Population Synthesis

Rob Izzard

Universiteit Utrecht

SINS Summer School 2007

Contents

Population

From Late Latin populationem (c.470, nom. populatio) "a people, multitude," from Latin populus "people", of unknown origin, possibly from Etruscan.

Population

- Any finite or infinite aggregation of individuals, not necessarily animate, subject to a statistical study.
- The set of individuals, items, or data from which a statistical sample is taken. Also called universe.
- The entire aggregation of items from which samples can be drawn
- The set of individuals, items, or data from which a statistical sample is taken.
- A group of individual persons, objects, or items from which samples are taken for statistical measurement.

Synthesis

*From L. synthesis "collection, set, composition (of a medication)," from Gk. synthesis "composition," from syntithenai "put together, combine," from syn- "together" + tithenai "put, place," from PIE base *dhe- "to put, to do" (also root of factitious [artificial, contrived] from L. factitius "artificial", English "to do" and French "faire")*

Synthesis

- The combining of the constituent elements of separate material or abstract entities into a single or unified entity (opposed to analysis).
- A complex whole formed by combining.
- The combining of separate elements or substances to form a coherent whole.
- The combination of ideas into a complex whole.
- The combination of thesis and antithesis in the Hegelian dialectical process whereby a new and higher level of truth is produced

Population Synthesis

In the astronomical/stellar astrophysical context

- The process of combining *stellar models* to make a *stellar population* upon which a *statistical analysis* can be performed and which can, hopefully!, be compared to *real-life observations.*
- No Wikipedia page!

Concepts:

- Stellar model: a simulation describing the stars you would like to investigate
- Stellar population: a group of stars with something in common
- Observable: something you can see and measure, both in the models and in the real world

The General Idea

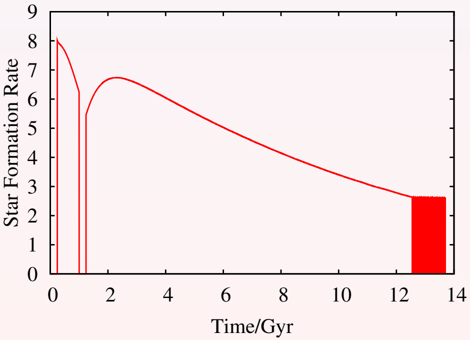
1. Make your stellar models for the stars you wish to test
2. From these extract the simulated value – probably a distribution of values – you would like to compare to observation(s)
3. Look at real stars to determine the “real-life” distribution (easy bit)
4. Compare the two and interpret the difference to see where you have gone wrong (requires brain!)
5. Refine your stellar model physics based on your interpretation, if you can
6. Return to step 1

Making a Synthetic Population

- Choose:
  - The stellar birth rate
  - An initial distribution
  - The initial values of many *freeish* parameters, some of which are relevant to your problem, some of which are not (how do you know?)
- The stars you are looking for could emerge from a rare evolutionary pathway: you may need many stellar models. These may take a long time to construct...
- Compare *like with like*, not  with !
- Selection effects could be important.

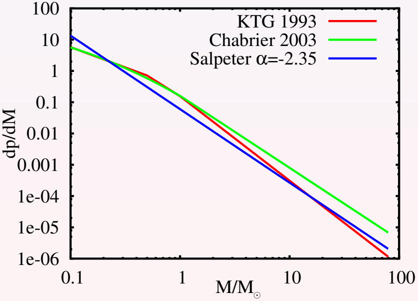
Stellar Birth: Rate

- Often we use a starburst or constant birth rate because this is the most simple.
- This may be justifiable for the Galactic disk. . .
- E.g. SFR (Chiappini et al 1997) for the solar neighbourhood . . .



Stellar Birth Functions

- Massive stars are *much* rarer than low-mass stars.
- The initial mass function (IMF) gives the relative number of stars of each mass
- Observation based, ~universal.



- Note: in binaries we require at least a mass distribution for the secondary star and the initial separation (or period) as well as the IMF.

Free(ish) Parameters

Each stellar model contains many parameters which you can set, these are usually called *free* parameters, even if you are not really free to choose them. . . so,
Either

- Set the parameter to a single value, if you do not think changing the value is important to your results, or
- Distribute the parameter according to some probability distribution function, if you think it is uncertain and will change your results.

Free(ish) Parameters

To distribute a parameter you must know how ignorant you really are about the parameter. Usually you don't have a clue, so it is customary to assume that all values of a parameter in a reasonable range are equally likely. (This is also the principle of least effort approach!)

