The statistical applications on the galaxies and AGNs in SDSS

Shiyin Shen 沈世银
2014/12/23  USTC

Shanghai Astronomical Observatory
Collaborators: Zhengyi Shao, Xiaohu Yang, Xiaoyi Xie, Houjun Mo et al.
Outline

- Introduction
  - The Sloan Digital Sky Survey
  - The large galaxy and QSO samples in SDSS legacy
  - The classical statistical studies in SDSS

- Our statistical applications
  - The statistical nature of the brightest group galaxies;
  - The tile between accretion disk and stellar disk;
  - The redshift dependence of the composite quasar spectra: implications on the cosmic dust extinction.
Why statistical studies?

- Common properties
  - Constraints to physical models
  - Physical implications

- Astronomical units: no experiment
  - Observe a galaxy from different direction 😞
  - Observe many galaxies from the same direction 😊

- Cosmological principle
Shortcomings of statistical studies

- Correlation ≠ causality

- A city with more bars also with more book stores (Antonucci)
  - Doesn’t mean people like reading also like drinking
Modern galaxy survey

- Developed together with the increasing of computational power and storage capability.

- Sloan Digital Sky Survey (SDSS, Since 2000)
  - 2.5 m special designed survey telescope
  - (u,g,r,i,z) five bands photometry
  - Multi-fiber spectroscopy
    - 640 fibers in SDSS I,II
    - 1000 fibers in SDSS III
  - ~50 TB data (DR8)

- Next generation
  - LSST: 1 PB raw data per year

http://www.sdss.org/

- **Photometry**
  - ~8400 sq deg
  - u,g,r,i,z to (22.0, 22.2,22.2,21.3,20.5) mag

- **Spectroscopy**
  - Main galaxy sample (r<17.77)
    - ~700k, z~0.1
    - completeness ~90% (another 10% LAMOST)
  - Luminous red galaxies
    - ~300k, Volume-limited to z~0.38 and up to z~0.55
  - QSO sample
    - >100k QSOs with z up to ~5
    - $M_i < -22$
SDSS III and IV

- SDSS III (2009-2014)
  - APOGEE, MaNGA, eBOSS
  - We are members! (SHAO, NAO, SJTU, NJU)
Classical statistical studies in SDSS (I) Basic distributions

- Luminosity function
- Stellar mass function

\( L^* \sim 3 \times 10^{10} L_\odot \)

\( \phi(L) \) vs. Galaxy luminosity

e.g. Croton et al. 2006

Blanton et al. 2003

Li & White 2007
(II) Conditional distributions

- Conditional luminosity function
  - Halo mass $P(L|M_h)$ (Yang et al. 2005)
  - Morphology (Nakamura et al. 2003)
  - Inclination (Shao, Xiao & Shen 2007)
  - Environment (Tempel et al. 2010)
  - See Blanton & Moustakas (ARA&A, 2009)

- Size distribution $f(R|M)$ (Shen et al. 2003)
  - Surface brightness distribution
(III) Scaling relations

- Mass – metallicity relation
  - Fundamental Metallicity Relation
    - SFR-Mass-Metallicity (Mannuci et al. 2010)
  - The star formation history of galaxies

- Tully-Fisher relation and Fundamental Plane
  - TF: spiral galaxies
    - $L = A V_{rot}^a$ (Shen et al. 2002, 2009)
  - FP: elliptical galaxies
    - $\log R_e = a \log \sigma + b \log I_e + c$ (e.g. Bernardi 2003)
  - Dark matter halo VS stellar mass

- M-sigma relation (e.g. Gu 2009)
  - The co-evolution between the galaxy and black hole
Can statistical studies tell us more?

