

Esperienza 5

Stima di massa totale e frazione di massa stellare per ammassi di galassie

Nicola Bianchetti
Matteo Bianchi
Diana Capelli
Riccardo Loddo

Università degli Studi di Milano

11 agosto 2022

Sommario

L'obiettivo dell'esperienza è quello di stimare la massa totale di un ammasso di galassie utilizzando i dati ottenuti dalla *Sloan Digital Sky Survey* [1] e dal telescopio *Subaru*[4]. Per stabilirne il valore sono state applicate tecniche di analisi Bayesiana, sfruttando un modello basato sul teorema del viriale.

Indice

1. Modello	2	
1.1 Modello fisico		2
1.2 Modello statistico.....		3
2. Dati	3	
3. Analisi dati e risultati	3	
4. Conclusioni	6	

1. Modello

1.1 Modello fisico

Dalle osservazioni astronomiche risulta che le galassie non si distribuiscono uniformemente nello spazio, ma tendono a disporsi in gruppi e ammassi. Tipicamente gli ammassi contengono un numero di galassie dell'ordine delle centinaia o delle migliaia e pertanto hanno un valore di massa maggiore di $10^{14} M_{\odot}$. Per identificare le galassie appartenenti ad un determinato ammasso in una data regione di cielo vengono confrontati i valori di redshift z . Per piccoli z ¹ è possibile approssimare la velocità di una singola galassia con $v \approx c \cdot z$ (dove c è la velocità della luce) e dalla legge di Hubble $v = H_0 \cdot D$ (dove H_0 è la costante di Hubble, presa pari a $70 \text{ km Mpc}^{-1} \text{ s}^{-1}$) si può stimare la sua distanza:

$$D \simeq \frac{c}{H_0} \cdot z. \quad (1)$$

Il redshift misurato per una galassia è la somma del contributo dovuto all'espansione cosmologica dell'Universo e di un termine di moto peculiare della stessa nell'ammasso. Questi due termini possono essere tenuti in considerazione separatamente supponendo che il segnale ricevuto dagli strumenti, λ_{oss} , sia la somma di un segnale emesso dalla galassia e misurato nel centro dell'ammasso, λ_0 , e del segnale proveniente dal centro dell'ammasso e misurato sulla Terra, λ_{cl} , secondo la relazione:

$$1 + z_{\text{oss}} = \frac{\lambda_{\text{oss}}}{\lambda_{\text{cl}}} \cdot \frac{\lambda_{\text{cl}}}{\lambda_0} = (1 + z_{\text{pec}}) \cdot (1 + z_{\text{cl}}), \quad (2)$$

dove z_{pec} è il redshift causato dalla velocità peculiare della galassia nell'ammasso e z_{cl} quello complessivo del cluster.

Le velocità delle galassie dell'ammasso nel sistema di riferimento dell'ammasso si possono così stimare:

$$v_{\text{pec}} = \frac{z_{\text{oss}} - z_{\text{cl}}}{1 + z_{\text{cl}}} \cdot c \quad (3)$$

e nel nostro modello assumiamo che seguano una distribuzione normale. Oltre a considerare valori di z piccoli secondo quanto indicato in precedenza, assumiamo che anche la distribuzione degli z sia una distribuzione gaussiana centrata in z_{cl} . Infine poniamo $\langle v_{\text{pec}} \rangle = 0$, supponendo che l'ammasso non ruoti o si espanda. La dispersione delle velocità lungo la linea di vista si ottiene poi dalla formula generale per la varianza²:

$$\sigma_v = \sqrt{\langle v_{\text{pec}}^2 \rangle} = \frac{c}{1 + \langle z \rangle} \cdot \sigma_z. \quad (4)$$