- Often, free parameters are based on observations - do you trust these?
- Or specialized models of a difficult evolution phase (e.g. SPH or 1/2/3D hydrodynamics) - do you trust these?
- Or previous population synthesis studies! (Dangerous!)
- Perhaps you're trying to determine a parameter for the first time (hard, but somebody has to do it)

Rare Pathways - Population Resolution

- Many events or types of star are “rare”, in that they happen in only a small fraction of systems in the total population of stars and/or they do not last for long.
- e.g. type Ia supernovae, gamma ray bursts, NS-NS mergers, AGB phase.
- You should ensure that your resolution is high enough both:
  - In the number of models you have (e.g. mass-grid resolution) and
  - Your model timestep is short enough (time resolution).
- You just have to be careful.

Pilar with David, not Pilar with John

You must compare like with like.

- An example: if you are going to model the distribution of abundances in CH stars in your models, do not compare to solar neighbourhood red dwarfs!
- Often a certain type of star is observed, e.g. with a specific spectral type, and you have to select the appropriate star from your models.
- This is a simple and obvious point – but...
- Also, be very careful when combining observations from different papers into one survey for comparison with your models:
  - Systematic errors vary between surveys (next slide)
  - However, you may be able to do no better!

Model Errors

Statistical:

- Limited resolution: Poisson
- . . . run more models.

Systematic:

- Free parameters are uncertain.
  - Can you quantify these uncertainties?
  - Often not, but you can use a reasonable range of each parameter.
  - Each parameter adds a dimension to the grid of initial stars which in turn slows down your total simulation time.
- Model limitations, pushing model too far.
- Simple models cannot predict everything you want.

Selection Effects (Observer Errors)

GOLDEN RULE: Hope that the observers *have already taken selection effects into account*. Otherwise, modelling selection effects can be very hard, you may have to take into account:

- Small number statistics
- Luminosity (magnitude/flux) limit (Malmquist bias) – due to distance, dust etc.
- Volume limited samples are better, but be careful when the volume exceeds the local solar neighbourhood!
- If you are lucky the survey was carried out with one telescope and you might be able to remove some of its quirks, if you can find out what they are.

Simple Observable Accountancy

Comparing models to observables involve lots of adding up!

Accountant: $66,445, anti-dignity

Astronomer (Me): Much less, (some) dignity

1. Number/event counts
2. Use Ratios!
3. Distributions
4. SSPs (spectra!)

Observables 1: Number Counts

Simplest statistic we can calculate.

- 1 Model a population with n stars.
- 2 You want to compare your models to stars during some phase, so define

$$\begin{aligned}\delta(\text{phase}) &= 1 \quad \text{during the phase,}\\ &= 0 \quad \text{otherwise}\,.\end{aligned}$$

  (For *events* e.g. supernovae, $\delta = \delta(\text{time})$ really is a delta function)
- 3 Include your selection effects in $\delta(\text{phase})$!

Observables 1: Number Counts

- 4 For each model star (labelled i) add up the time spent in that phase

$$\Delta t_i = \sum_{t_{min}}^{t_{max}} \delta(\text{phase})_i \, \delta t \,,$$

  where the sampling is between times t_{min} and t_{max}.
- 5 In general, on a grid of n different masses M_i separation by dM, the birth function is given by

$$\Psi_i \;=\; \psi(M_i)\, dM$$

  where $\psi(M_i)$ is the initial mass function.
- 6 Modulate this with the birth probability Ψ_i and star formation S rate

$$S\Psi_i\Delta t_i \,.$$

Observables 1: Number Counts

- 7 Sum this for all the stars to get

$$\begin{aligned} \text{count} &= \sum_i \mathcal{S}\Psi_i \Delta t_i \\ &= \sum_i \mathcal{S}\Psi_i \sum_{t_{\text{min}}}^{t_{\text{max}}} \delta(\text{phase})_i \, \delta t \end{aligned}$$

- 8 Statistical error is $\propto 1/\sqrt{n}$ (computer-time limited)
- 9 Systematic error is ???

