Analysis of Conjoint Data
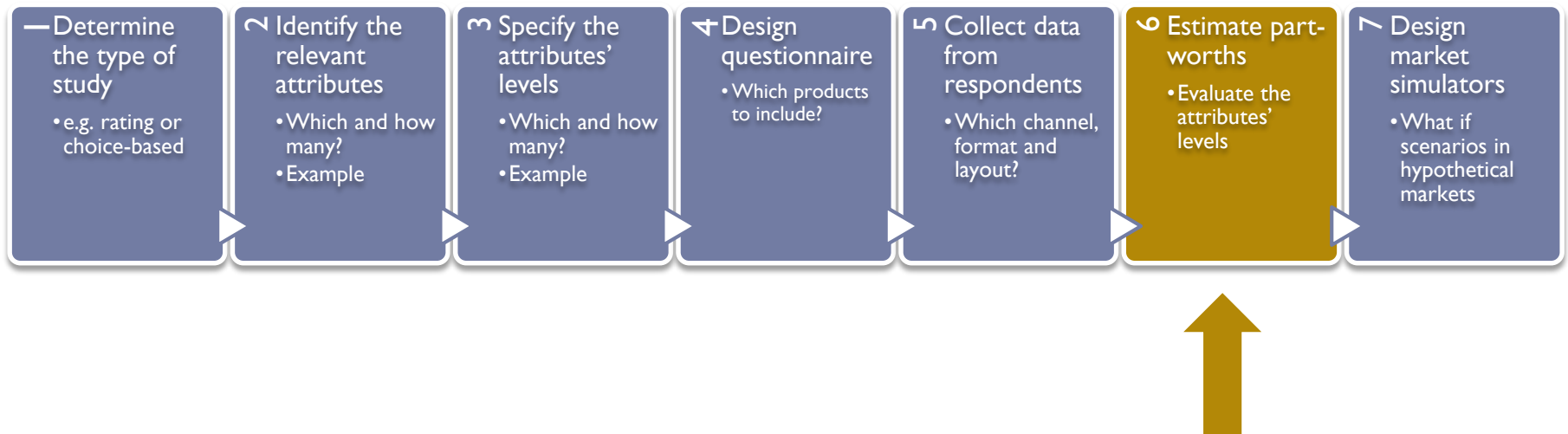# Part IV: Hierarchical Bayes

Aurélie Lemmens

What Have We Learned Last Week

‣ **3 methods to analyze the CBC data**
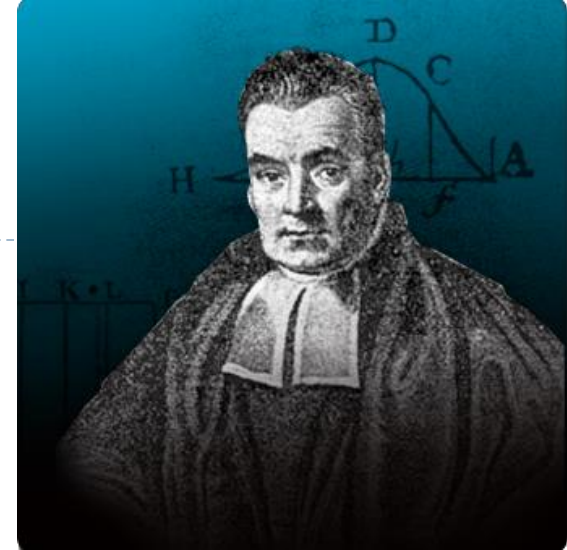
 ‣ Counting Analysis

 ‣ Logit

 ‣ Latent Class Analysis

1 Determine the type of study	2 Identify the relevant attributes	3 Specify the attributes' levels	4 Design questionnaire	5 Collect data from respondents	6 Estimate part-worths	7 Design market simulators
• e.g. rating or choice-based	• Which and how many? • Example	• Which and how many? • Example	• Which products to include?	• Which channel, format and layout?	• Evaluate the attributes' levels	• What if scenarios in hypothetical markets

Hierarchical Bayesian Analysis

Hierarchical Bayes

Thomas Bayes (1701-1763)

