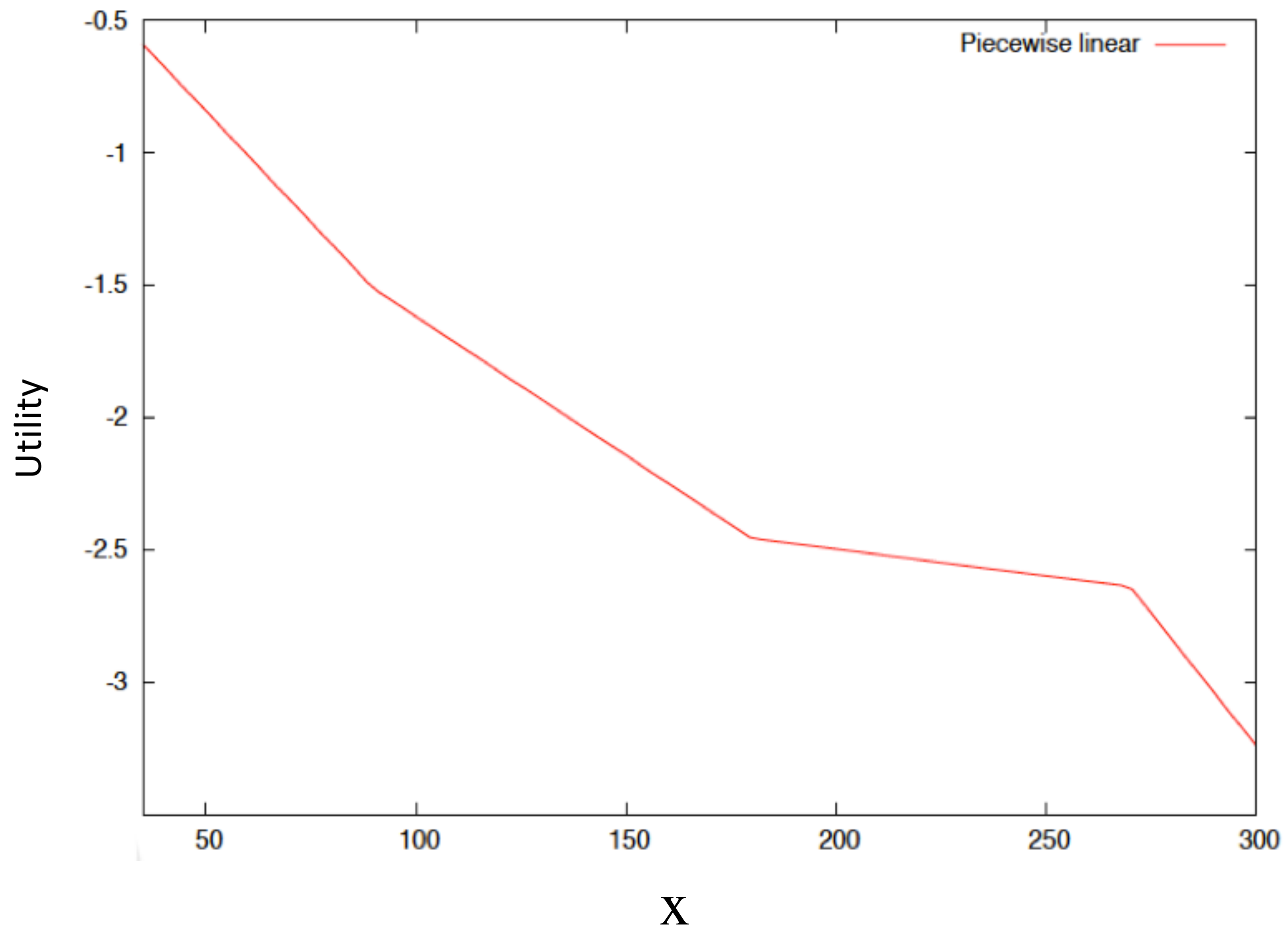
$$x_4 = \begin{cases} 0 & \text{if } x < b_3 \\ x - b_3 & \text{otherwise} \end{cases}$$

Piecewise linear specification (cont.)

- Examples: $b_1 = 90$, $b_2 = 180$, $b_3 = 270$

x	x_1	x_2	x_3	x_4
40	40	0	0	0
100	90	10	0	0
200	90	90	20	0
300	90	90	90	30

Piecewise linear specification (cont.)



