Reconstructing the selection function

Jan Rybizki - MPIA Heidelberg, Germany ISSS 2019 L'Aquila 06/04/19

Outline

- Historic examples of sample bias
- Gaia detection limits
- Your own selection
- Gaia's intrinsic completeness
- RVS example using mock data

Malmquist bias



Distance



Luminosity

Volume-complete sample

- Hipparcos 100 pc sample
- 10 mas, 2 mas uncertainty



Gaia detection limits



Gaia detection limits

Measurement with uncertainty



Gaia detection limits

Measurement with uncertainty



- Brightness limitation
- Data transfer limitation
- Scanning law
- Contrast sensitivity

Magnitude distributions





Slide from Coryn Bailer-Jones

Gaia DR2 down to 20.7 G

Gaia Magnitude limit

GDR2all: Gmaglimit



16.8 mag_limit (hp level 7) 21.4

Contrast sensitivity



Your own selection



Your own selection



- parameter availability
- quality cuts
- specific stellar tracer

Apsis completeness

- Stellar parameters up to 17th Gmag
- Teff, A_G

Teff completeness

GDR2 APSIS teff completeness



)	completeness	

A_G completeness

GDR2 APSIS A_G completeness



)	completeness	0.940299

Photometry cut

Photometry cut

GDR2





Photometry cut per Gmag bin



completeness







- Age distribution of red clump stars, create your own using: https://bit.ly/2W1y53S
- Example Chempy application: https://bit.ly/2MliEnl







Slide from Ronald Drimmel

- From independent measurements
 - Short time in between
- Only in sufficiently small bins of:
 - Magnitude
 - Colour
 - Position



- From independent measurements
 - Short time in between
- Only in sufficiently small bins of:
 - Magnitude
 - Colour
 - Position
- Problems:
 - Crossmatch
 - Colour-transformation
 - Spurious sources (ruwe)
 - Duplicated sources
- https://github.com/jan-rybizki/gdr2_completeness



Slide from Ronald Drimmel





Credit: FrancescoA - Near-stars-past-future-de.svg, CC BY-SA 3.0

Encounter diagramm (1995 stars) LMA



Encounter diagramm (1995 stars) LMA



RVS sample (7.2 mio stars)

• 3550 < Teff [K] < 6900





True Galaxy

Besancon model-







https://github.com/jan-rybizki/Gaia_data_analysis_HGSFP2018/blob/master/notebooks/Day%203.ipynb

Completeness

completeness is: 16 percent



Summary

 Your selection should not depend on unobserved data

Summary

- Your selection should not depend on unobserved data
- Try to incorporate your data selection into your forward model and compare in the space of observables

Summary

- Your selection should not depend on unobserved data
- Try to incorporate your data selection into your forward model and compare in the space of observables
- You can/should do your inference on mock data

Resources

- Spectroscopic surveys http://arxiv.org/abs/1902.10485
- Extinction effects on Selection function https://arxiv.org/abs/1509.06751
- Application for TGAS https://arxiv.org/abs/1704.05063
- GDR2 completeness and selection https://github.com/jan-rybizki/gdr2_completeness

Resources

- GDR2mock catalog (TAP queries, like DR2) https://arxiv.org/abs/1804.01427
- FIRE simulation and mock observables https://arxiv.org/abs/1806.10564
- AuriGaia simulation and mock observables https://arxiv.org/abs/1804.08549
- Equation 2 and 3 give GRVS from G and GRP https://arxiv.org/abs/1804.09365
- GAVO Gaiadr2.light contains GRVS and ruwe

Thank you for your attention

Ignorability condition

Selection effects are ignorable if:

- The posterior for the parameters of interest can be written strictly in terms of the observed data equivalent to:
- The prior distribution for ϕ is independent of the prior distribution for all other parameters.
- Selection does not depend on unobserved (or potentially unobserved) data