Per stimare la massa totale dell'ammasso è stato utilizzato il *teorema del viriale*, che permette di descrivere le proprietà medie di un sistema in condizioni di equilibrio, mettendo in relazione energia cinetica K e potenziale U :

$$\langle U \rangle + 2\langle K \rangle = 0. \quad (5)$$

Per un ammasso di galassie è possibile utilizzare il modello di sistema sferico di raggio \tilde{r} , contenente N galassie, tutte di massa m . Dalla relazione (5) e assumendo che la distribuzione di velocità sia isotropa per le galassie con $\langle v_{\text{pec}} \rangle = 0$, si ottiene la relazione:

$$\sigma_v^2 = \frac{\langle v^2 \rangle}{3}. \quad (6)$$

Confrontando l'energia potenziale gravitazionale per una distribuzione sferica uniforme di massa totale M e raggio \tilde{r} ³ con l'energia puramente cinetica, si ha:

$$-3m\sigma_v^2 \approx -\frac{3GM^2}{5N\tilde{r}}; \quad (7)$$

¹Sono considerati redshift piccoli valori di z minori di 0.2.

² $\sigma_v^2 = \langle v_{\text{pec}}^2 \rangle - \langle v_{\text{pec}} \rangle^2$.

³ottenuta integrando $dU = \frac{-GM_r dm}{r} = \frac{-GM_r 4\pi r^2 \rho dr}{r}$ sull'intervallo $[0, \tilde{r}]$.

Considerando poi che $M = Nm$, perveniamo a:

$$M_{vir} \approx \frac{5\sigma_v^2 \tilde{r}}{G}, \quad (8)$$

che fornisce la massa totale dell'ammasso.

1.2 Modello statistico

Il modello appena descritto dipende dalla dispersione di velocità σ_v delle galassie nell'ammasso, la quale a sua volta dipende dal valor medio del redshift dell'ammasso \bar{z} . La distribuzione del redshift dell' n -esima galassia appartenente all'ammasso è ben approssimata da una distribuzione normale con media pari al valor medio del redshift dell'ammasso e varianza σ_z^2 , quindi $z_{gal,n} \sim N(\bar{z}, \sigma_z^2)$. La distribuzione a posteriori di tali parametri è stata ottenuta tramite il teorema di Bayes, a partire dai dati di redshift $z_{gal} = \{z_{gal,n}\}$:

$$P(\bar{z}, \sigma_z^2, f | z_{gal}) = \frac{P(z_{gal} | \bar{z}, \sigma_z^2, f) P(\bar{z}, \sigma_z^2, f)}{P(z_{gal})}. \quad (9)$$

Assumendo che le misure di redshift siano tra loro indipendenti, la distribuzione di verosimiglianza, ovvero $P(z_{gal} | \bar{z}, \sigma_z^2, f)$, è data da:

$$P(z_{gal} | \bar{z}, \sigma_z^2, f) = \prod_n \left[f \frac{1}{\sqrt{2\pi(\sigma_z^2)}} \exp\left\{-\frac{(z_{gal,n} - \bar{z})^2}{2\sigma_z^2}\right\} + (1-f)D(z_{gal,n}) \right], \quad (10)$$

dove si è introdotto un terzo parametro f che rappresenta la probabilità che la galassia associata alla misura $z_{gal,n}$ appartenga all'ammasso, mentre $D(z_{gal,n})$ è una distribuzione, scelta uniforme, che descrive i valori di redshift di galassie che non appartengono all'ammasso considerato che potrebbero fornire del rumore di fondo nei dati di redshift. La distribuzione a priori $P(\bar{z}, \sigma_z^2, f)$ è stata costruita assumendo che i parametri siano distribuiti in modo uniforme in opportuni intervalli. L'evidenza $P(z_{gal})$ fornisce la normalizzazione alla distribuzione a posteriori⁴.