▸ Latent Class: segment-specific preferences

$$p_s(i|C) = \frac{exp(U_{is})}{\sum_{j=1}^{m} exp(U_{js})} = \frac{exp(x_i\beta_s)}{\sum_{j=1}^{m} exp(x_j\beta_s)}$$

▸ HB: individual-specific preferences: *k = respondent*

Assumes that the probability than an individual will choose one of the *m* alternatives *i* from the choice set *C* is:

$$p_k(i|C) = \frac{exp(U_{ik})}{\sum_{j=1}^{m} exp(U_{jk})} = \frac{exp(x_i\boldsymbol{\beta_k})}{\sum_{j=1}^{m} exp(x_j\boldsymbol{\beta_k})} \implies \beta_k \sim N(\beta_0, \sigma)$$

LCA vs HB

- LCA uses the information about all respondents

- LCA identifies customer segments & *segments are easier to target*

- If the customer base is truly segmented in groups, LCA is more accurate.

- Choice simulators are less accurate than with HB.

- HB uses the information about all respondents

- HB allows for *differences between all respondents*

- HB can be used in combination with *clustering* but it is less efficient than LCA

- If preferences are more continuously distributed, HB is more accurate

- Choice simulators are more accurate than with LCA.

Thought Experiment

▸ I pick a random number between 0 and 10 and you have to guess whether it is higher than 5

  ▸ Your guess?

▸ I pick a second number between 0 and 10 and inform you it is higher than the first number I picked.

  ▸ Your guess?

▸ I pick a third number between 0 and 10 and inform you it is higher than the first number I picked.

  ▸ Your guess?

Original number: **?**

Thought Experiment

▸ I pick a random number between 0 and 10 and you have to guess whether it is higher than 5

  ▸ Your guess? Equal chance   *Prior belief*

▸ I pick a second number between 0 and 10 and inform you it is higher than the first number I picked.   *New information*

  ▸ Your guess? Updated belief decreases

▸ I pick a third number between 0 and 10 and inform you it is higher than the first number I picked.   *New information*

  ▸ Your guess? Updated belief decreases even more

Original number: **?**

Mechanism: Bayesian Updating

▸ In each new round of belief updating, the most recent posterior becomes the prior for the new calculation.

Prior Belief → New Data → Improved Belief

Estimating Customer Preferences

β? Path-worth of a customer (preference)

Data: probability of observing a
choice given the preferences

$$P(\boldsymbol{\beta}|\mathbf{y}) = \frac{P(\mathbf{y}|\boldsymbol{\beta}) \times P(\boldsymbol{\beta})}{P(\mathbf{y})}$$

