Bayesian Inference on the Logit Demand Model

Yudong Cao
Bayesian Final Presentation
4/20/2016

The Logit Demand Model

- Mathematical Modeling in Economics (Steven Tschantz, 2015)
- Derived by *Mathematica*

$$\pi_i(p_1, \ldots, p_n) = \frac{\exp((\eta_i - p_i)/\lambda)}{1+\exp((\eta_1 - p_1)/\lambda)+\ldots+\exp((\eta_n - p_n)/\lambda)}$$

- π_i is the choice probability for product *i*
- p_i is the price for product *i* (inside good *1, …, n*)
- η_i is the mean consumer values for product *i*
- λ is a scale parameter in proportion to the standard deviation of consumer values

The Logit Demand Model

- M is the total consumer population or market size
- q_i is the demand for product i (inside good)

$$q_i(p_1, \ldots, p_n) = M\pi_i(p_1, \ldots, p_n) = \frac{M\exp((\eta_i - p_i)/\lambda)}{1+\exp((\eta_1 - p_1)/\lambda)+\ldots+\exp((\eta_n - p_n)/\lambda)}$$

- Outside good:

The choice probability and demand for the outside good are

$$\pi_0(p_1, \ldots, p_n) = \frac{1}{1+\exp((\eta_1 - p_1)/\lambda)+\ldots+\exp((\eta_n - p_n)/\lambda)}$$

and

$$q_0(p_1, \ldots, p_n) = M\pi_0(p_1, \ldots, p_n) = \frac{M}{1+\exp((\eta_1 - p_1)/\lambda)+\ldots+\exp((\eta_n - p_n)/\lambda)}$$

Variables & Parameters

$$\pi_i(p_1, \ldots, p_n) = \frac{\exp((\eta_i - p_i)/\lambda)}{1+\exp((\eta_1 - p_1)/\lambda)+\ldots+\exp((\eta_n - p_n)/\lambda)}$$

- π_i → outcome variable
- p_i → independent variable
- *Parameters to be estimated:*
  - η_i the mean consumer values for product *i*
  - λ the scale parameter in proportion to the standard deviation of consumer values

Problem

Use maximum-likelihood to estimate a logit demand model for three products from the following individual choice data for specific prices. Here the *i*-th element of data is a 4-tuple, the first three entries being the prices of three products, and the last being an integer *j* in the range 1 to 3 if the consumer picks product *j* and $j = 4$ if the consumer chooses no purchase. The parameters of the model are eta1, eta2, eta3, and lambda. The probability of the consumer choosing $j = 1$ is

$$\frac{e^{(eta1-p1)/lambda}}{e^{(eta1-p1)/lambda}+e^{(eta2-p2)/lambda}+e^{(eta3-p3)/lambda}+e^{(eta4-p4)/lambda}}$$

where eta4=p4=0, and similarly for other *j*.

Dataset

- N=200
- 4-tuple:

  ($p1$, $p2$, $p3$, X)
- p_i is the price for product i (inside good 1, 2, 3)
- X is the product chosen (indicator)
- X=0: outside product chosen

```
data = {{11.23, 14.37, 8.33, 3}, {13.33, 9.32, 8.4, 3},
   {11.52, 10., 8.81, 3}, {11.42, 15.71, 10.74, 4}, {11.18, 10.41, 12.91, 3},
   {8.26, 14.16, 11.12, 1}, {11.94, 8.06, 11.77, 2}, {8.75, 9.95, 13.97, 1},
   {15.94, 9.49, 10.27, 3}, {9.68, 14.97, 15.01, 4}, {10.8, 12.53, 11.48, 4},
   {14.26, 14.88, 10.66, 3}, {10.87, 12.52, 15.1, 4}, {10.46, 14.63, 8.59, 3},
   {14.45, 14.11, 12.23, 4}, {13.99, 14.74, 13.16, 4}, {12.03, 9.22, 11.07, 3},
   {8.57, 13.57, 12.49, 1}, {13.4, 14.49, 13.5, 4}, {12.03, 13.8, 10.22, 3},
   {8.94, 11.14, 14.81, 1}, {11.08, 10.96, 9.32, 3}, {13.99, 13.25, 10.55, 3},
   {14.08, 11.45, 11.17, 3}, {14.61, 11.08, 8.21, 3}, {14.67, 8.6, 9.92, 2},
   {13.28, 12.36, 12.44, 3}, {12.31, 14.25, 9.97, 3}, {12.81, 10.64, 14.97, 3},
   {9.07, 11.4, 11.19, 1}, {11.79, 14.23, 12.99, 4}, {8.77, 15.43, 15.47, 1},
   {12.24, 9.23, 12.84, 2}, {11.88, 15.45, 9.76, 3}, {10.52, 9.2, 11.03, 3},
   {8.61, 14.93, 11.67, 3}, {11.09, 11.51, 13.45, 3}, {13.65, 8.71, 10.26, 3},
   {9.81, 14.25, 15.18, 1}, {15.17, 10.16, 11.21, 3}, {12.17, 15.98, 13.98, 4},
   {11.79, 15.57, 12.06, 1}, {11.38, 15.53, 13.93, 4}, {12.15, 8.72, 12.91, 2},
   {13.69, 13.62, 8.99, 3}, {9.32, 11.96, 13.26, 1}, {10.32, 12.15, 10.62, 3},
   {9.54, 11.28, 15.01, 4}, {14.28, 11.07, 8.31, 3}, {13.21, 11.92, 13.61, 4},
   {13.67, 15.32, 14.52, 4}, {10.79, 11.95, 14.59, 4}, {14.78, 9.81, 10.53, 3},
   {13.75, 11.61, 14.69, 4}, {10.8, 12.68, 11.07, 3}, {15.7, 11.63, 12.54, 4},
   {12.29, 9.69, 15.2, 2}, {13.8, 14.36, 8.6, 3}, {10.43, 9.04, 13.86, 2},
   {9.25, 15.3, 15.87, 4}, {13., 13.04, 12.22, 3}, {9.19, 8.88, 14.73, 4},
   {8.15, 10.14, 9.56, 3}, {14.84, 14.2, 10.03, 3}, {12.71, 13.06, 15.66, 4},
   {12.74, 13.13, 9.59, 3}, {11.92, 10.22, 15.66, 2}, {12.19, 11.13, 12.23, 2},
   {9.57, 11.24, 10.04, 3}, {11.22, 10.9, 10.11, 3}, {10.55, 8.52, 15.86, 2},
   {8.86, 8.73, 13.18, 2}, {8.09, 12.15, 13.62, 1}, {9.41, 15.82, 12.59, 1},
   {11.61, 9.99, 9.76, 3}, {13.32, 13.89, 14.7, 4}, {13.45, 10.14, 12.66, 3},
   {15.93, 11.28, 12.97, 2}, {9.9, 13.73, 8.09, 3}, {9.77, 13.91, 13.37, 1},
   {15.01, 15.98, 8.68, 3}, {9.77, 8.2, 8.44, 3}, {14.69, 15.02, 9.56, 3},
   {11.83, 12.3, 10.7, 3}, {9.12, 10.66, 9.07, 3}, {11.28, 12.71, 10.33, 3},
   {12., 10.03, 9.95, 3}, {12.72, 9.9, 11.92, 2}, {9.52, 9.84, 14.87, 2},
   {11.15, 9.42, 9.35, 3}, {13.52, 8.06, 13.54, 2}, {12.59, 9.7, 13.9, 2},
   {11.72, 15.27, 8.42, 3}, {9.23, 9.37, 14.11, 2}, {13.7, 11.4, 11.34, 3},
   {11.58, 15.55, 14.56, 4}, {13.73, 15.78, 8.26, 3}, {12.98, 12.75, 14.91, 4},
   {12.27, 11.7, 8.23, 3}, {9.9, 9.98, 15.05, 2}, {14.85, 13.96, 13.88, 4},
   {12.71, 9.43, 8.33, 3}, {9.84, 13.09, 14.88, 4}, {11.89, 10.17, 12.58, 2},
   {12.82, 14.56, 15.8, 4}, {12.42, 15.73, 11.18, 3}, {14.34, 14.14, 15.48, 4},
   {13.27, 11.42, 15.37, 4}, {9.79, 11.28, 13.43, 2}, {14.02, 14.67, 11.23, 3},
   {9.76, 11.51, 13.87, 3}, {10.62, 10.23, 13.7, 2}, {15.71, 15.05, 9.03, 3},
   {14.63, 13.06, 11.8, 3}, {8.01, 15.52, 11.65, 1}, {12.76, 12.1, 10.04, 3},
   {13.21, 10.12, 11.5, 3}, {10.5, 14.49, 13.64, 1}, {8.91, 13.12, 11.02, 3},
   {15.65, 14.58, 14.56, 4}, {13.01, 9.98, 8.65, 3}, {10.69, 15.02, 12.62, 4},
```

