On causal questions and statistical answers

Saskia le Cessie

Work by STRATOS TG 7, Causal Inference

Bianca De Stavola	University College London, England
Els Goetghebeur	Ghent University, Belgium
Saskia Le Cessie	Leiden University Medical Centre, the Netherlands
Erica Moodie	McGill University, Canada
Ingeborg Waernbaum	Upsala University, Sweden

The Stratos Initiative (http://www.stratos-initiative.org/)

- STRengthening Analytical Thinking for Observational Studies
- Tremendous amount of development of new statistical methods
- In practice only few of these methods are used
- Big gap between methods development and methods applications
- Aim: provide guidence for statisticians and other data analysists

STRATOS: 9 topic groups

1. Missing data
2. Selection of variables and functional forms in multivariable analysis
3. Initial data analysis
4. Measurement error and misclassification
5. Study design
6. Evaluating diagnostic tests and prediction models
7. Causal inference
8. Survival analysis
9. High-dimensional data

And several panels (data panel, glossary panel, publication panel)

Aims of TG 7: causal inference

Offer **guidance** in the changing and challenging landscape of causal inference from observational studies by presenting intuitive formal and practically principled building blocks, tools for analysis through ...

Aims of TG 7: causal inference

Offer **guidance** in the changing and challenging landscape of causal inference from observational studies by presenting intuitive formal and practically principled building blocks, tools for analysis through ...

- courses (ISCB, IBC (next year), SSC, Master PH in Comparative Effectiveness Research Paris, U. Decartes)
- didactical material (case studies)
- a simulation learner
- website. (ofcaus.org)
- talks at meetings (+ organizing meetings)
- papers

emphasis on insight and tools

Tutorial: Causal questions and principled answers

- Statistical literature is exploding
- Several formalisms and schools of thought
- Expanding tool kit
- Seemingly 'easy' to do $->$ software
- Transparency on fundamentals and important 'detail' lacking

Our tutorial provide guidance and insights

- Focus on point exposures

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.
4. Formalise potential outcomes

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.
4. Formalise potential outcomes
5. Specify the target causal effect

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.
4. Formalise potential outcomes
5. Specify the target causal effect
6. State assumptions validating the causal effect estimation from the available data.

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.
4. Formalise potential outcomes
5. Specify the target causal effect
6. State assumptions validating the causal effect estimation from the available data.
7. Estimate target causal effect.

Key steps of causal Inference

1. Define treatment with relevant levels/values corresponding to scientific question of study.
2. Define outcome.
3. Define the population(s) of interest.
4. Formalise potential outcomes
5. Specify the target causal effect
6. State assumptions validating the causal effect estimation from the available data.
7. Estimate target causal effect.
8. Evaluate the validity of the assumptions and perform sensitivity analyses as needed.

Promotion of Breastfeeding Intervention Trial - PROBIT

(Kramer et al, 2001)

- Pregnant women from low income area of Belarus, 1996-1997
- (cluster) randomised to
  - invitation for breast feeding encouraging educational program
  - no invitation

  during last term of pregnancy.
- Primary outcome: weight of baby at age 3 months.
- in total 17,044 women included
  (8,667 in the active arm and 8,377 in the control arm).

Many different causal questions of interest

Policy maker/health insurer may ask about impact on weight at 3m of:

- offering a BF encouragement program
    - The intention to treat effect

Many different causal questions of interest

Policy maker/health insurer may ask about impact on weight at 3m of:

- offering a BF encouragement program
  - The intention to treat effect
- following the BF program
  - The per protocol effect

Many different causal questions of interest

Policy maker/health insurer may ask about impact on weight at 3m of:

- offering a BF encouragement program
  - The intention to treat effect
- following the BF program
  - The per protocol effect
- starting BF

Many different causal questions of interest

Policy maker/health insurer may ask about impact on weight at 3m of:

- offering a BF encouragement program
    - The intention to treat effect
- following the BF program
    - The per protocol effect
- starting BF
- continuing BF for a full 3 months

in different (targeted) population strata

Similar discussion by European Medicines Agency (EMA)

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

30 August 2017
EMA/CHMP/ICH/436221/2017
Committee for Human Medicinal Products

ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials

Step 2b

EMA estimands

Accounting for intercurrent events (e.g, change treatment, death)

- Treatment policy strategy
  - Intention to treat

EMA estimands

Accounting for intercurrent events (e.g, change treatment, death)

- Treatment policy strategy
  - Intention to treat
- Composite strategy
  - Combine outcome of interest and intercurrent event

EMA estimands

Accounting for intercurrent events (e.g, change treatment, death)