Posterior: updated guess
about preferences after
observing choice

Prior: guess about preferences
before observing choice

Iterative Process

$$\beta_k \sim N(\beta_0, \sigma)$$

Sawtooth Software CBC/HB - cbc
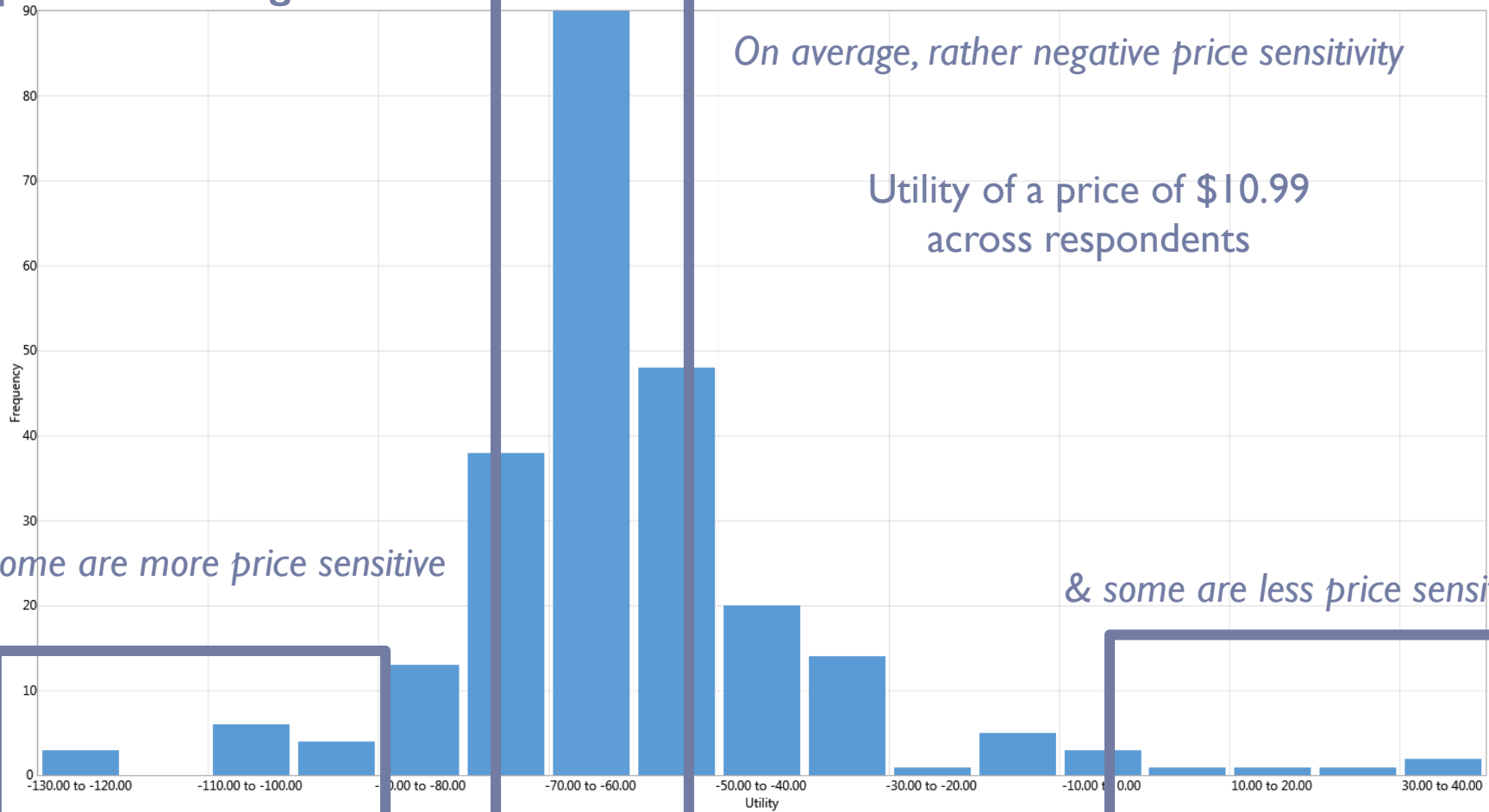
Preliminary iterations		10000
Draws used per respondent		10000
Total iterations		20000
Number of respondents		250
Parameters per respondent		9

No covariates in use
No constraints in use
Random draws not saved

Iteration		3987
	Current	Average
Pct. Cert.	0.386	0.381
RLH	0.427	0.424
Avg Variance	1.486	1.475
Parameter RMS	1.372	1.372
Iterations/Sec		166.13
Time Elapsed		0:00:24
Time Remaining		0:01:36

Parameter Estimates

Mean Beta

Iterations   20000

Burn-in

| 0.86 | 0.59 | -0.61 | -0.84 | -0.72 | 0.16 | 0.56 | 1.09 | 0.25 | -0.09 |
| -1.25 | -0.78 | | | | | | | | |

Stop estimation

Sample (Average) Utilities

$$\beta_0 \qquad \qquad \sigma$$

Average Utilities (Zero-Centered Diffs)	Average Utilities	Standard Deviation
High-Flyer Pro, by Smith and Forester	41.01852	26.03931
Magnum Force, by Durango	32.01098	24.75692
Eclipse+, by Golfers, Inc.	-31.64712	29.40097
Long Shot, by Performance Plus	-41.38238	20.71599
Drives 5 yards farther than the average ball	-38.20113	16.16097
Drives 10 yards farther than the average ball	11.80400	15.87702
Drives 15 yards farther than the average ball	26.39712	19.40170
$4.99 for package of 3 balls	51.93810	19.56939
$6.99 for package of 3 balls	15.22258	20.62804
$8.99 for package of 3 balls	-6.11965	18.02962
$10.99 for package of 3 balls	-61.04103	21.95092
NONE	-118.64638	236.55102
Average Importances	**Average Importances**	**Standard Deviation**
Brand:	35.10127	9.30883
Performance:	24.64572	6.81803
Price:	40.25301	7.75260