Box-Cox Transforms

- $V = \beta x(\lambda) + \cdots$

  where

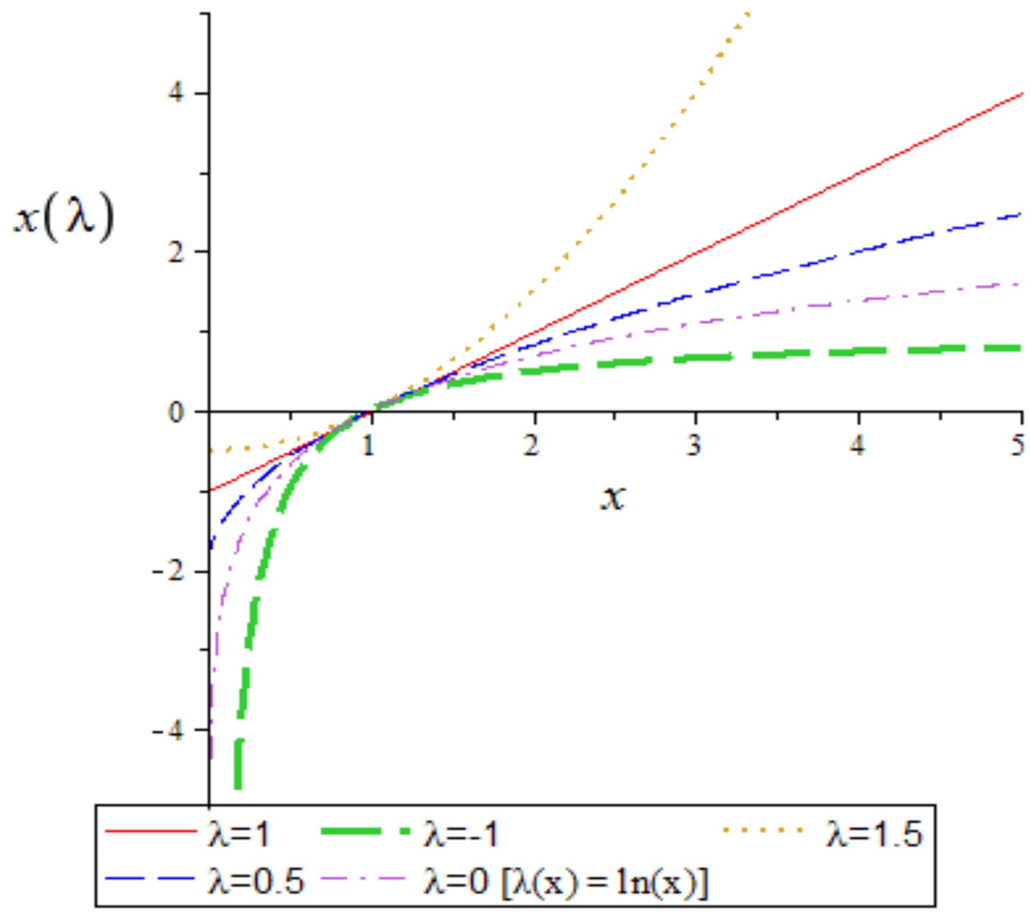
  $$x(\lambda) = \begin{cases} (x^{\lambda} - 1) / \lambda & \text{if } \lambda \neq 0 \\ \ln x & \text{if } \lambda = 0 \end{cases}$$

  where $x > 0$

- If $x < 0$, let α such that $x + \alpha > 0$ and

  $$x(\lambda,\alpha) = \begin{cases} ((x + \alpha)^{\lambda} - 1\ ) / \lambda & \text{if } \lambda \neq 0 \\ \ln\,(x + \alpha) & \text{if } \lambda = 0 \end{cases}$$

Box-Cox Transforms (cont.)



Power Series

$$V = \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \ldots$$

In practice, these terms could be correlated

Difficult to interpret

Risk of over-fitting

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

	Non stop flight (1)	One stop flight with the same airline (2)	One stop flight with a change of airline (3)
β_1	0	1	0
β_2	0	0	1
β_3	Round trip Fare ($100) of (1)	Round trip fare ($100) of (2)	Round trip fare ($100) of (3)
β_4	Elasped time (hours) for (1)	Elapsed time (hours) for (2)	Elapsed time (hours) for (3)
β_5	Leg room in (1) (inches), if male	Leg room in (2) (inches), if male	Leg room in (3) (inches), if male
β_6	Leg room in (1) (inches), if female	Leg room in (2) (inches), if female	Leg room in (3) (inches), if female
β_7	Being early (hours) for (1), at departure or arrival, depending on the preference of the respondent	Being early (hours) for (2), at departure or arrival, depending on the preference of the respondent	Being early (hours) for (3), at departure or arrival, depending on the preference of the respondent
β_8	Being late (hours) for (1), at departure or arrival, depending on the preference of the respondent	Being late (hours) for (2), at departure or arrival, depending on the preference of the respondent	Being late (hours) for (3), at departure or arrival, depending on the preference of the respondent
β_9	0	1 if the respondent makes more than two air trips per year	0
β_{10}	0		1 if the respondent makes more than two air trips per year
β_{11}	0	1 if male, 0 otherwise	
β_{12}	0	0	1 if male, 0 otherwise