- Treatment policy strategy
  - Intention to treat
- Composite strategy
  - Combine outcome of interest and intercurrent event
- Hypothetical strategy
  - What would happen if patients would not switch treatment/ would not die

EMA estimands

Accounting for intercurrent events (e.g, change treatment, death)

- Treatment policy strategy
    - Intention to treat
- Composite strategy
    - Combine outcome of interest and intercurrent event
- Hypothetical strategy
    - What would happen if patients would not switch treatment/ would not die
- Principal stratum strategy
    - Effect in those who would not switch treatment

More causal questions of interest

Pregnant woman may ask about impact on weight at 3m of:

- following a BF program that is offered
- starting BF
- continuing BF for the full 3 months

in 'her' population stratum (e.g.highly educated, 30 year old, ...)
A very conditional treatment effect

Many of these causal estimands are not straight forward to estimate

Problems of

- Confounding
- Selection bias

Even harder when observational data ("real life data") are used

Our tutorial uses a simulated dataset

The Simulation Learner

- We simulated data based on the Probit trial
- "Real life simulation"
- Illustrates concepts and methods on data

Advantage

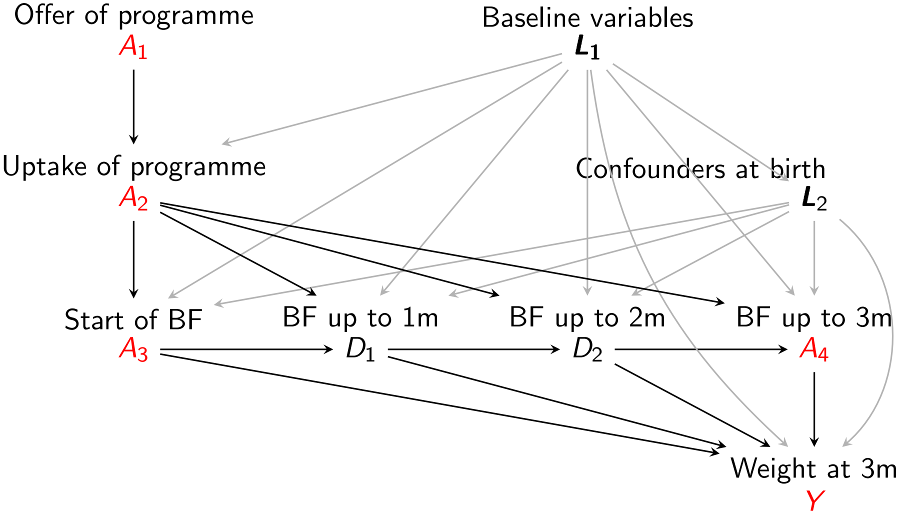
- Truth is known in simulated data
- We can extend data with counterfactual outcomes

We generated data inspired by PROBIT trial

- A simulated version of individually *randomised* women.
- 50 % received offer for the breast feeding educational programme
- 50 % received no offer
- Outcome: weight of baby after 3 months

Observed data is enriched by generation of potential outcome data

The data generation



Simulate potential intermediate events and outcomes

For each women:

- We simulated potential intermediate events
  - Will she take up the programme if she is invited?
  - Will she start breastfeeding if she is invited for the programmer?
  - Will she start breastfeeding without an offer
  - Will she start breastfeeding if she followed the program?

Simulate potential intermediate events and outcomes

For each women:

- We simulated potential intermediate events
  - Will she take up the programme if she is invited?
  - Will she start breastfeeding if she is invited for the programmer?
  - Will she start breastfeeding without an offer
  - Will she start breastfeeding if she followed the program?
- And potential outcomes under different "interventions"
  The weight of the baby after 3 months under
  - no offer
  - offer of the programme
  - following the programme
  - starting breastfeeding after following the programme
  - starting breastfeeding without a programme, etc

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

Potential intermediate events

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

Potential intermediate events

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

Potential intermediate events

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

Potential intermediate events

- Would she take up the programme if she received an offer? YES

An example

Woman nr 7:

- 22 years old, higher educated, rural area, no smoker, baby boy with birth weight 2667 g
- $A_1 = 0$ She was randomized for control
- $A_2 = 0$ She did not follow the programme
- $A_3 = 0$ She did not start breastfeeding

Potential intermediate events

- Would she take up the programme if she received an offer? YES
- Would she start breastfeeding if she did receive an offer? YES, AND CONTINUE FOR 3 MONTHS

Potential outcomes for this woman

Weight of baby after 3 months

- Y_{obs} was 5813 grams

Potential outcomes

- $Y_{\mathfrak{a}_1(0)} = 5813g$, the potential outcome under no intervention.
- $Y_{\mathfrak{a}_1(1)} = 6133g$ the potential outcome under intervention.
- $Y_{\mathfrak{a}_2(1)} = 6133g$ the potential outcome under actually following the programme.
- $Y_{\mathfrak{a}_3(1),\mathfrak{a}_1(0)} = 6133g$ the potential outcome when she would start breastfeeding but not received an offer
- etc.