2. Dati

In questa esperienza sono stati elaborati inizialmente i dati provenienti dall'ottava pubblicazione (DR8) della *Sloan Digital Sky Survey* (SDSS [1]), per lo studio dell'ammasso di galassie *Abell 2142*. Questi dati comprendono le coordinate celesti ascensione retta RA e declinazione Dec , i valori di magnitudine in banda r m_r e in banda u m_u , i valori di redshift z e la massa stellare totale in unità di masse solari M_* con il relativo errore⁵ σ_{M_*} . Per studiare l'ammasso sono state considerate le galassie con un redshift z , senza problemi nella determinazione, contenute in un cerchio di raggio 20 arcmin, centrato nelle coordinate celesti: $RA = 239.5833^\circ$ e $Dec = 27.2334^\circ$. Successivamente sono stati analizzati i dati provenienti dal telescopio *Subaru* [4], relativi all'ammasso identificato come *MACS J1206* sulla base delle sue coordinate di riferimento. Il catalogo comprende le coordinate celesti di ascensione retta RA e declinazione Dec , il valore di redshift z e la magnitudine in banda r m_r . Le galassie considerate in questo secondo caso sono quelle contenute in un riquadro di larghezza angolare pari a circa 25 arcmin centrato nelle coordinate celesti: $RA = 181.5508$ e $Dec = -8.8002$.

3. Analisi dati e risultati

Considerando il primo ammasso di galassie, *Abell 2142*, la distribuzione a priori, necessaria per l'eq.(9), è stata costruita come prodotto delle distribuzioni uniformi sui parametri \bar{z} e σ_z , all'interno degli intervalli

⁴si veda la sezione Analisi dati [3] per maggiori dettagli.

⁵calcolato come la semi differenza tra l'84-esimo e il 16-esimo percentile.

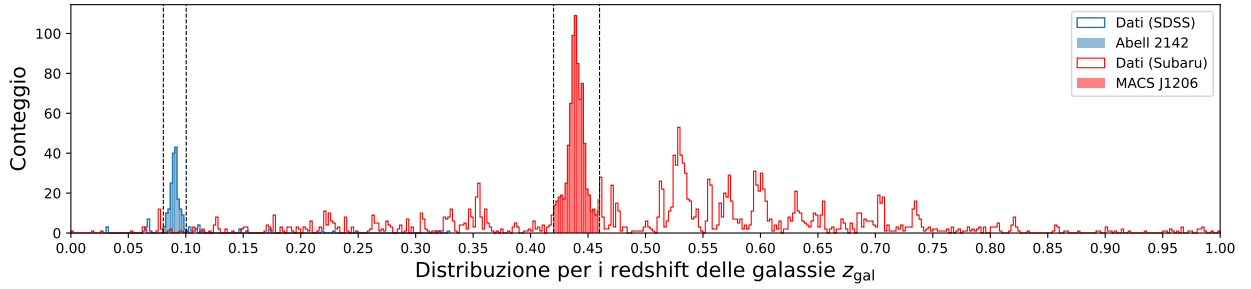


Figura 1: Il grafico riporta la distribuzione sperimentale per i dati di redshift osservati z_{gal} , ristretti all'intervallo $[0,1]$. I due ammassi sono rappresentati sulla stessa scala per un confronto diretto dei due insiemi di dati.

$\bar{z} \in [0.03, 0.33]$ e $\sigma_z \in (0, 0.3]$, scelti in modo tale da risultare compatibili con gli intervalli di redshift selezionati nel corso delle analisi.

Per il secondo ammasso i dati sono stati selezionati tenendo in considerazione il fatto che, secondo il modello statistico (sez.[1.2]), z deve seguire una distribuzione gaussiana. Osservando l'andamento del numero di galassie in funzione del redshift (Fig. 1) riscontriamo un picco evidente per $z = 0.44$ nell'intorno del quale è stata svolta l'analisi.⁶ Considerando tutti i dati in Fig.1, infatti, non sarebbe stato possibile sfruttare una semplice distribuzione uniforme per la mistura di probabilità senza prima trascurare gli altri picchi, probabilmente dovuti a ulteriori gruppi di galassie nella stessa porzione di cielo. Si è dunque considerato l'intervallo di dati $\bar{z} \in [0.42, 0.46]$ e $\sigma_z \in (0, 0.04]$, in modo tale che la distribuzione ipotizzata sia adeguatamente rappresentativa dell'andamento sperimentale del redshift.