Posterior Distribution

Population average = -61.04

On average, rather negative price sensitivity

Utility of a price of $10.99 across respondents

But some are more price sensitive

& some are less price sensitive

Attribute Importances (HB)

Scaling	Zero-Centered Differences	Bin Selection	Fixed Width
Precision	2	Bin width	10.0
✓ Display importances			Graph Panel Options

ACTIONS — Refresh, Save To Excel
SEGMENTATION — Segments, Weights
OPTIONS
HISTOGRAMS

	A	B	C
1	**Average Utility Values**		
2	Utility Scaling Method	Zero-Centered Differences	
3	Respondent Count	250	
4			
5	**Label**	**Utility**	**Std Deviation**
6	High-Flyer Pro, by Smith and Forester	41.02	26.04
7	Magnum Force, by Durango	32.01	24.76
8	Eclipse+, by Golfers, Inc.	-31.65	29.40
9	Long Shot, by Performance Plus	-41.38	20.72
10			
11	Drives 5 yards farther than the average ball	-38.20	16.16
12	Drives 10 yards farther than the average ball	11.80	15.88
13	Drives 15 yards farther than the average ball	26.40	19.40
14			
15	$4.99 for package of 3 balls	51.94	19.57
16	$6.99 for package of 3 balls	15.22	20.63
17	$8.99 for package of 3 balls	-6.12	18.03
18	$10.99 for package of 3 balls	-61.04	21.95
19			
20	None	-118.65	236.55
21			
22	**Average Importances**		
23			
24	**Attribute**	**Importance**	**Std Deviation**
25	Brand:	35.10	9.31
26	Performance:	24.65	6.82
27	Price:	40.25	7.75

Attribute Importances

HB

Brand: / P... / Price:

Attribute Importances (logit)

Individual Importances

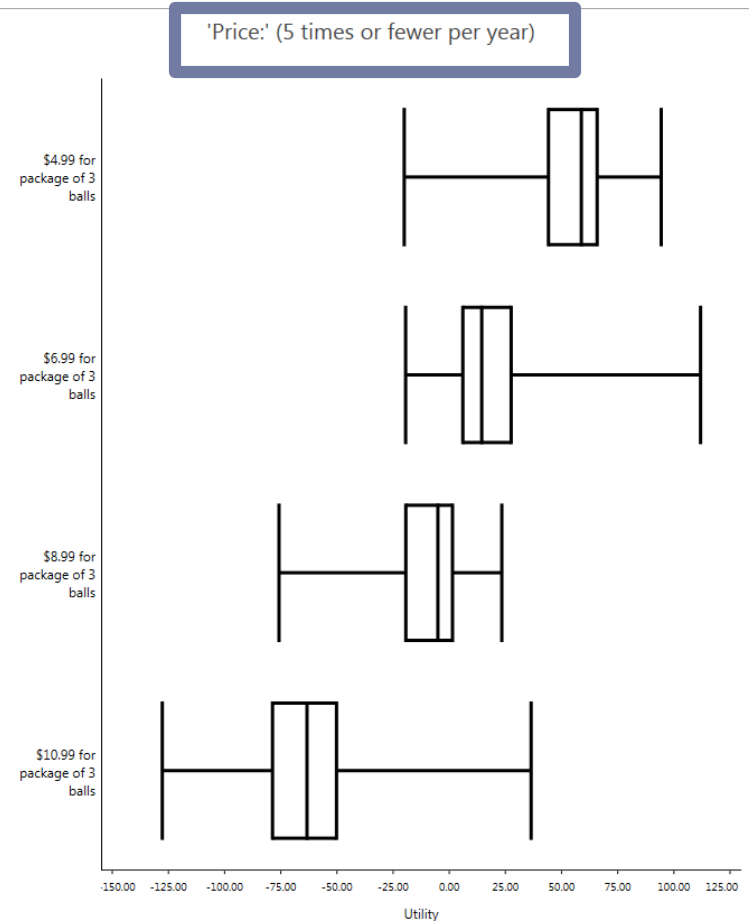|	A	B	C	D	
1	Respondent	Brand:	Performance:	Price:	
2	5567	34.43505	17.19110	48.37384	
3	11256	36.20776	19.72213	44.07011	
4	23367	24.80663	25.97027	49.22310	
5	23564	40.11239	22.07293	37.81467	
6	25789	42.31632	18.46612	39.21756	
7	30528	41.76997	30.43953	27.79050	
8	33721	45.55148	3.21971	51.22880	
9	37637	30.08919	16.88332	53.02750	
10	38497	31.98142	25.87104	42.14755	
11	44049	31.93421	25.79057	42.27523	
12	45291	49.18955	23.95452	26.85593	
13	48913	46.67177	14.81106	38.51718	
14	64220	22.84824	29.33912	47.81264	
15	65474	68.74008	9.41605	21.84387	
16	71137	31.71566	26.63424	41.65010	
17	78501	40.16893	24.46276	35.36831	
18	82146	35.64526	14.03149	50.32325	
19	82607	32.68975	26.32097	40.98929	
20	87036	37.47996	27.20533	35.31471	
21	87066	26.80919	27.83667	45.35414	
22	90281	20.77256	31.95991	47.26752	
23	97076	31.72125	29.43524	38.84351	
24	97352	44.89915	20.95528	34.14557	