- Theoretical models
  - Basic distributions
  - Conditional distributions
  - Scaling relations

Data

Statistical applications

Questions
Key ingredients of statistical applications

- Large sample
  - Case I: the statistical nature of the brightest group galaxies

- Selection effects: control sample
  - Case II: the tilt between the accretion disk and stellar disk

- Statistical hypothesis
  - Case III: the cosmic dust extinction: implications from quasar composite spectrum

- Brightest group/cluster galaxies (BGGs)
  - Center of group/cluster scale haloes
    - Extreme brightness in rich cluster, small scatter (0.3 mag)
  - Red ellipticals
    - Old and passively evolved stellar population
  - Formation
    - Subsequent merge of small systems, assemble most of their mass at early time

- BGGs show specialties in observations
  - More extended surface brightness profile than $r^{1/4}$ law
    - (Bernardi et al. 2007)
  - On fundamental plane, different from the normal ellipticals
    - (Von der Linden 2007)
  - More likely to be radio-loud AGN (Von der Linden 2007)
Bernardi et al. 2007

Von der Linden 2007
**Physical view of BGGs**

- BGGs are the central galaxies of a halo, the other galaxies are satellites.
- The centrals are distinct from satellites

We split the CLF in a **central** and a **satellite** term:

\[
\Phi(L|M) = \Phi_c(L|M) + \Phi_s(L|M)
\]

For **centrals** we adopt a log-normal distribution:

\[
\Phi_c(L|M) \, dL = \frac{1}{\sqrt{2\pi} \sigma_c} \exp\left[-\left(\frac{\ln(L/L_c)}{\sqrt{2}\sigma_c}\right)^2\right] \frac{dL}{L}
\]

For **satellites** we adopt a modified Schechter function:

\[
\Phi_s(L|M) \, dL = \frac{\phi_s}{L_s} \left(\frac{L}{L_s}\right)^{\alpha_s} \exp\left[-(L/L_s)^2\right] \, dL
\]
$M_h - M_1$ relation

The BGGs are brighter in higher mass halos.

$$\bar{L}_1 = L_0 \frac{(M_h/M_1)^{\gamma_1}}{[1 + (M_h/M_1)]^{\gamma_1 - \gamma_2}},$$

e.g. Zheng et al. 2007
**Statistical View of BGGs**

- BGGs may be special only because of their definition.

- **Order statistics (OS):** statistics on the $j$-th largest members of groups with $N$ members
  - Extreme value statistics (EVS) correspond to OS of $j=1$
  - Basic assumption: all group members are drawn from the same population
  - $N$: sample size

- **Basic conclusion of the OS/EVS**
  - For the same population, if the sample size $N$ is larger, the extreme value is larger
    - The BGGs are brighter in higher mass (richer) halos!
Chinese basketball team  
Average height: 2.0m  
Population: 1.3 billion

Japanese basketball team  
Average height: 1.95m  
Population: 0.13 billion
Other conclusions of EVS/OS

- The scatter of the extreme values
  - decrease with the increasing of the sample size $N$
    - the brightest members of rich clusters: standard candle (0.3 mag, Postman 1995)
  - decrease with the increasing of the order $j$
    - The scatter of the Second brightest member galaxies is smaller than BGGs

Dobos & Csabai 2010
SDSS DR7 group catalog (Yang et al.)

- **Richness** $N_{19.5}$
  - Number of galaxies $M_r < -19.5$

- **complete to** $z < 0.09$
  - $M_r = -19.5$ (r=17.77 at $z=0.09$)

- **Number of groups** $N_{19.5} > 1$
  - 16,014

- **Number of group members**
  - 62,081
The statistical groups have the same richness and underlying distribution as the real groups.
OS prediction VS observation: BGG

The OS predicted BGGs are systematically fainter than the real BGGs.

The OS predicted correlation between the BGG luminosity and group richness (halo mass) is similar as observations.
The second brightest group galaxies
Model the BGGs

- At the time of group formation, the stochastic merging process makes the BGGs and other members follow the OS;

- After BGG settling into the center of the group, the local environmental effects further brighten the BGG, e.g. galaxy cannibalism, star formation in cooling flow etc.

