What is a hierarchical choice model?

CHOICE MODELING FOR MARKETING IN R

Elea McDonnell Feit
Assistant Professor of Marketing, Drexel University

Heterogeneity in preferences

het·er·o·ge·ne·i·ty

/ˌhedərəjəˈnēədē,ˌhedərəjəˈnāədē/

noun

noun: **heterogeneity**

> the quality or state of being diverse in character or content.
> "the genetic heterogeneity of human populations"

Origin

mid 17th century: from medieval Latin *heterogeneitas*, from *heterogeneus* (see heterogeneous) + -ity.

Hierarchical choice models (random coefficients models)

```
for (i in 1:n_resp) {
  beta[i] <- mvrnorm(1, beta_0, Sigma)  # Random normal vector
  for (j in 1:n_task[i]) {
    X <- X[X$resp == i & X$task == j, ]
    u <- X %*% beta[i]
    p[i,] <- exp(u) / sum(exp(u))
  }
}
```

Fitting a hierarchical multinomial logit model

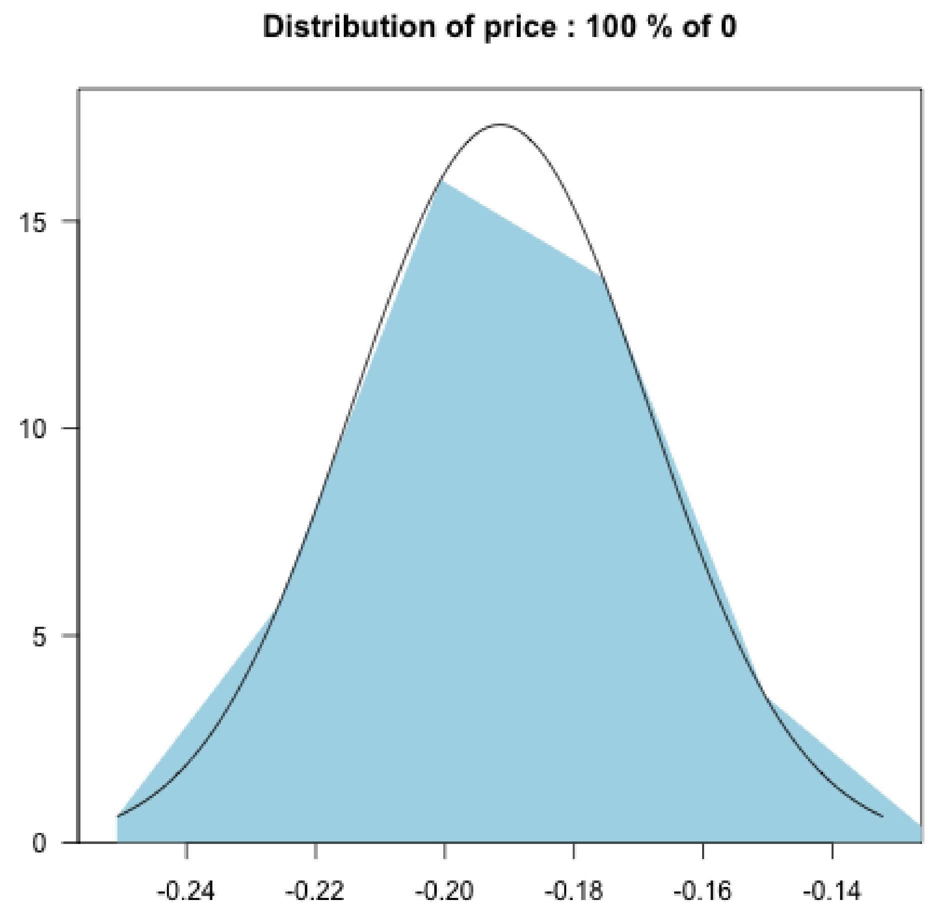
```
sportscar <- mlogit.data(sportscar,
                         choice = "choice",
                         shape = "long",
                         varying = 5:8,
                         alt.var = "alt",
                         id.var = "resp_id")
m7 <- mlogit(choice ~ 0 + seat + trans + convert + price,
             data = sportscar,
             rpar = c(price = "n"),
             panel = TRUE)
```

```r
summary(m7)
```

```
...
Coefficients :
              Estimate Std. Error  z-value  Pr(>|z|)
seat4       -0.0185815  0.0762964  -0.2435 0.8075843
seat5        0.4259317  0.0751681   5.6664 1.458e-08 ***
transmanual -1.2206527  0.0650133 -18.7754 < 2.2e-16 ***
convertyes   0.2013760  0.0603982   3.3341 0.0008556 ***
price       -0.1914656  0.0092325 -20.7382 < 2.2e-16 ***
sd.price     0.0230365  0.0327214   0.7040 0.4814209

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log-Likelihood: -1709.8

random coefficients
      Min.    1st Qu.     Median       Mean    3rd Qu. Max.
price -Inf -0.2070035 -0.1914656 -0.1914656 -0.1759277  Inf
...
```