In entrambe le analisi, la distribuzione a priori per il parametro f presente nella sez.[1.2], è stata considerata entro il suo intervallo di definizione $f \in [0, 1]$.

Per stimare le distribuzioni a posteriori dei parametri \bar{z} , σ_z e f relativi a entrambi gli ammassi, sono stati utilizzati algoritmi di tipo Monte Carlo basati sulle Catene di Markov (MCMC), implementati tramite la libreria *emcee*[3]. Grazie al campionamento discreto della distribuzione a posteriori, non è stato necessario calcolare l'evidenza data dalla sua normalizzazione. I risultati così ottenuti sono rappresentati tramite il *corner plot*[2] riportato in Fig.2.

Dalla distribuzione a posteriori delle \bar{z} del primo ammasso, sfruttando l'eq.(1), è stato possibile ottenere la distribuzione per la distanza. Nel caso del secondo ammasso, essendo esso caratterizzato da un redshift medio \bar{z} più elevato, è stato necessario utilizzare un calcolatore cosmologico [5] in cui è stata assunta una costante $\Omega_M = 0.3$ per un universo piatto, convertendo quindi i percentili ottenuti dalla distribuzione di \bar{z} . Il raggio viriale del primo ammasso è stato calcolato tramite la relazione trigonometrica $R = D \tan \theta \approx D\theta$ dove θ è l'apertura angolare (20 arcmin) e D la distribuzione della distanza dell'ammasso. Per il secondo ammasso, si è considerato un raggio viriale $R = 2$ Mpc. La dispersione di velocità delle galassie per entrambi gli ammassi è stata calcolata usando l'eq.(4), usufruendo delle distribuzioni per z e σ_z calcolate tramite *emcee*. Usando l'eq.(8), si sono ottenute poi le distribuzioni per i valori di massa totale degli ammassi. In Tab.1 sono riportati i risultati dell'analisi per entrambi gli ammassi.

Infine, è stato possibile stimare il rapporto fra la massa stellare e la massa totale del primo ammasso, considerando soltanto le galassie che si trovano entro una distanza di $3\sigma_z$ dal centro della distribuzione di redshift dell'ammasso⁷, arrivando a $N_{\text{Abell}} = 174$ galassie. Per tale rapporto è stato quindi ottenuto un

⁶Tra i dati rappresentati nell'istogramma non sono presenti valori di redshift identicamente uguali a $z = 0$, probabilmente dovuti a stelle della nostra galassia identificate per errore come sorgenti estese. Per valori elevati di redshift ($z > 5$) si rilevano invece immagini multiple delle stesse sorgenti, i cui spettri di emissione risultano duplicati a causa del fenomeno di *lensing gravitazionale* esercitato dall'ammasso.

⁷In tal caso, i valori di \bar{z} e σ_z utilizzati per la selezione sono stati presi come il 50-esimo percentile delle rispettive distribuzioni.