Table 5.2: Specification table of the model for the choice of airline itinerary

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-0.879	0.219	-4.02	0.00
2	One stop–multiple airlines dummy	-1.27	0.227	-5.60	0.00
3	Round trip fare ($100)	-1.81	0.151	-11.99	0.00
4	Elapsed time (hours)	-0.303	0.0778	-3.90	0.00
5	Leg room (inches), if male (non stop)	0.100	0.0330	3.04	0.00
6	Leg room (inches), if female (non stop)	0.182	0.0318	5.71	0.00
7	Leg room (inches), if male (one stop)	0.113	0.0297	3.80	0.00
8	Leg room (inches), if female (one stop)	0.0931	0.0273	3.41	0.00
9	Being early (hours)	-0.151	0.0189	-7.99	0.00
10	Being late (hours)	-0.0975	0.0167	-5.83	0.00
11	More than 2 air trips per year (one stop–same airline)	-0.300	0.141	-2.12	0.03
12	More than 2 air trips per year (one stop–multiple airlines)	-0.0847	0.157	-0.54	0.59
13	Male dummy (one stop–same airline)	0.100	0.133	0.75	0.45
14	Male dummy (one stop–multiple airlines)	0.189	0.144	1.31	0.19
15	Round trip fare / income ($100/$1000)	-23.8	8.09	-2.94	0.00

Summary statistics

Number of observations = 2544

$$\begin{aligned}
\mathcal{L}(0) &= -2794.870 \\
\mathcal{L}(c) &= -2203.160 \\
\mathcal{L}(\hat{\beta}) &= -1640.525 \\
-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] &= 2308.689 \\
\rho^2 &= 0.413 \\
\bar{\rho}^2 &= 0.408
\end{aligned}$$

Table 5.4: Specification of the airline itinerary choice model with an interaction between the traveling fare and the income, as well as alternative specific leg room coefficients

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation

Maximum Likelihood Estimation

- Log Likelihood function:

$$L(\beta) = \ln L^*(\beta) = \sum_{n=1}^{N} \ln P\left(\mathrm{y_n} \middle| X_n, \beta\right) = \sum_{n=1}^{N} \left(\sum_{i \in C_n} y_{in} \ln P\left(i \middle| C_n\right) \right)$$

where $yin = 1$ if n chose alternative i, 0 otherwise

- Logit:

$$\mathrm{P}(i \mid C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \quad , \quad \mu = 1$$

$$L(\beta) = \sum_{n=1}^{N} \left(\sum_{i \in C_n} y_{in} \left(V_{in} - \ln \sum_{j \in C_n} e^{V_{jn}} \right) \right)$$

Maximum Likelihood Estimation (cont.)

- The maximum likelihood estimation problem:

$$\hat{\beta} = \text{argmax}_{\beta} L\,(\beta_1, \beta_2, \ldots, \beta_K)$$

- FOC for linear in parameters Logit

$$L(\beta) = \sum_{n=1}^{N}\left(\sum_{i\in C_n} y_{in}\left(V_{in} - \ln\sum_{j\in C_n} e^{V_{jn}}\right)\right) \qquad V_{in} = \sum_{k=1}^{K}\beta_k X_{ink}$$

$$\frac{\partial L(\beta)}{\partial \beta_k} = \sum_{n=1}^{N}\left(\sum_{i\in C_n} y_{in}\left(x_{ink} - \frac{\sum_{j\in C_n} x_{jnk} e^{V_{jn}}}{\sum_{j\in C_n} e^{V_{jn}}}\right)\right) = 0$$

$$\sum_{n=1}^{N}\sum_{i\in C_n}\left[y_{in} - P_n\left(i \mid x_n, \hat{\beta}\right)\right] x_{ink} = 0$$

Think about the constant xink = 1
and sum_n yin = sum_n Pni
meaning
<u>the observed shares in the estimation sample is equal to the predicted shares</u>
Since sum_iCn yin=1 .
This will always be true for logit.

