A Generalization of Credal Networks

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

June 24, 2009

foundations of statistics

frequentist approach	**likelihood approach**	**Bayesian approach**
empirical	empirical	personalistic
repeated-sampling	conditional	conditional

foundations of statistics

frequentist approach	**likelihood approach**	**Bayesian approach**
empirical	empirical	personalistic
repeated-sampling	conditional	conditional

can be interpreted as an **imprecise probability** approach:

(profile) likelihood function $=:$ membership function of fuzzy probability

foundations of statistics

frequentist approach

empirical
repeated-sampling

likelihood approach

empirical
conditional

Bayesian approach

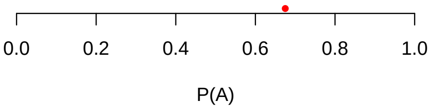
personalistic
conditional

can be interpreted as an **imprecise probability** approach:

(profile) likelihood function =: membership function of fuzzy probability

generalizations:

precise probability



foundations of statistics

frequentist approach	**likelihood approach**	**Bayesian approach**
empirical	empirical	personalistic
repeated-sampling	conditional	conditional

can be interpreted as an **imprecise probability** approach:

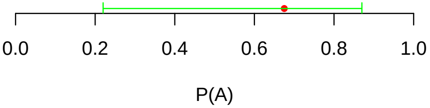
(profile) likelihood function =: membership function of fuzzy probability

generalizations:

precise probability

$\downarrow$

interval probability



foundations of statistics

frequentist approach	**likelihood approach**	**Bayesian approach**
empirical	empirical	personalistic
repeated-sampling	conditional	conditional

can be interpreted as an **imprecise probability** approach:

(profile) likelihood function =: membership function of fuzzy probability

generalizations:

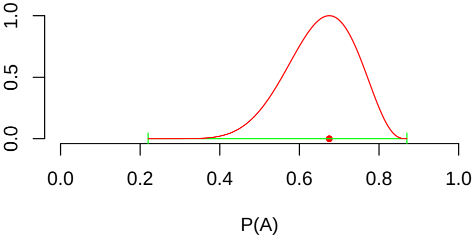
precise probability

$\downarrow$

interval probability

$\downarrow$

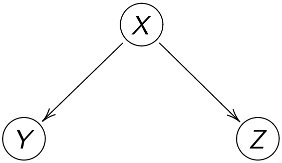
fuzzy probability



probabilistic graphical models

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

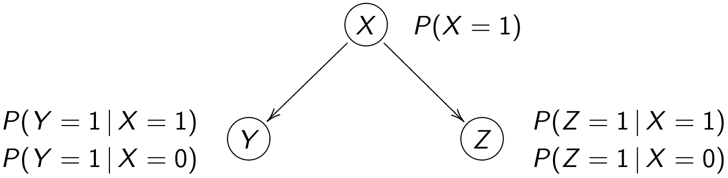
Y and Z independent conditional on X:



probabilistic graphical models

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

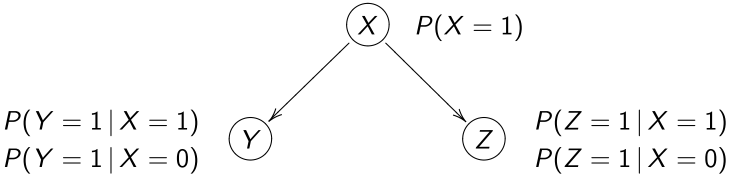
Y and Z independent conditional on X:



probabilistic graphical models

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

Y and Z independent conditional on X:



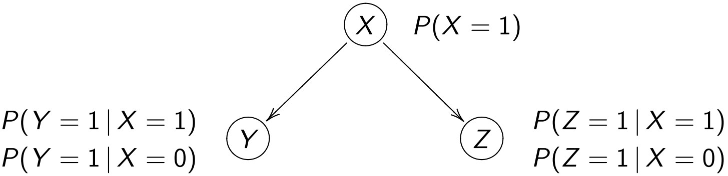
generalizations:

precise probabilities: Bayesian networks

probabilistic graphical models

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

Y and Z independent conditional on X:



generalizations:

precise probabilities: Bayesian networks

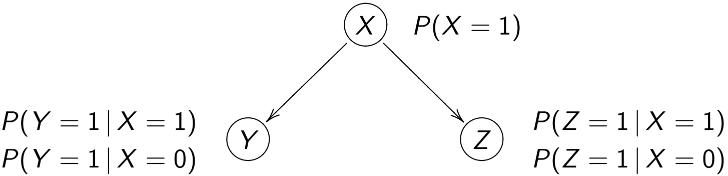
$\downarrow$

interval probabilities: credal networks

probabilistic graphical models

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

Y and Z independent conditional on X:



generalizations:

precise probabilities: Bayesian networks

$\downarrow$

interval probabilities: credal networks

$\downarrow$

fuzzy probabilities: hierarchical networks

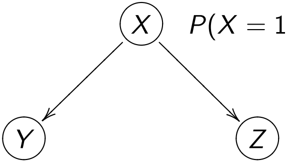
training data

X	Y	Z	#
0	0	0	21
0	0	1	6
0	1	0	30
0	1	1	7
1	0	0	9
1	0	1	15
1	1	0	5
1	1	1	7
			100

training data

X	Y	Z	#
0	0	0	21
0	0	1	6
0	1	0	30
0	1	1	7
1	0	0	9
1	0	1	15
1	1	0	5
1	1	1	7
			100

simulated according to:

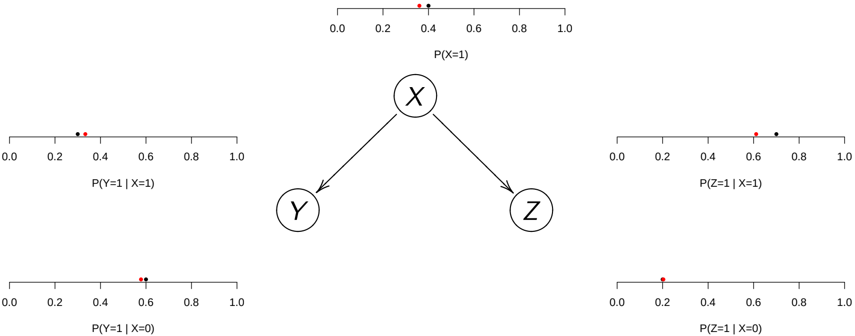


$P(X = 1) = 0.4$

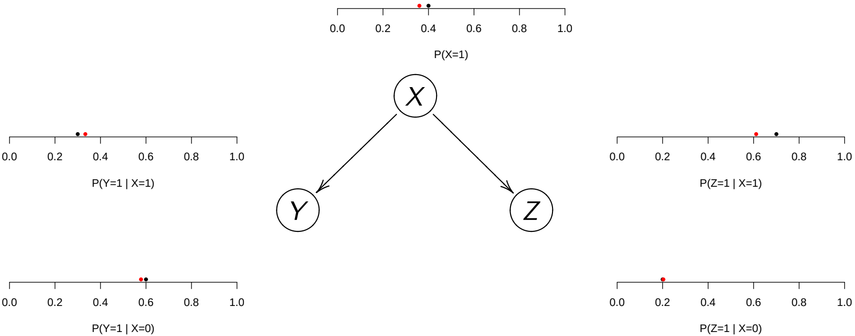
$P(Y = 1 \mid X = 1) = 0.3$
$P(Y = 1 \mid X = 0) = 0.6$

$P(Z = 1 \mid X = 1) = 0.7$
$P(Z = 1 \mid X = 0) = 0.2$

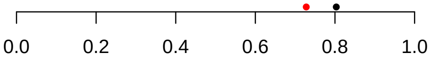
Bayesian network via MLE



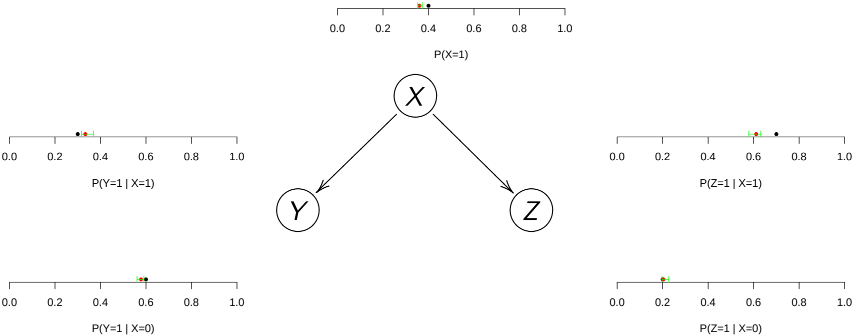
Bayesian network via MLE



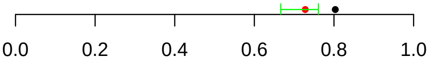
$\Rightarrow \; P(X = 1 \mid Y = 0, \, Z = 1)$:



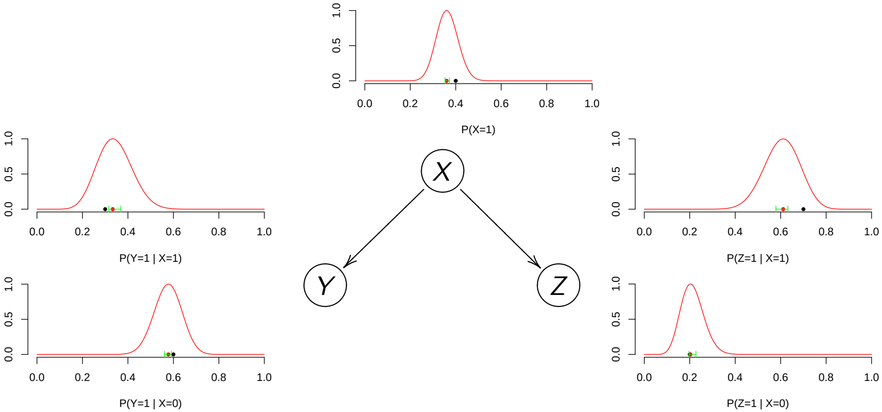
credal network via IDM (with $s = 2$)