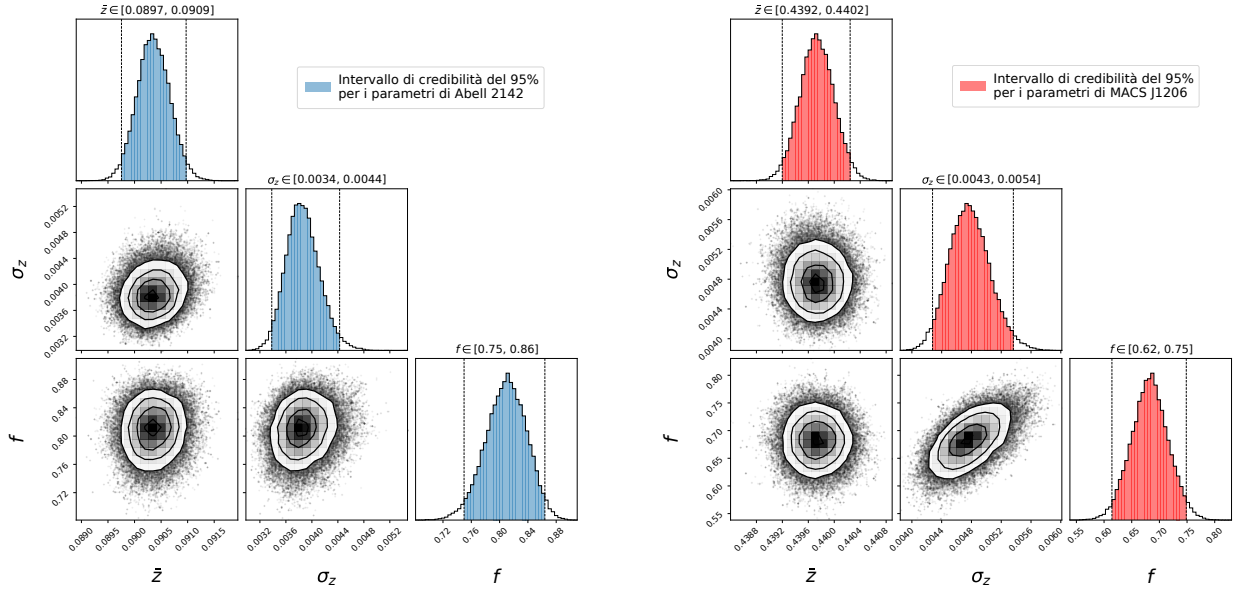


Figura 2: In figura sono riportate le distribuzioni a posteriori dei parametri z , σ_z e f assieme ai grafici delle relative covarianze per entrambi gli ammassi. In particolare, nel grafico di sinistra vengono riportate le distribuzioni per l'ammasso *Abell 2142* e a destra per *MACS J1206*.

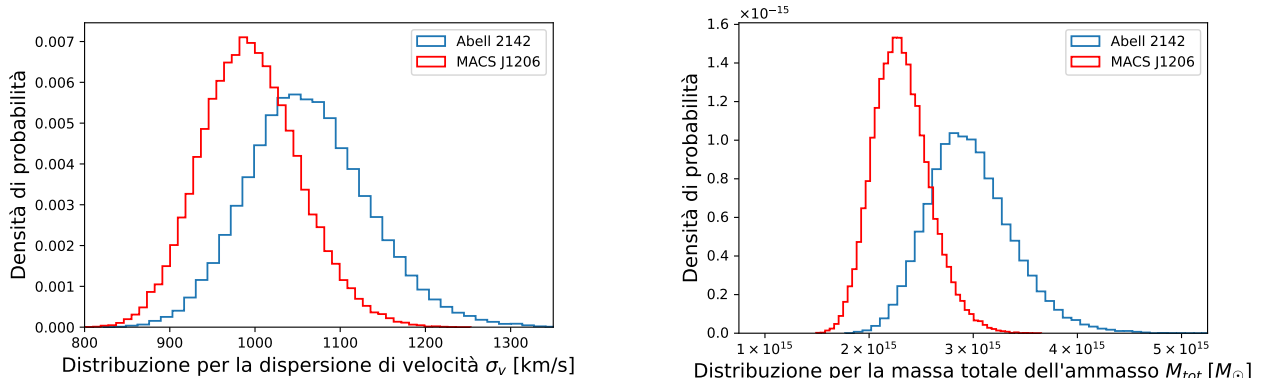


Figura 3: Nei grafici sono riportate le distribuzioni a posteriori ottenute per la dispersione di velocità σ_v e la massa totale M_{tot} di entrambi gli ammassi. Ognuna delle distribuzioni è stata ricavata a partire dal campionamento effettuato da *emcee*

Ammasso	\bar{z}	σ_z	f	D [Mpc]	M_{tot} [$\log M_\odot$]
Abell 2142	[0.0897, 0.0909]	[0.0034, 0.0044]	[0.75, 0.86]	[384.7, 389.9]	[15.36, 15.88]
MACS J1206	[0.4392, 0.4402]	[0.0043, 0.0054]	[0.62, 0.75]	[1171.5, 1173.1]	[15.26, 15.46]

