

Pleiadi

Francesca Brivio, Riccardo Loddo, Dalila Signorelli

1 Introduzione

Le Pleiadi sono un ammasso aperto visibile nella costellazione del Toro. Lo scopo di questa esperienza è misurare la distanza tra la Terra e le Pleiadi e quantificare l'estensione spaziale dell'ammasso attraverso l'inferenza Bayesiana. Sono stati utilizzati i dati sulla parallasse, misurati dalla missione *Hipparcos*[1].

2 Modello

2.1 Modello fisico

Il metodo per calcolare la distanza tra un corpo celeste e la Terra è quello della parallasse trigonometrica. Tale fenomeno sfrutta il movimento della Terra attorno al Sole, lungo un'orbita approssimata circolare, per osservare un effetto di spostamento angolare della stella, supposta fissa. Tale metodo è applicabile per stelle relativamente vicine (dell'ordine al massimo di qualche decine di kpc), in quanto stelle molto più lontane hanno un angolo di parallasse non misurabile. Esistono sempre due¹ configurazioni, nel corso di un anno, in cui l'angolo di parallasse è massimo, ovvero quando la stella è perpendicolare alla congiungente Terra-Sole. Tale angolo corrisponde allo spostamento apparente della stella. Noto l'angolo di parallasse massimo, tramite considerazioni trigonometriche, si ricava la relazione che intercorre tra la parallasse massima π e la distanza della stella d :

$$\frac{d}{1000 \text{ pc}} = \frac{1 \text{ mas}}{\pi}. \quad (1)$$

2.2 Modello statistico

Per misurare la distanza e l'estensione dell'ammasso possiamo supporre che la distanza misurata \hat{d}_n per la n -esima stella sia una distribuzione normale con media pari alla distanza vera della stella d_n e varianza ϵ_n^2 , quindi $\hat{d}_n \sim N(d_n, \epsilon_n^2)$. Supponiamo, inoltre, che la distanza vera di una stella sia anch'essa una distribuzione normale con media pari alla distanza del cluster dell'ammasso d_{cl} e con varianza σ_{cl}^2 , quindi $d_n \sim N(d_{cl}, \sigma_{cl}^2)$. Tuttavia non si conosce la distanza vera della stella, quindi occorre esprimere la distanza misurata in funzione della distanza del cluster. Sfruttando il fatto che \hat{d}_n dato d_n è indipendente da d_{cl} e marginalizzando per d_{cl} si ottiene la distribuzione di probabilità delle distanze \hat{d}_n delle stelle data d_{cl} . La distribuzione delle distanze \hat{d}_n delle stelle, quindi, è una distribuzione normale centrata in d_{cl} con varianza data dalla somma delle varianze delle due gaussiane da cui è costituita, $\hat{d}_n \sim N(d_{cl}, \epsilon_n^2 + \sigma_{cl}^2)$. L'analisi verrà effettuata con le parallassi, in quanto la dimensione del cluster in parallasse è minore rispetto alla dimensione in distanze, quindi si commette un errore minore.

Il modello appena descritto dipende da due parametri d_{cl} e σ_{cl} , quindi in termini di parallasse π_{cl} e σ_{cl} . La distribuzione a posteriori di tali parametri, che forniscono una stima della distanza e delle dimensioni

¹Questa condizione riduce notevolmente gli errori di misura che si commettono nel misurare tale spostamento.

dell'ammasso, è stata ottenuta tramite il teorema di Bayes, a partire dai dati di parallasse $\hat{\pi} = \{\hat{\pi}_n\}$ e i relativi errori $\epsilon = \{\epsilon_n\}$:

$$P(\pi_{\text{cl}}, \sigma_{\text{cl}} | \hat{\pi}, \epsilon) = \frac{P(\hat{\pi} | \pi_{\text{cl}}, \sigma_{\text{cl}}, \epsilon) P(\pi_{\text{cl}}, \sigma_{\text{cl}} | \epsilon)}{P(\hat{\pi} | \epsilon)}. \quad (2)$$

Assumendo che le misure di parallasse siano tra loro indipendenti, la distribuzione di verosimiglianza $P(\hat{\pi} | \pi_{\text{cl}}, \sigma_{\text{cl}}, \epsilon)$ assume la forma:

$$P(\hat{\pi} | \pi_{\text{cl}}, \sigma_{\text{cl}}, \epsilon) = \prod_n \left[f \frac{1}{\sqrt{2\pi(\sigma_{\text{cl}}^2 + \epsilon_n^2)}} \exp - \frac{(\hat{\pi}_n - \pi_{\text{cl}})^2}{2(\sigma_{\text{cl}}^2 + \epsilon_n^2)} + (1 - f) D(\hat{\pi}_n) \right], \quad (3)$$

dove si è introdotto un terzo parametro f che rappresenta la probabilità che la stella associata alla misura $\hat{\pi}_n$ appartenga all'ammasso, mentre $D(\hat{\pi}_n)$ è una distribuzione, scelta uniforme, che descrive la presenza di stelle che in realtà non appartengono all'ammasso considerato, esplicitata nella sezione Analisi Dati. La distribuzione a priori $P(\pi_{\text{cl}}, \sigma_{\text{cl}} | \epsilon)$ è stata costruita assumendo che i parametri siano distribuiti in modo uniforme in opportuni intervalli, descritti nella sezione Analisi Dati. Infine l'evidenza $P(\hat{\pi} | \epsilon)$ fornisce la normalizzazione alla distribuzione a posteriori.