BGG brightening: $M_1 = M_0 - \Delta M$

Dim the luminosities of BGGs with magnitude correction $\Delta M$

For each richness bin, shuffle the members of the BGG-dimmed sample

Build $K$ statistical groups, the same $N_i (i=1,K)$ distribution as the real groups

Model groups: apply BGG brightening $\Delta M$ on the statistical BGGs

test the distributions of the real group members VS the model groups
The average brightening of BGGs

- The BGGs are averagely brightened ~0.2 mag
- Corresponding to a stellar mass increment of about 20 percent
- Consistent with the minor merge scenario of the central galaxies (e.g. Edwards & Patton 2012)

The BGG brightening ($\Delta M$) is not stochastic, but depends on how prominent is the BGG inside the group $\Delta M = 0.3 \ G_{1,2}$
A simple HOD and CLF modelling

- For given halo $M_h$, estimate the group richness $N$;

- Assume a conditional luminosity function;

- Assign $N$ members to halo according to CLF;

- Pick up the brightest one as the central;

- Brighten the luminosity of the central according to $\Delta M \approx 0.5G_{1,2}^*$
Summary on the statistical nature of BGGs

- The brightest group/cluster galaxies (BGGs) are special, but not distinct from other members.

- The BGGs are consistent with a scenario that they are formed in a statistical process and later brightened (~0.2 mag) by local processes.
(II) The tilt between accretion disk and stellar disk (Shen, Shao & Gu, 2010, ApJL)

Any alignment between the Kpc scale stellar disk and sub-pc scale dust tours (accretion disk)?
**AGN unification model**

**Orientation:**
- the key parameter of the AGN schema

**Type II AGNs:**
- edge-on view from obscuring torus (accretion disk)

**Hypothesis:** if the stellar disks show alignment with accretion disk, the host galaxies of type II AGNs will also be biased to edge-on
AGNs hosted by disk galaxies

- $b/a$ of disk galaxies
  - approximate the inclination of disk
- $\text{fracDev}<0.5$: select disk galaxies
  - $F_{\text{comp}} = \text{fracDev} F_{\text{dev}} + (1 - \text{fracDev}) F_{\text{exp}}$
  - $\sim 40\%$ of galaxies are spiral
- AGNs hosted by disk galaxies
  - BPT diagram
  - $\sim 30\%$ of AGNs hosted by spirals

$N_{\text{AGN, spiral}} = 32,618$
~32,600 type 2 AGNs hosted by disk galaxies in SDSS DR7

Distinctive physical properties of AGN host

Spatial correlation between accretion disk and stellar disk

B/a distribution

A control sample with the same physical properties as AGN hosts
Control sample of non-AGN galaxies

The same distributions of

- Stellar mass
- Redshift
- Concentration
- D4000
- R50
- FracDev
Type 2 AGN hosts are biased to low inclination (edge-on) disks!
The orientation of the accretion disk and stellar disk is correlated.

How to quantify this result?
The shape of spiral galaxies

- Incl = f(b/a | γ, ε, θ, ψ)

- (γ, ε) stellar disk structure parameter
  - γ = C/A : disk height
  - ε = (1 - B/A) : disk ellipticity
  - AGN hosts have the same (γ, ε) distribution as control galaxies!

- (θ, ψ) viewing angle
  - θ: inclination  ψ: position angle
  - Control galaxies: random
  - AGN hosts: preferred angle?
\( \gamma, \varepsilon \) distributions

Non-negative least square linear regression
viewing angle of AGN hosts

- Viewing angle of stellar disk \((\theta_s, \phi_s)\)

- Viewing angle of accretion disk \((\theta_A, \phi_A)\)

- Tilt angle between accretion disk and Tilt stellar disk \(\Delta_{AS}\)

- Opening angle of torus \(\Theta_T\)
  - Type II AGN: \(90 - \theta_A < \Theta_T\)
**Toy model**

Random viewing angle of stellar disk \((\theta_S, \phi_S)\)

Tilt angle \(\Delta_{AS}\)

Viewing angle of accretion disk \((\theta_{A0}, \phi_{A0})\)

\((\gamma, \varepsilon)\) distribution

\(b/a\) of galaxy

Viewing angle of observed AGN \((\theta_A, \phi_A)\)

\((\gamma, \varepsilon)\) distribution

\(90 - \theta_{A0} < \Theta_T\)

Type II AGN fraction \(f_2\)

\(N_{AGN}/N_{AGN}(b/a)\)