Tabella 1: In tabella sono riportati i valori del redshift medio e la sua deviazione standard, il parametro di mistura della distribuzione normale, la distanza dell'ammasso e la massa totale stimata dall'analisi per entrambi gli ammassi. Tutti i risultati hanno un intervallo di credibilità corrispondente al 95% delle distribuzioni a posteriori per i parametri.

valore pari a $(0.40 \pm 0.05)\%$, il cui errore è stato ottenuto propagando l'incertezza sulle misure di massa stellare e la deviazione standard ricavata dalla distribuzione di massa totale.

4. Conclusioni

I risultati numerici dell'analisi, sotto forma di intervalli di credibilità del 95% sono riassunti nella Tab.1. Nel caso dell'ammasso *Abell 2142*, come descritto nella sezione precedente, è stato possibile calcolare la percentuale di massa stellare su massa totale. Tipicamente sappiamo che il $\sim 75 - 85\%$ della massa totale di un ammasso è costituita da materia oscura e il $\sim 15\%$ da ICM⁸, mentre la percentuale restante dovrebbe essere di massa stellare. Tuttavia il rapporto così ottenuto per la frazione di massa stellare ($f_* = 0.4\%$) risulta sottostimato, principalmente a causa del fatto che i valori di massa stellare riportati nel catalogo SDSS potrebbero non corrispondere alla totalità delle galassie appartenenti all'ammasso, oltre alla selezione effettuata per incompletezza nella determinazione del redshift. L'ipotesi iniziale che lo spettro di emissione osservato per le galassie fosse un tracciante valido dell'effettiva massa stellare, ha originato una stima per la massa di $M_* = 10^{13} M_\odot$, risultante dalle sole $N_{\text{Abell}} = 174$ galassie considerate che, tuttavia, per un ammasso con massa totale $M_* = 10^{15} M_\odot$ non sono sufficienti per fornire un campione completo e rappresentativo. Osservando i risultati raffigurati nel *corner plot* per il secondo ammasso (Fig.2), è invece possibile notare una correlazione diretta tra i parametri σ_z e f . L'introduzione di quest'ultimo parametro all'interno del modello di mistura consente infatti di assegnare un peso diverso alla distribuzione gaussiana rispetto al contributo di quella uniforme, in modo tale da permettere a σ_z di assumere valori più aderenti al reale andamento della distribuzione per i redshift. Tuttavia, il fatto che il parametro risulti appartenere all'intervallo $f \in [0.62, 0.75]$ è indice di un moderato disturbo presente nell'insieme di dati considerato, non riducibile a partire da un semplice taglio in corrispondenza di $3\sigma_z$, che porterebbe a un numero $N_{\text{MACS}} = 706$ di galassie in confronto alle 780 inizialmente considerate sull'intervallo $z \in [0.42, 0.46]$.

Riferimenti bibliografici

- [1] H. Aihara et al. «The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III». In: 193, 29 (apr. 2011), p. 29. DOI: 10.1088/0067-0049/193/2/29. arXiv: 1101.1559 [astro-ph.IM].
- [2] Daniel Foreman-Mackey. «corner.py: Scatterplot matrices in Python». In: *The Journal of Open Source Software* 1.2 (giu. 2016), p. 24. DOI: 10.21105/joss.00024. URL: <https://doi.org/10.21105/joss.00024>.
- [3] Daniel Foreman-Mackey et al. «emcee: The MCMC Hammer». In: 125.925 (mar. 2013), p. 306. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM].
- [4] Miyazaki Iye. «Current Performance and On-Going Improvements of the 8.2 m Subaru Telescope». In: 56.381 (2004).
- [5] E. L. Wright. «A Cosmology Calculator for the World Wide Web». In: 118.850 (dic. 2006), pp. 1711–1715. DOI: 10.1086/510102. arXiv: astro-ph/0609593 [astro-ph].

⁸Intra Cluster Medium, plasma caldo costituito principalmente da idrogeno ionizzato ed elio.