Mean potential weight at 3 months in all women under different treatments

outcome	interventions	population
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182

Mean potential weight at 3 months in all women under different treatments

outcome	interventions	population
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182

Causal questions:

(i) What is the overall mean change in Y due to inviting expectant women to attend the BF program?
$\rightarrow$ ITT$=E(Y(\mathfrak{a}_1(1) - Y(\mathfrak{a}_1(0))$=98g

Mean potential weight at 3 months in all women under different treatments

outcome	interventions	population
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182

Causal questions:

(i) What is the overall mean change in Y due to inviting expectant women to attend the BF program?
$\rightarrow$ ITT$=E(Y(\mathfrak{a}_1(1) - Y(\mathfrak{a}_1(0))$=98g

(ii) What is the overall mean change in Y if all women would attend training?
$\rightarrow E(Y(\mathfrak{a}_2(1)) - Y(\mathfrak{a}_2(0)) = 165g$

Different (sub)populations could be of interest

Women:

(a) with babies for whom BF is not counter-indicated (overall population)

(b) who attended the training (the “treated”)

(c) who would BF if invited to the training but not otherwise (the “BF compliers”)

(d) from rural areas

Causal effects of different interventions in these subpopulations could be of interest

Mean potential weight at 3 months in treated and not treated

outcome	interventions	population	$A_2 = 1$	$A_1 = 1$ $A_2 = 0$
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017	6047	5964
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115	6200	5964
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182	6200	6149

Mean potential weight at 3 months in treated and not treated

outcome	interventions	population	$A_2 = 1$	$A_1 = 1$ $A_2 = 0$
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017	6047	5964
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115	6200	5964
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182	6200	6149

Causal questions:

(i) Among the women who followed the program, what estimated difference did it make (on average)?
ATT=Average treatment effect in the treated 6200-6047=153 g.

Mean potential weight at 3 months in treated and not treated

outcome	interventions	population	$A_2 = 1$	$A_1 = 1$ $A_2 = 0$
$Y_{\mathfrak{a}_1(0)}$	programme not offered	6017	6047	5964
$Y_{\mathfrak{a}_1(1)}$	programme offered	6115	6200	5964
$Y_{\mathfrak{a}_2(1)}$	programme followed	6182	6200	6149

Causal questions:

(i) Among the women who followed the program, what estimated difference did it make (on average)?
ATT=Average treatment effect in the treated 6200-6047=153 g.

(ii) And expected potential impact for women who did not follow the program? ATNT = 185 g

Methods to estimate these causal effects from observed data

Different approaches (with different assumptions, which may estimate different causal effects)

- Outcome regression
  - Fit a regression model, with interactions between treatment and covariates L
  - Marginalizing over the distribution of L

Methods to estimate these causal effects from observed data

Different approaches (with different assumptions, which may estimate different causal effects)

- Outcome regression
    - Fit a regression model, with interactions between treatment and covariates L
    - Marginalizing over the distribution of L
- Propensity score methods
    - Estimate probability to receive treatment(propensity)
    - Then perform propensity matching, weighting, double robust estimation, or regression

Methods to estimate these causal effects from observed data

Different approaches (with different assumptions, which may estimate different causal effects)

- Outcome regression
  - Fit a regression model, with interactions between treatment and covariates L
  - Marginalizing over the distribution of L
- Propensity score methods
  - Estimate probability to receive treatment(propensity)
  - Then perform propensity matching, weighting, double robust estimation, or regression
- Instrumental variable analysis
  - Randomisation as instrument

Methods to estimate these causal effects from observed data

Different approaches (with different assumptions, which may estimate different causal effects)

- Outcome regression
    - Fit a regression model, with interactions between treatment and covariates L
    - Marginalizing over the distribution of L
- Propensity score methods
    - Estimate probability to receive treatment(propensity)
    - Then perform propensity matching, weighting, double robust estimation, or regression
- Instrumental variable analysis
    - Randomisation as instrument
- In R, Stata, and SAS

Different methods with different assumptions

Method	No unmeasured Confounding	Correct specification Y model	PS model	Core IV assumption
Outcome regression				
conditional on L	$\surd$	$\surd$		
conditional on PS	$\surd$	$\surd^{(*)}$	$\surd^{(*)}$	
Stratification by PS	$\surd$		$\surd$	
Matching by PS	$\surd$		$\surd$	
IPW by PS	$\surd$		$\surd$	
Double robust with PS	$\surd$	either	or	
Instr. variable				$\surd^{(**)}$

$^{(*)}$ Either of these if the outcome model is linear.