\(b/a\) of AGN host

\(\frac{b}{a}\) of galaxy

\(N_{AGN}/N_{AGN}(b/a)\)
**Model results**

- If $\Delta_{AS}$ random
  - $\cos\Delta_{AS}$: uniform distribution
  - $N_{AGN}/N_{GAL} = 1$
- Non-random Assumption
  - $\cos\Delta_{AS}$ follows Gaussian distribution, with most probable value $\Delta_{AS} = 0$, scatter $\Delta_{AS,m}$
  - $\Delta_{AS,m} = 37$ deg
- $\Theta_T$: free parameter
  - $\Theta_T = 58$ deg
  - $f_2 = 80\%$: in excellent agreement with the results in literature!
Summary on the tilt angle between accretion disk and stellar disk

- The AGN host galaxies (if they are spirals) are more likely to be viewed as edge-on.
  - Consistent result from radio AGNs (new, arXiv:1412.5602)

- The bias in the axis ratio (inclination) of the AGN hosts can be explained with a simple model that
  - the opening angle of the torus is ~60 degree
  - the mean tilt angle between the accretion disk and stellar disk is ~30 degree

- SN Ia cosmology
  - Standard candle

- Cosmic dust extinction?
  - How to measure it?
    - Reddening of background sources (Menard et al. 2010)
    - Compare luminosity distance and angular distance (More et al. 2009)
    - Quasar color (Johanson & Mortsell 2012)
The quasar composite spectrum

- The mean quasar spectrum
  - High S/N
  - large wave-length coverage
  - Quasar spectra are similar

- Composite spectrum from SDSS DR7
  - Better redshifts (Hewett & Wild 2010)
  - Better Sky subtraction (Wild & Hewett 2010)
  - Physical parameters of each quasar (Shen et al. 2010)
  - Exclude radio-loud and broad absorption line quasars
Quasar spectrum: any dependence?

- Luminosity dependence
  - UV continuum of higher luminosity quasars is harder (Telfer et al. 2002; Davis et al. 2007; Carballo et al. 1999).

- Redshift dependence (evolution)?
  - slight hardening of the spectra towards higher redshift (Kuhn et al. 2001; Carballo et al. 1999)
  - no evolution (Pentericci et al. 2003)
  - Shallow (redder) at high redshifts (Wright 1981)
Our results

Luminosity dependence!

SDSS quasars are flux-limited sample, correlation between luminosity and redshift!

No redshift dependence?
Redshift dependence after luminosity constrained
What makes the quasars at higher redshifts being redder?

- Physical evolution?
- Quasar host galaxies?
  - High z galaxies are more transparent
- Selection effects?
- Cosmic dust extinction
  - $A_B \sim 0.1 \, z$
  - Up to redshift $z \sim 3$
Summary on the redshift dependence of quasar composite spectrum

- We find a redshift dependence of the UV slope of the quasar spectrum
  - The quasars are redder at higher redshift.

- This reddening can be explained by the cosmic dust extinction
  - The average extinction is about $A_B \approx 0.1z$. 
Summary

- Modern galaxy survey, like the Sloan Digital Sky Survey (SDSS), provides unprecedented large sample of galaxies and AGNs (QSOs) for statistical study.

- Besides the classical statistical studies, e.g. (conditional) distribution function, scaling relations,
  - the large sample of galaxies and AGNs can be used to test many hypothesis in the modelling of the galaxy formation and evolution
  - once the selection effects have been properly taken into account
Key ingredients of statistical applications

- **Large sample**
  - Case I: the statistical nature of the brightest group galaxies
    - Thanks to the largest group catalog of Yang et al.

- **Selection effects: control sample**
  - Case II: the tilt between the accretion disk and stellar disk
    - A control sample of non-AGN galaxies with the other properties controlled

- **Statistical hypothesis**
  - Case III: the cosmic dust extinction: implications from quasar composite spectrum
    - Our key assumption: Quasars are redder at higher redshifts due to the cosmic dust extinction.
Out statistical applications show

1. The brightest group/cluster galaxies are special, but not distinct from other members. They are formed in a statistical process and later brightened (~0.2 mag) by local processes.

2. There is an alignment between the accretion disk and stellar disk for AGNs hosted by spirals, the typical tilt angle ~30 deg.

3. The cosmic dust extinction can be probed from the composite quasar spectra up to z~3, the typical extinction $A_B \sim 0.1z$. 
Thanks for your attention

MERRY CHRISTMAS