Can we characterize the differences between respondents?

If a new customer comes in, what can we say about his own preferences?

Profiling

‣ We can segment after HB, e.g. based on <u>play frequency</u>

	A	B	C	D	
1	**Average Utility Values (by Segment)**				
2	Utility Scaling Method	Zero-Centered Differences			
3					
4		**Total**	**5 times or fewer per year**	**more than 50 times per year**	**Other**
5	Respondent Count	250	66	56	
6					
7			**Utility**		
8	**Label**	**Total**	**5 times or fewer per year**	**more than 50 times per year**	**Other**
9	High-Flyer Pro, by Smith and Forester	41.02	38.72	50.07	
10	Magnum Force, by Durango	32.01	26.58	34.23	
11	Eclipse+, by Golfers, Inc.	-31.65	-33.86	-37.92	
12	Long Shot, by Performance Plus	-41.38	-31.44	-46.39	
13					
14	Drives 5 yards farther than the average ball	-38.20	-36.34	-43.96	
15	Drives 10 yards farther than the average ball	11.80	4.18	12.28	
16	Drives 15 yards farther than the average ball	26.40	32.15	31.69	
17					
18	$4.99 for package of 3 balls	51.94	54.48	56.94	
19	$6.99 for package of 3 balls	15.22	17.87	8.20	
20	$8.99 for package of 3 balls	-6.12	-11.52	-1.03	
21	$10.99 for package of 3 balls	-61.04	-60.83	-64.11	
22					
23	None	-118.65	-190.28	114.26	
24					
25	**Average Importances (by Segment)**				
26					
27			**Importance**		
28	**Attribute**	**Total**	**5 times or fewer per year**	**more than 50 times per year**	**Other**
29	Brand:	35.10	33.05	33.42	
30	Performance:	24.65	25.07	26.04	
31	Price:	40.25	41.88	40.54	

'Price:' (5 times or fewer per year)

Profiling

- We can segment after HB, e.g. based on <u>play frequency</u>

'Price:' (more than 50 times per year)

Conclusions

- Count analysis and Logit are useful analyses as a first step, but cannot be used to design a marketing strategy.

  - They can also give us a hint on the potential interaction effects in the data

- LCA and HB are appropriate to design effective marketing strategies

  - LCA is managerially more intuitive and allows the firm to focus on a segment of the customers

  - HB is closer to the truth but can be harder to understand
    - ➜ always use in combination with profiling variables