3 Dati

L'analisi è stata effettuata utilizzando i dati forniti dalla missione *Hipparcos* relativi al catalogo "*Hipparcos, the New Reduction (van Leeuwen, 2007)*" [2]. I dati sono stati selezionati in un cono di apertura 10° centrato sulle Pleiadi, situate in: $RA = 56.601^\circ$ e $DEC = 24.114^\circ$, ottenendo un numero totale di $N_{\text{tot}} = 817$ stelle. Tuttavia, i dati contengono delle contaminazioni di oggetti che si trovano nelle vicinanze delle Pleiadi, per ovviare a tale problema sono state analizzate le velocità relative ai moti propri $pmRA$ e $pmDEC$ in modo da identificare le Pleiadi come l'ammasso con estensione minore, dovuto al moto proprio coerente delle stelle. Per determinare le stelle appartenenti all'ammasso delle Pleiadi si è utilizzato l'algoritmo di *soft-clustering* contenuto nella libreria *sklearn* [3] di *Python*. In questo modo è possibile selezionare tutte le stelle che hanno una probabilità superiore al 10% di appartenere a tale ammasso, ottenendo un campione di stelle $N = 78$, come mostrato in figura (1), su cui verrà effettuata l'analisi dati.

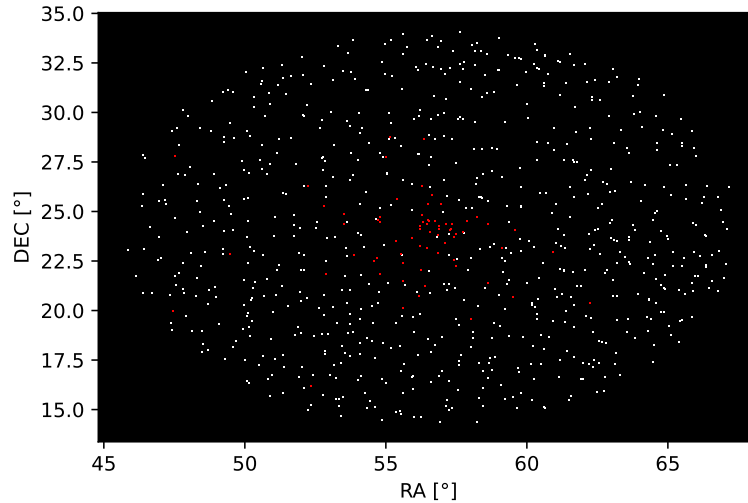


Figura 1: In bianco il numero di stelle totali $N_{\text{tot}} = 817$, in rosso il campione di $N = 78$ stelle.

4 Analisi Dati

Per stimare la distanza e l'estensione dell'ammasso è stato realizzato un campionamento basato sulle catene di Markov, facendo riferimento alla libreria *emcee*[4]. La distribuzione a priori è stata costruita come prodotto delle distribuzioni uniformi sui parametri π_{cl} e σ_{cl} , all'interno dei rispettivi intervalli: (1 mas, 768 mas) e (0 mas, 100 mas). Tale scelta è giustificata dal fatto che, essendo la distanza proporzionale all'inverso della parallasse, il minimo spostamento angolare, corrisponde alla massima distanza, che per la sensibilità di *Hipparcos* è pari a 1 kpc. Mentre, il massimo spostamento angolare è stato scelto considerando che la stella a noi più vicina dista 1.302 pc, dunque in parallasse 768 mas. L'intervallo per σ_{cl} è stato scelto considerando che l'ampiezza del cluster non può essere superiore alla distanza stessa. La distribuzione di verosimiglianza, inizialmente, è stata calcolata considerando che le stelle appartengano tutte all'ammasso, ovvero considerando $f = 1$ nell'equazione (3). Da qui, sfruttando il teorema di Bayes, si sono calcolate le distribuzioni a posteriori. L'algoritmo computazionale è stato eseguito iniziando 7 camminatori e facendo compiere al campionario 5000 passi. Siccome i valori assunti dai parametri sono dipendenti dai valori iniziali dei camminatori, l'algoritmo è stato ripetuto una seconda volta, utilizzando come valori iniziali dei camminatori le posizioni di convergenza degli stessi durante la prima esecuzione e rigettando i primi 100 passi. In modo analogo, l'analisi è stata ripetuta ampliando l'inferenza a tre parametri. Così facendo, la distribuzione di verosimiglianza è stata calcolata utilizzando f come terzo parametro in modo da scartare le ulteriori stelle di campo. A tal fine, la distribuzione uniforme $D(\hat{\pi}_n)$, definita nell'equazione (3), è stata costruita in modo tale da essere normalizzata rispetto all'intervallo scelto per le parallassi π_{cl} :

$$D(\hat{\pi}_n) = \begin{cases} \frac{1}{767} & \text{se } 1 \leq \hat{\pi}_n \leq 768 \\ 0 & \text{altrimenti} \end{cases} \quad (4)$$