Observables 2: Use Ratios When You Can

- Use ratios to compare number counts.
- E.g. consider two number counts $[\sum_i S\Psi_i \Delta t_i]_1$ and $[\sum_i S\Psi_i \Delta t_i]_2$
- If S is constant, often $\sim$ true, then

$$\text{ratio} = \frac{[\sum_i \Psi_i \Delta t_i]_1}{[\sum_i \Psi_i \Delta t_i]_2}$$

- Removes need for star formation rate in calculations
- Requires S to be constant or starburst (set $t_{\text{min,max}}$ appropriately)
- e.g. Galactic disk in last 5 Gyr, globular clusters, starburst galaxies, ellipticals with care

Observables 3: Synthetic Distributions

- Often you want more than a count, you want a distribution of something
- Just extra details in the accounting process
- Replace $\delta(\text{phase})$ with a variable w to weight the output
- Add up, bin results, et voila
- For some examples, see later slides...

Observables 4: Single Stellar Populations

- This technique is used to calculate synthetic spectra, e.g. to calculate the flux at $x\,\text{Å}$:

$$\text{flux at}\,x\,\text{Å}(t) \;=\; \sum_{i} \Psi_i \sum_{t}^{t+\delta t} (\text{flux at}\,x\,\text{Å from star}\,i)\,\delta t\,,$$

  although note that calculating "flux from star $i(x)$" is non-trivial (Spectral library or model)
- This is what you will often see referred to *(spectral) population synthesis* in the literature.
- Often calculated for a starburst - these are called *simple/single stellar populations (SSPs).* (then convoluted with a SFH to make a galactic model)
- These are *not* what I do: Stellar evolution is much more exciting!

Model Problems: The Need For Synthetic Codes...

- In an ideal world, you have a good, detailed model to compare to observations:
  - accurate
  - high resolution (e.g. for rare objects)
- In the real world, you do not
- e.g. MSSSP code + nucleosynthesis for *one star* may take weeks! (just CPU)
- Impractical for simulations with large parameter space
- Hence *Synthetic Models*

Synthetic (Single) Stellar Models

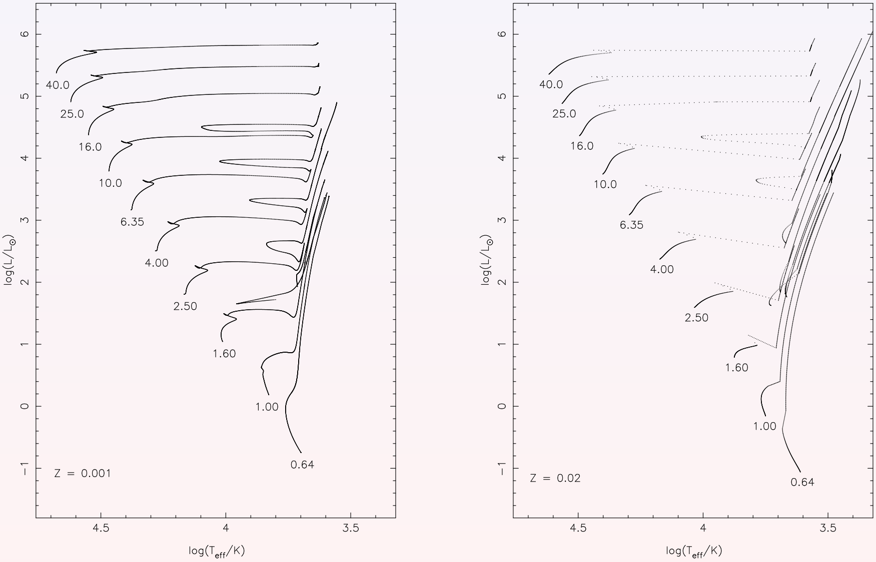
- Combine:
  - Pre-computed stellar models
  - "Extra" algorithms
- To make a fast "synthetic" code which reproduces the slow code results and extends the physics
- e.g. Fit stellar model results e.g. burning lifetimes, L, R, M_c to "simple" functions
- Code is faster, $\sim 10^7$ times!, but contains $\gg\gg\gg$ information

SSE code

- Full stellar evolution for $0.1 \leqslant M/M_{\odot} \leqslant 100$, $10^{-4} \leqslant Z \leqslant 0.03$
  - e.g. SSE code (Hurley et al 2000 MNRAS 315 543)
  - Fitted to models constructed with Eggleton's stellar evolution code
  - Added variable stellar wind
  - But the AGB phase is approximate (skips pulses)

Fitting example: HR Diagram from SSE



Pros and Cons of Synthetic Models

Pros

- Fast, stable
- Extra algorithms probe new physics
- Lacks details

Cons

- Only as “good” as (wrong) input models
- Dangerous to interpolate - be careful!
- Very dangerous to extrapolate!
- Lacks details