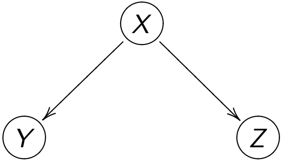
Likelihood-based inference for probabilistic graphical models: Some preliminary results

Marco Cattaneo
Department of Statistics, LMU Munich
cattaneo@stat.uni-muenchen.de

PGM 2010, Helsinki, Finland
14 September 2010

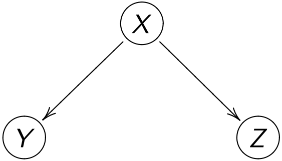
example

$X, Y, Z \in \{0, 1\}$



example

$X, Y, Z \in \{0, 1\}$

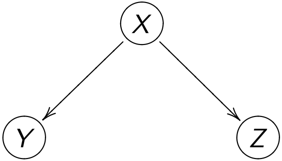


data:

X	Y	Z	#
0	0	0	15
0	0	1	25
0	1	0	7
0	1	1	5
1	0	0	6
1	0	1	35
1	1	0	3
1	1	1	4
			100

example

$X, Y, Z \in \{0, 1\}$



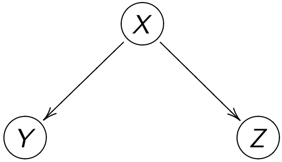
data:

X	Y	Z	#
0	0	0	15
0	0	1	25
0	1	0	7
0	1	1	5
1	0	0	6
1	0	1	35
1	1	0	3
1	1	1	4
			100

inference about $P(X = 1 \mid Y = 1,\ Z = 1)$:

example

$X, Y, Z \in \{0, 1\}$



data:

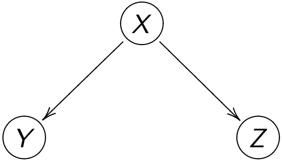
X	Y	Z	#
0	0	0	15
0	0	1	25
0	1	0	7
0	1	1	5
1	0	0	6
1	0	1	35
1	1	0	3
1	1	1	4
			100

inference about $P(X = 1 \mid Y = 1,\ Z = 1)$:

- ML estimate: 0.45

example

$X, Y, Z \in \{0, 1\}$



data:

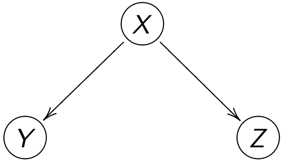
X	Y	Z	#
0	0	0	15
0	0	1	25
0	1	0	7
0	1	1	5
1	0	0	6
1	0	1	35
1	1	0	3
1	1	1	4
			100

inference about $P(X = 1 \mid Y = 1,\ Z = 1)$:

- ML estimate: 0.45
- Bayesian estimate
  with uniform priors: 0.46

example

$X, Y, Z \in \{0, 1\}$

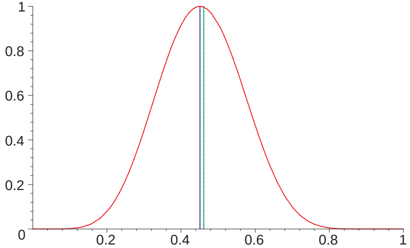


data:

X	Y	Z	#
0	0	0	15
0	0	1	25
0	1	0	7
0	1	1	5
1	0	0	6
1	0	1	35
1	1	0	3
1	1	1	4
			100

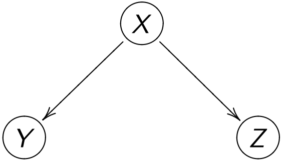
inference about $P(X = 1 \mid Y = 1,\ Z = 1)$**:**

- ML estimate: 0.45
- Bayesian estimate with uniform priors: 0.46
- profile likelihood function:



example $\times$ 100

$X, Y, Z \in \{0, 1\}$



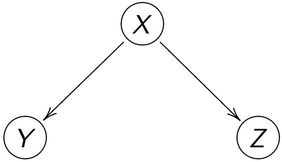
data:

X	Y	Z	#
0	0	0	1500
0	0	1	2500
0	1	0	700
0	1	1	500
1	0	0	600
1	0	1	3500
1	1	0	300
1	1	1	400
			10000

inference about $P(X = 1 \mid Y = 1,\ Z = 1)$:

example $\times$ 100

$X, Y, Z \in \{0, 1\}$



data:

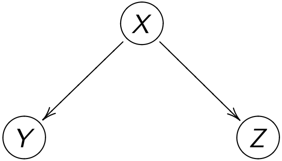
X	Y	Z	#
0	0	0	1500
0	0	1	2500
0	1	0	700
0	1	1	500
1	0	0	600
1	0	1	3500
1	1	0	300
1	1	1	400
			10000

inference about $P(X = 1 \mid Y = 1, Z = 1)$:

- ML estimate: 0.45

example $\times$ 100

$X, Y, Z \in \{0, 1\}$



data:

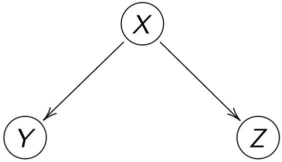
X	Y	Z	#
0	0	0	1500
0	0	1	2500
0	1	0	700
0	1	1	500
1	0	0	600
1	0	1	3500
1	1	0	300
1	1	1	400
			10000

inference about $P(X = 1 \mid Y = 1,\ Z = 1)$:

- ML estimate: 0.45
- Bayesian estimate
  with uniform priors: 0.46−0.01

example × 100

$X, Y, Z \in \{0, 1\}$

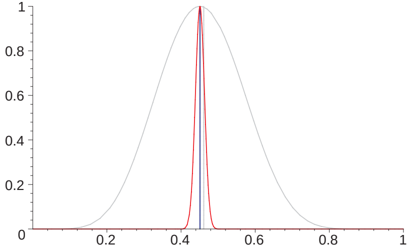


data:

X	Y	Z	#
0	0	0	1500
0	0	1	2500
0	1	0	700
0	1	1	500
1	0	0	600
1	0	1	3500
1	1	0	300
1	1	1	400
			10000

inference about $P(X = 1 \mid Y = 1, Z = 1)$:

- ML estimate: 0.45
- Bayesian estimate with uniform priors: 0.46—0.01
- profile likelihood function:



likelihood

- **profile likelihood function for $P(X = 1 \mid Y = 1,\, Z = 1)$:**

$$\boxed{lik(p) \propto \max_{\substack{bn \in \mathcal{BN}\,: \\ P_{bn}(X=1 \mid Y=1,\, Z=1)=p}} P_{bn}(\text{data})} \qquad \text{for all } p \in [0,1],$$

where $\mathcal{BN}$ is the set of all Bayesian networks *bn* compatible with the given graph

likelihood

- **profile likelihood function for $P(X = 1 \mid Y = 1,\ Z = 1)$:**

$$\boxed{lik(p) \propto \max_{\substack{bn \in \mathcal{BN}\,: \\ P_{bn}(X=1 \mid Y=1,\, Z=1)=p}} P_{bn}(\text{data})} \qquad \text{for all } p \in [0,1],$$

  where $\mathcal{BN}$ is the set of all Bayesian networks *bn* compatible with the given graph

- for each $\beta \in [0,1]$,

$$\{p \in [0,1] : lik(p) \geq \beta\}$$

  is a **confidence interval** for $P(X = 1 \mid Y = 1,\ Z = 1)$ with (approximate) level $F_{\chi^2_1}(-2 \log \beta)$

likelihood

- **profile likelihood function for $P(X = 1 \mid Y = 1,\ Z = 1)$:**

$$\boxed{lik(p) \propto \max_{\substack{bn \in \mathcal{BN}\,: \\ P_{bn}(X=1 \mid Y=1,\, Z=1)=p}} P_{bn}(\text{data})} \qquad \text{for all } p \in [0,1],$$

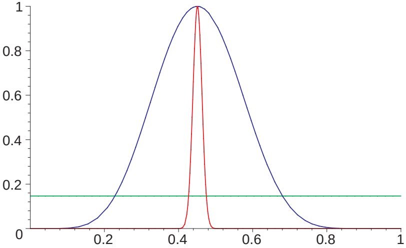
where $\mathcal{BN}$ is the set of all Bayesian networks *bn* compatible with the given graph

- for each $\beta \in [0,1]$,

$$\{p \in [0,1] : lik(p) \geq \beta\}$$

is a **confidence interval** for $P(X = 1 \mid Y = 1,\ Z = 1)$ with (approximate) level $F_{\chi^2_1}(-2 \log \beta)$

- example: [0.23, 0.68] and [0.43, 0.48] are (approximate) 95% confidence intervals for $P(X = 1 \mid Y = 1,\ Z = 1)$ in the cases with 100 and 10000 data, respectively



conclusion

- **main result of the paper:** a simple method for calculating profile likelihood functions in an important class of problems

conclusion

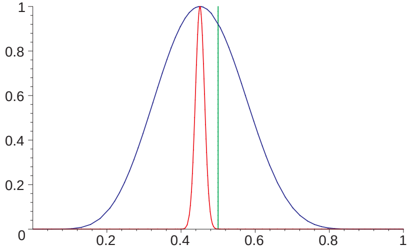
- **main result of the paper:** a simple method for calculating profile likelihood functions in an important class of problems
- **example of application:** a naive classifier assigning a degree of uncertainty to each preference between classes

conclusion

- **main result of the paper:** a simple method for calculating profile likelihood functions in an important class of problems

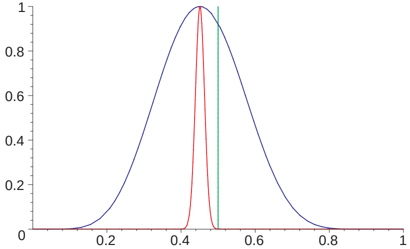
- **example of application:** a naive classifier assigning a degree of uncertainty to each preference between classes

- example: 0.92 and 0.00 are the degrees of uncertainty $lik(0.5)$ of the classification $X = 1$ (given $Y = 1$ and $Z = 1$) in the cases with 100 and 10000 data, respectively



conclusion

- **main result of the paper:** a simple method for calculating profile likelihood functions in an important class of problems
- **example of application:** a naive classifier assigning a degree of uncertainty to each preference between classes
- example: 0.92 and 0.00 are the degrees of uncertainty $lik(0.5)$ of the classification $X = 1$ (given $Y = 1$ and $Z = 1$) in the cases with 100 and 10000 data, respectively



- **future research:** generalization to larger classes of problems and consideration of the uncertainty about the graph