Maximum Likelihood in *Mathematica*

1. Define a function llh1[{p1_,p2_,p3_,j_}]:=... that evaluates to the log-likelihood of a consumer choosing *j*.

```
prob1[p1_, p2_, p3_] := Exp[(eta1 - p1) / lambda] /
   (1 + Exp[(eta1 - p1) / lambda] + Exp[(eta2 - p2) / lambda] + Exp[(eta3 - p3) / lambda])
prob2[p1_, p2_, p3_] := Exp[(eta2 - p2) / lambda] /
   (1 + Exp[(eta1 - p1) / lambda] + Exp[(eta2 - p2) / lambda] + Exp[(eta3 - p3) / lambda])
prob3[p1_, p2_, p3_] := Exp[(eta3 - p3) / lambda] /
   (1 + Exp[(eta1 - p1) / lambda] + Exp[(eta2 - p2) / lambda] + Exp[(eta3 - p3) / lambda])
prob4[p1_, p2_, p3_] :=
  1 / (1 + Exp[(eta1 - p1) / lambda] + Exp[(eta2 - p2) / lambda] + Exp[(eta3 - p3) / lambda])

prob1[p1, p2, p3] + prob2[p1, p2, p3] + prob3[p1, p2, p3] + prob4[p1, p2, p3] // Simplify

1

llh1[{p1_, p2_, p3_, j_}] :=
  Log[{prob1[p1, p2, p3], prob2[p1, p2, p3], prob3[p1, p2, p3], prob4[p1, p2, p3]}][[j]]

llh1[{12, 14, 16, 1}]
```

$$\mathrm{Log}\left[\frac{e^{\frac{-12+\mathrm{eta1}}{\mathrm{lambda}}}}{1+e^{\frac{-12+\mathrm{eta1}}{\mathrm{lambda}}}+e^{\frac{-14+\mathrm{eta2}}{\mathrm{lambda}}}+e^{\frac{-16+\mathrm{eta3}}{\mathrm{lambda}}}}\right]$$

Maximum Likelihood in *Mathematica*

2. Define llh to be the average log-likelihood for the data set, applying llh1 to each element of data and averaging.

```
llh = 1 / Length[data] * Sum[llh1[data[[i]]], {i, 1, Length[data]}]
```

$$\frac{1}{200}\left(\text{Log}\left[\frac{1}{1+e^{\frac{-9.78+\text{eta1}}{\text{lambda}}}+e^{\frac{-12.11+\text{eta2}}{\text{lambda}}}+e^{\frac{-15.93+\text{eta3}}{\text{lambda}}}}\right]+\text{Log}\left[\frac{e^{\frac{-8.13+\text{eta2}}{\text{lambda}}}}{1+e^{\frac{-12.08+\text{eta1}}{\text{lambda}}}+e^{\frac{-8.13+\text{eta2}}{\text{lambda}}}+e^{\frac{-15.88+\text{eta3}}{\text{lambda}}}}\right]+\right.$$

$$\text{Log}\left[\frac{1}{1+e^{\frac{-9.25+\text{eta1}}{\text{lambda}}}+e^{\frac{-15.3+\text{eta2}}{\text{lambda}}}+e^{\frac{-15.87+\text{eta3}}{\text{lambda}}}}\right]+\text{Log}\left[\frac{e^{\frac{-8.52+\text{eta2}}{\text{lambda}}}}{1+e^{\frac{-10.55+\text{eta1}}{\text{lambda}}}+e^{\frac{-8.52+\text{eta2}}{\text{lambda}}}+e^{\frac{-15.86+\text{eta3}}{\text{lambda}}}}\right]+$$

...... (200 Log[XXX]'s, 5 pages long)

Maximum Likelihood in *Mathematica*

- Results: (no standard errors)

3. Find the parameters that maximize llh.

```
FindMaximum[llh, {{eta1, 5}, {eta2, 5}, {eta3, 5}, {lambda, 4}}]
{-0.468355, {eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}}
```

Bayesian Inference

- Polytomous Nominal Response Model
  - 1 item
  - 4 choices
    - 1, 2, 3 are inside products
    - 4 is outside product
  - → WinBUGS!

WinBUGS

$$\pi_i(p_1, \ \ldots, p_n) = \frac{\exp((\eta_i - p_i)/\lambda)}{1+\exp((\eta_1-p_1)/\lambda)+\ldots+\exp((\eta_n-p_n)/\lambda)}$$

with indicator function Equals(x[j], i)

```
#############
### Model  ###
#############

model{
        for (j in 1:n) {
                pr1[j]<-(exp(eta1-p1[j])/lambda)/((1+exp(eta1-p1[j])/lambda)+(exp(eta2-p2[j])/lambda)+(exp(eta3-p3[j])/
lambda))*equals(x[j],1)

                pr2[j]<-(exp(eta2-p2[j])/lambda)/((1+exp(eta1-p1[j])/lambda)+(exp(eta2-p2[j])/lambda)+(exp(eta3-p3[j])/
lambda))*equals(x[j],2)

                pr3[j]<-(exp(eta3-p3[j])/lambda)/((1+exp(eta1-p1[j])/lambda)+(exp(eta2-p2[j])/lambda)+(exp(eta3-p3[j])/
lambda))*equals(x[j],3)

                pr4[j]<-1/((1+exp(eta1-p1[j])/lambda)+(exp(eta2-p2[j])/lambda)+(exp(eta3-p3[j])/lambda))*equals(x[j],4)
        }

# priors
eta1 ~ dnorm(10,0.1)
eta2 ~ dnorm(10,0.1)
eta3 ~ dnorm(10,0.1)
lambda ~ dnorm(0,0.001) I(0,)
```