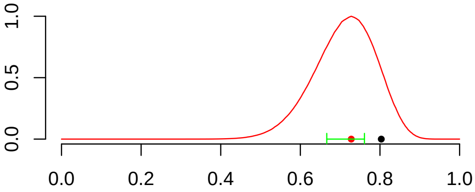
$\Rightarrow \; P(X = 1 \mid Y = 0, \; Z = 1):$



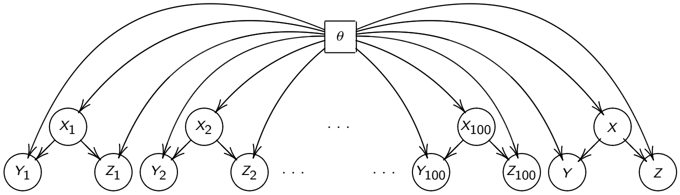
hierarchical network



$\Rightarrow \; P(X = 1 \mid Y = 0,\; Z = 1) :$

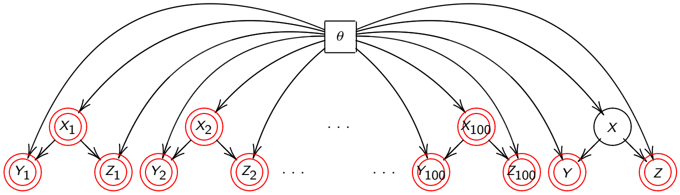


global models



prior ignorance about $\theta \in [0, 1]^5$

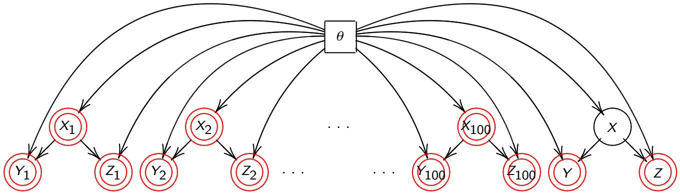
global models



prior ignorance about $\theta \in [0, 1]^5$

302 variables observed

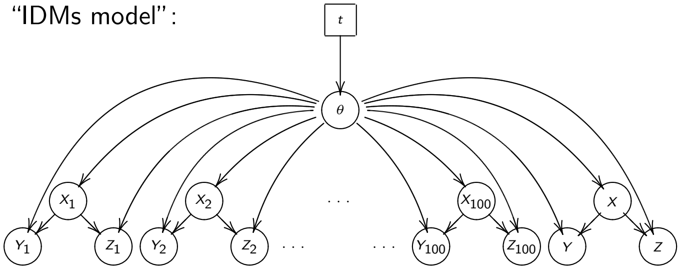
global models



prior ignorance about $\theta \in [0, 1]^5$

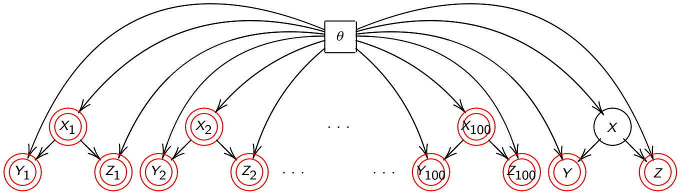
302 variables observed

"IDMs model":



prior ignorance about $t \in [0, 1]^5$

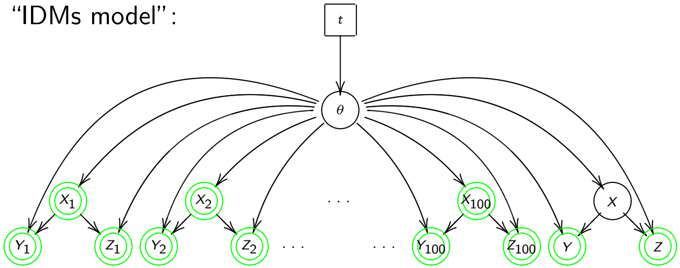
global models



prior ignorance about $\theta \in [0,1]^5$

302 variables observed

"IDMs model":



prior ignorance about $t \in [0,1]^5$

302 variables observed

conclusions and results

advantages of (likelihood-based) **fuzzy probability** over interval probability:

- **more expressive** (relative plausibility of different values in the probability interval)
- **more powerful updating rule** (extracts more information from the data)
- **more robust updating rule** (less sensitive to small perturbations of the model)

conclusions and results

advantages of (likelihood-based) **fuzzy probability** over interval probability:

- **more expressive** (relative plausibility of different values in the probability interval)
- **more powerful updating rule** (extracts more information from the data)
- **more robust updating rule** (less sensitive to small perturbations of the model)

mathematical results of the paper:

- **d-separation** implies conditional irrelevance in hierarchical networks
- hierarchical networks can be described by **convex sets of measures**, and it suffices to consider the extreme points