Distribution of the price coefficient

```
plot(m7)
```



Let's practice!

CHOICE MODELING FOR MARKETING IN R

Heterogeneity in preference for other features

CHOICE MODELING FOR MARKETING IN R

Elea McDonnell Feit
Assistant Professor of Marketing, Drexel University

A different way to code factors

Dummy coding (what we've been doing)

```
  seat4 seat5
2     0     0
4     1     0
5     0     1
```

Effects coding (better for hierarchical models)

```
   seat4  seat5
2     -1     -1
4      1      0
5      0      1
```

Changing the coding for a factor

```
contrasts(sportscar$seat) <- contr.sum(levels(sportscar$seat))
```

```
dimnames(contrasts(sportscar$seat))[[2]] <- levels(sportscar$seat)[1:2]
```

```
contrasts(sportscar$seat)
```

```
   4  5
2 -1 -1
4  1  0
5  0  1
```

Making all the coefficients heterogeneous

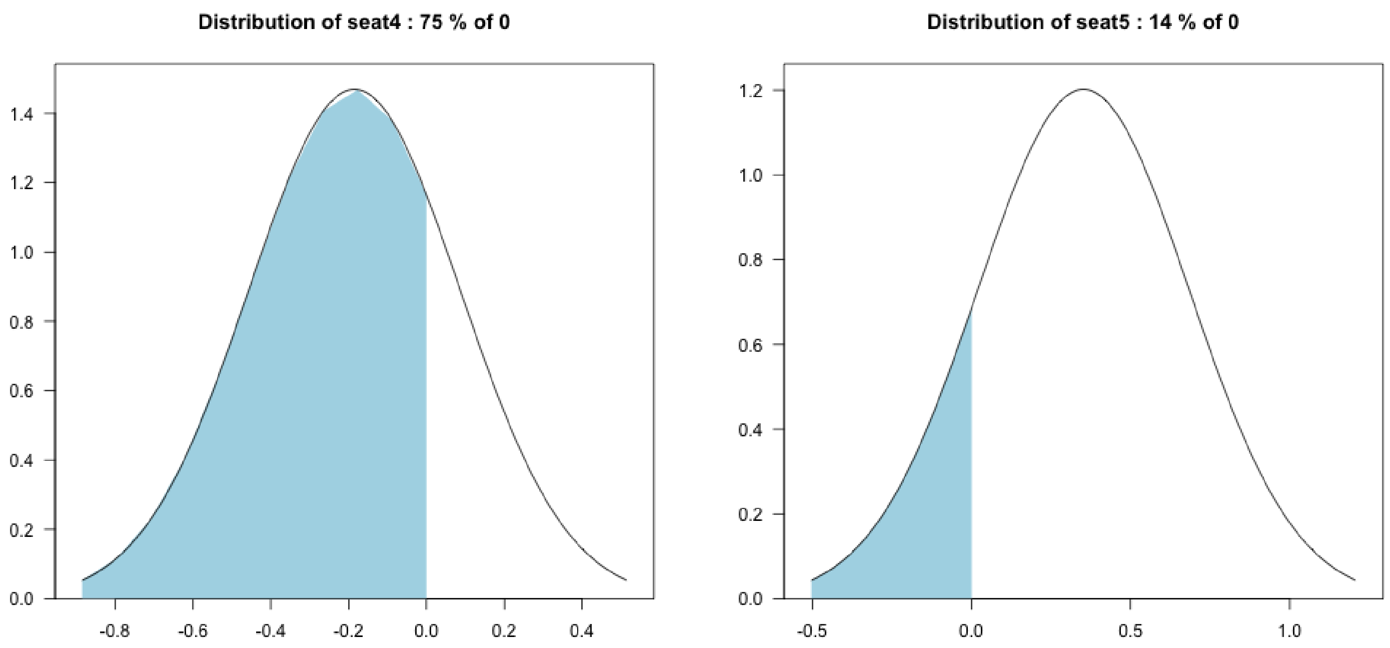
```
my_rpar <- c("n", "n", "n", "n", "n")
m3 <- mlogit(choice ~ 0 + seat + trans + convert + price,
             data = sportscar)
names(my_rpar) <- names(m3$coefficients)
my_rpar
```

```
seat4        seat5 transmanual  convertyes       price
  "n"          "n"         "n"         "n"         "n"
```

```
m8 <- mlogit(choice ~ 0 + seat + trans + convert + price,
             data = sportscar,
             panel = TRUE, rpar = my_rpar)
```