Simple Model – Null Model

$U_i = \varepsilon_i$, *for all i* (β=0)

$$\mathrm{P}\left(i \mid C_n\right) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} = \frac{e^0}{\sum_{j \in C_n} e^0} = \frac{1}{\# C_n}$$

$$L(0) = \sum_n \ln \frac{1}{\# C_n} = -\sum_n \ln(\# C_n)$$

Simple Model – Constants Only

Assume $C_n = C$

$$\mathrm{P}\left(i \mid C_n\right) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} = \frac{e^{c_i}}{\sum_j e^{c_j}}$$

$U_i = c_i + \varepsilon_i$, for all i

$$\ln P(i) = c_i - \ln \sum_j e^{c_j}$$

In the sample of size N, there are N_i persons choosing alternative i

The log likelihood for all people choosing i is $L_i = N_i c_i - N_i \ln \sum_j e^{c_j}$

Simple Model – Constants Only (cont.)

- Total log likelihood $L(c) = \sum_i N_i c_i - N \ln \sum_j e^{c_j}$

- At the maximum, the derivatives must be zero

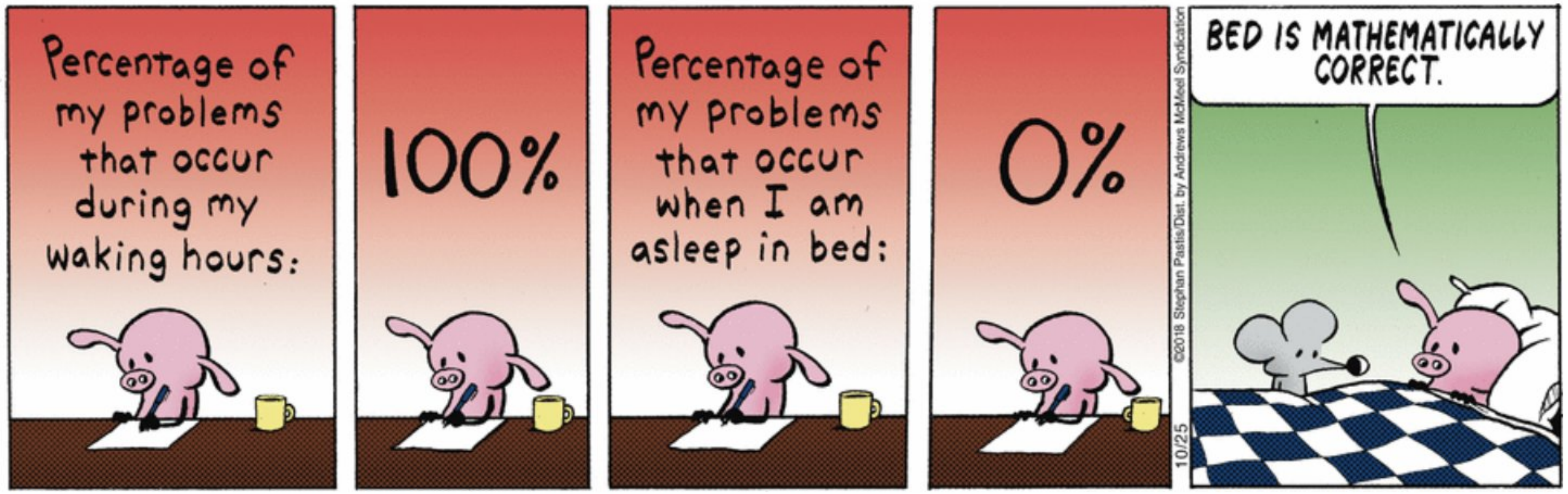
$$\frac{\partial L}{\partial c_i} = N_i - N \frac{e^{c_i}}{\sum_j e^{c_j}} = N_i - NP(i) = 0$$

$$P(i) = \frac{N_i}{N}$$

With a constants only model, the probability of an alternative (i) is equal to the share of people in the estimation sample who chose that alternative (i). Always true for logit.

Outline

- Theory
  - Review: RUM and Binary Choice
  - Extension to more than 2 alternatives
  - Probit
  - Logit
  - Properties of Logit
- Application
  - Specification of Vs
  - Boeing SP Model Example
- Maximum Likelihood Estimation


Percentage of my problems that occur during my waking hours:
100%
Percentage of my problems that occur when I am asleep in bed:
0%
©2018 Stephan Pastis/Dist. by Andrews McMeel Syndication
10/25
BED IS MATHEMATICALLY CORRECT.