$^{(**)}$ Plus no Z-L interaction, or monotonicity assumption

Gain in weight@3months, following the programme ($A_2 = 1$), versus no programme

Estimand	Estimation method	Estimate	(SE)
ATE			
	True value	165.1 g	
	Crude regression	196.0	(9.6)
	Regression adjustment (simple)	155.4	(9.5)
	Regression adjustment (with interactions)	165.0	(9.7)
	PS stratification (6 strata)	165.0	(9.4)
	Regression with PS	156.2	(9.0)
	PS matching (1 match)	155.7	(10.1)
	PS matching (3 matches)	154.9	(10.1)
	PS IPW	164.7	(9.7)
	PS Double robust IPW	164.7	(9.7)
	Instrumental variable	146.2	(14.0)

Gain in weight@3months, following the programme ($A_2 = 1$), versus no programme

Estimand	Estimation method	Estimate	(SE)
ATE			
	True value	165.1 g	
	Crude regression	196.0	(9.6)
	Regression adjustment (simple)	155.4	(9.5)
	Regression adjustment (with interactions)	165.0	(9.7)
	PS stratification (6 strata)	165.0	(9.4)
	Regression with PS	156.2	(9.0)
	PS matching (1 match)	155.7	(10.1)
	PS matching (3 matches)	154.9	(10.1)
	PS IPW	164.7	(9.7)
	PS Double robust IPW	164.7	(9.7)
	Instrumental variable	146.2	(14.0)

Results for A_3 (Starting breastfeeding)

	$A_1 = 0$		$A_1 = 1$	
Estimation method	Estimate	(SE)	Estimate	(SE)
ATE				
True value	386.8		422.3	
Crude regression	503.2	(11.6)	582.0	(12.2)
Regression (simple)	384.3	(2.8)	428.0	(3.3)
Regression (with interactions)	384.7	(3.2)	425.3	(2.7)
Regression with PS	384.4	(3.2)	425.9	(3.3)
PS stratification (6 strata)	392.2	(4.1)	442.0	(6.5)
PS matching (1 match)	386.5	(8.1)	429.0	(10.6)
PS matching (3 matches)	380.7	(5.5)	437.2	(7.8)
PS IPW	384.7	(3.8)	426.6	(6.7)
PS DR IPW	384.8	(3.9)	426.7	(7.0)
IV	513.3	(44.4)	–	–

Results for A_3 (Starting breastfeeding)

	$A_1 = 0$		$A_1 = 1$	
Estimation method	Estimate	(SE)	Estimate	(SE)
ATE				
True value	386.8		422.3	

Results for A_3 (Starting breastfeeding)

	$A_1 = 0$		$A_1 = 1$	
Estimation method	Estimate	(SE)	Estimate	(SE)
ATE				
True value	386.8		422.3	
Crude regression	503.2	(11.6)	582.0	(12.2)
Regression (simple)	384.3	(2.8)	428.0	(3.3)
Regression (with interactions)	384.7	(3.2)	425.3	(2.7)
Regression with PS	384.4	(3.2)	425.9	(3.3)
PS stratification (6 strata)	392.2	(4.1)	442.0	(6.5)
PS matching (1 match)	386.5	(8.1)	429.0	(10.6)
PS matching (3 matches)	380.7	(5.5)	437.2	(7.8)
PS IPW	384.7	(3.8)	426.6	(6.7)
PS DR IPW	384.8	(3.9)	426.7	(7.0)
IV	513.3	(44.4)	–	–

Conclusion

- Think about the causal question of interest before starting a causal analysis
- Important to clearly state the assumptions made when estimating
- Simulation learner is useful because:
  - Generates dataset with observed data, augmented with potential outcomes
  - Gives more insight in process of data generation
  - Actual causal effects are known
  - Great help in finding correct ways of analysis (which turned out to be different for A_2 and A_3)
  - Enables to compare different analytic methods.
  - It is helpful in teaching causal methods
  - Code of generation and analysis of data is available

www.ofcaus.org

STRATOS INITIATIVE Topic Group 7

Home About Us Causal inference literature Courses Estimation approaches and data e... R implementatio

Causal Inference

STRATOS Initiative

Topic group 7 is a member of the STRATOS Initiative (STRengthening Analytical Thinking for Observational Studies) which is a large collaboration of experts in many different areas of biostatistical research. Ongoing research, discussions and activities within STRATOS are conducted in nine topic groups and several cross-cutting panels.