Hierarchical model parameters

```
m8 <- mlogit(choice ~ 0 + seat + trans + convert + price,
              data = sportscar,
             panel = TRUE, rpar = my_rpar)
plot(m8, par = c("seat4", "seat5"))
```



Coefficient for the base level

```
m8$coef[1:2]
```

```
     seat4      seat5
-0.1852167  0.3519204
```

```
-sum(m8$coef[1:2])
```

```
-0.1667037
```

Let's try it with the `chocolate` data!

CHOICE MODELING FOR MARKETING IN R

Predicting shares with hierarchical models

CHOICE MODELING FOR MARKETING IN R

Elea McDonnell Feit
Assistant Professor of Marketing, Drexel University

Hierarchical model with correlations

```
m10 <- mlogit(choice ~ 0 + seat + trans + convert + price,
              data = sportscar,
              rpar = myrpar, panel = TRUE, correlation = TRUE)
cor.mlogit(m10)
```

```
                 seat4      seat5 transmanual convertyes      price
seat4        1.0000000 -0.3411867   0.1584436 -0.3129433  0.1551497
seat5       -0.3411867  1.0000000  -0.1124484  0.1187094 -0.3206838
transmanual  0.1584436 -0.1124484   1.0000000 -0.6231883  0.7710748
convertyes  -0.3129433  0.1187094  -0.6231883  1.0000000 -0.1165536
price        0.1551497 -0.3206838   0.7710748 -0.1165536  1.0000000
```

Products we want to predict shares for

```
prod
```

```
  seat  trans convert price
1    2 manual      no    35
2    2   auto      no    30
```

```
prod.coded
```

```
  seat4 seat5 transmanual convertyes price
1    -1    -1           1          0    35
2    -1    -1           0          0    30
```

Share prediction for hierarchical model

```
mean <- m10$coef[1:5]   # Hard coded
Sigma <- cov.mlogit(m10)
share <- matrix(NA, nrow = 1000, ncol = nrow(prod.coded))
for (i in 1:1000) {
  coef <- mvrnorm(1, mu = mean, Sigma = Sigma)
  utility <- prod.coded %*% coef
  share[i,] <- exp(utility) / sum(exp(utility))
}
cbind(colMeans(share), prod)
```

```
  colMeans(share) seat  trans convert price segment
1       0.1238315    2 manual      no    35   basic
2       0.8761685    2   auto      no    30   basic
```

Let's practice!

CHOICE MODELING FOR MARKETING IN R

Goodbye and good luck!

CHOICE MODELING FOR MARKETING IN R

Elea McDonnell Feit

Assistant Professor of Marketing, Drexel University

Choices in building models

- Which attributes to include
- Treating numeric attributes as factors
- Interactions between attributes
- Interactions between attributes and decision-maker characteristics
- Hierarchical models
- Correlations between coefficients

Other choice model features

- Distributions of random coefficients
- Probit models
- Nested logit
- Bayesian choice models (using the `bayesm` package or Stan)

Advice for building models

- Always start by computing choice counts to summarize the data
- Build up from simple models to more complex
- If estimated parameters have very large standard errors, then you've probably added too much model complexity. Back up to a simpler model.
- For models describing human behavior, heterogeneity is usually a good idea

Go fit some choice models!

CHOICE MODELING FOR MARKETING IN R