I risultati ottenuti da entrambe le analisi sono stati, infine, convertiti da parallassi in distanze tramite l'equazione (1) e tramite la propagazione degli errori si sono calcolate le rispettive estensioni dell'ammasso. I risultati sono riportati in figura 2.

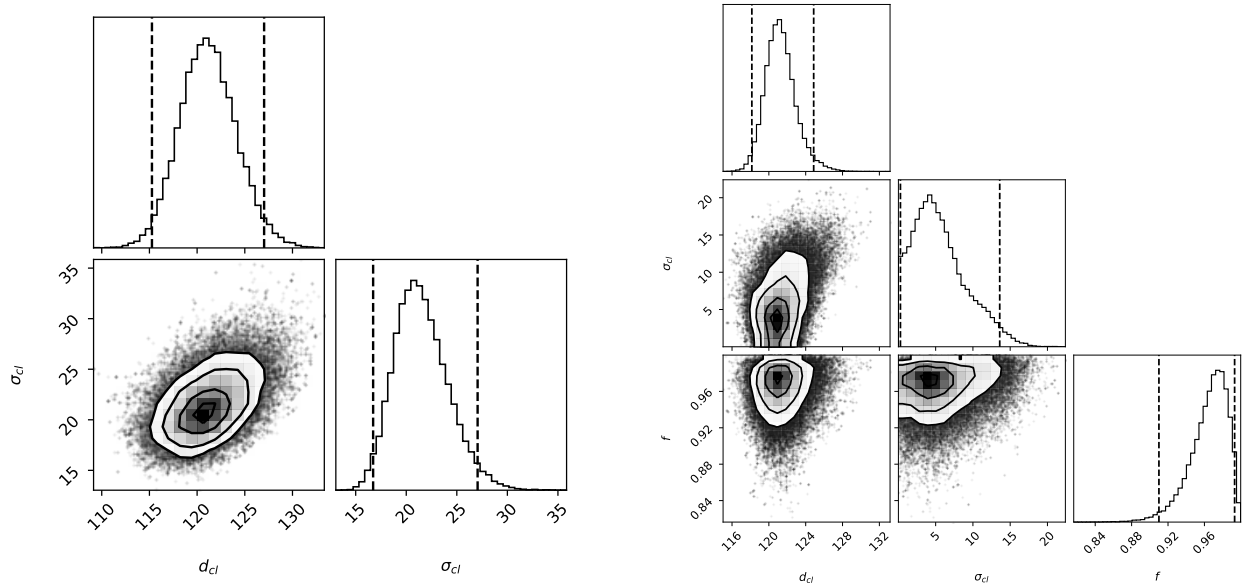


Figura 2: Distribuzione a posteriori dei parametri: a sinistra d_{cl} e σ_{cl} , a destra d_{cl} , σ_{cl} e f . Le linee tratteggiate corrispondono ai rispettivi intervalli di credibilità.

5 Discussione dei risultati

Gli intervalli di credibilità, riportati in tabella (1), mostrano entro dove è possibile determinare i valori più probabili dei parametri con confidenza del 95%. Tali intervalli sono stati calcolati con i rispettivi quantili 0.02 e 0.97. Dai risultati, ottenuti con i due modelli, si nota che le distanze risultano compatibili tra loro, mentre i valori trovati per l'estensione dell'ammasso non risultano compatibili tra loro. Ciò può essere dovuto al fatto che l'analisi a due parametri è stata effettuata considerando anche le possibili stelle di campo, che allargano l'ampiezza del cluster.

Tabella 1: Intervalli di credibilità dei parametri ottenuti con i due modelli.

Numero Parametri	Intervallo di credibilità		
	d_{cl} [pc]	σ_{cl} [pc]	f
2	(115.3, 127.0)	(16.7, 27.1)	-
3	(118.2, 124.9)	(0.27, 13.6)	(0.91, 0.99)

References

- [1] ESA, 1997, *The Hipparcos and Tycho Catalogues*, ESA SP-1200
- [2] <https://www.cosmos.esa.int/web/hipparcos/hipparcos-2>
- [3] *Scikit-learn: Machine Learning in Python*, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011
- [4] Dan Foreman-Mackey et al., 2020/21, arXiv:1202.3665 [astro-ph.IM]