EB: Informative

- eta1 ~ dnorm(10,0.1)
- eta2 ~ dnorm(10,0.1)
- eta3 ~ dnorm(10,0.1)
- lambda ~ dnorm(0,0.01)I(0,)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
eta1	9.991	3.137	0.03605	3.691	9.969	16.15	3000	6003
eta2	9.977	3.158	0.03837	3.716	10.01	16.21	3000	6003
eta3	9.986	3.189	0.03763	3.732	9.981	16.27	3000	6003
lambda	7.924	6.009	0.08414	0.3059	6.667	22.48	3000	6003

{eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}

EB: Noninformative

- eta1 ~ dnorm(10,0.001)
- eta2 ~ dnorm(10,0.001)
- eta3 ~ dnorm(10,0.001)
- lambda ~ dnorm(0,0.001) I(0,)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
eta1	9.947	31.36	0.3597	-53.01	9.69	71.67	3000	6003
eta2	9.767	31.58	0.3817	-52.84	10.14	72.1	3000	6003
eta3	9.873	31.89	0.3751	-52.68	9.87	72.73	3000	6003
lambda	25.06	18.98	0.2615	0.9674	21.09	70.97	3000	6003

{eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}

HB: Informative

- eta1 ~ dnorm(10,tau)
- eta2 ~ dnorm(10,tau)
- eta3 ~ dnorm(10,tau)
- lambda ~ dnorm(0,0.001) I(0,)
- tau ~ dgamma(.5,.5)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
eta1	4.709	800.4	5.952	-2.391	9.988	22.47	4000	18003
eta2	7.734	298.1	2.238	-2.676	10.0	22.41	4000	18003
eta3	-7.612	2437.0	18.11	-1.941	9.995	22.04	4000	18003
lambda	25.14	18.95	0.1492	1.007	21.27	70.5	4000	18003

```
{eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}
```

HB: Noninformative

- eta1 ~ dnorm(10,tau)
- eta2 ~ dnorm(10,tau)
- eta3 ~ dnorm(10,tau)
- lambda ~ dnorm(0,0.001) I(0,)
- tau ~ dgamma(.1,.1)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
eta1	-2.315E+18	3.617E+20	2.355E+18	-659200.0	10.03	908300.0	7000	24003
eta2	-6.407E+17	8.461E+19	5.504E+17	-914700.0	9.992	615700.0	7000	24003
eta3	-1.218E+14	1.673E+16	1.077E+14	-717700.0	10.01	8.26E+5	7000	24003
lambda	25.13	19.11	0.1302	0.9815	20.93	71.08	7000	24003

{eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}

Summary

	eta1	eta2	eta3	lambda
EB-informative	9.97 (3.69, 16.15)	10.01 (3.16, 16.21)	9.98 (3.73, 16.27)	6.67 (0.31, 22.48)
EB-noninformative	9.69 (-53.01, 71.67)	10.14 (-52.84, 72.1)	9.87 (-52.68, 72.73)	21.09 (0.97, 70.97)
HB-informative	9.99 (-2.39, 22.47)	10 (-2.68, 22.41)	9.99 (-1.94, 22.04)	21.27 (1.01, 70.5)
HB-noninformative	10.03	9.99	10.01	20.93 (0.98, 71.08)
ML (*Mathematica*)	9.9	10.67	12.77	0.61

- *WinBUGS* Bayesian inference results
  - “intuitive & realistic”
  - EB-informative: best results
  - highly dependent on priors
  - highly variable
  - Lambda very different (proportional to sd’s of etas)

Summary

- *Mathematica* maximum likelihood results speak for the data better

```
{eta1 → 9.89676, eta2 → 10.6662, eta3 → 12.7689, lambda → 0.613368}
```

Product	Count	Pr
1	25	0.125
2	33	0.165
3	99	0.495
4	43	0.215

- The model can be generalized to any combination of choices & products
  - Maybe Bayesian methods will work better when the model becomes too complicated

References

- Steven Tschantz (2015)
  - 09.1-LogitDemandModel
  - 22.1-MaximumLikelihood
  - 22.2-Assignment11

  - Dr. Cho's help on WinBUGS code