

第十四届国际天文与天体物理奥林匹克竞赛

数据分析试题

哥伦比亚 波哥大(线上) 2021年11月15日

数据分析1: 标度关系

在你开始做这个问题之前, 请阅读单独信封中的一般说明.

螺旋星系是盘状的旋转结构, 其动态状态可以通过所谓的旋转曲线得到相当的把握, 该曲线量化了盘在离中心不同距离的平均旋转速度(见图1, 曲线B). 一个有趣的特征是曲线的平坦区域, 这归因于暗物质的存在. 如果没有暗物质, 在离中心较远的地方, 旋转速度会稳步下降, 如曲线A所描述.

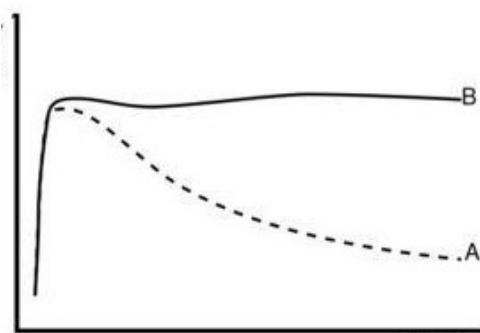


图1: 旋转曲线. 圆周速度(Y轴)与半径(X轴)的关系

在盘状星系中, 已经观察到整个星系的本征光度与渐进的旋转速度(由星系外缘的旋转曲线即 R_{\max} 给出)之间有很强的相关性, 这一结果被称为Tully-Fisher关系. 如果是特定波段的光度, 这个关系也是成立的. 图2中显示了一个星系团中的一些星系. 每个点都是一个星系, 实线是整个样本中K波段的绝对亮度与 $\log_{10}(V_{\max})$ 的最佳拟合线性关系.

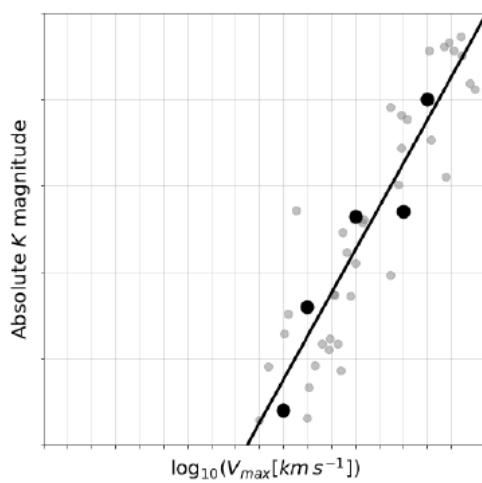


图2: K波段的绝对星等与 $\log_{10}(V_{\max}[\text{km s}^{-1}])$. 几个星系的Tully-Fisher关系. 每个点代表一个星系. 暗点是五个选定的星系, 我们将在1.2部分提供一些数据.

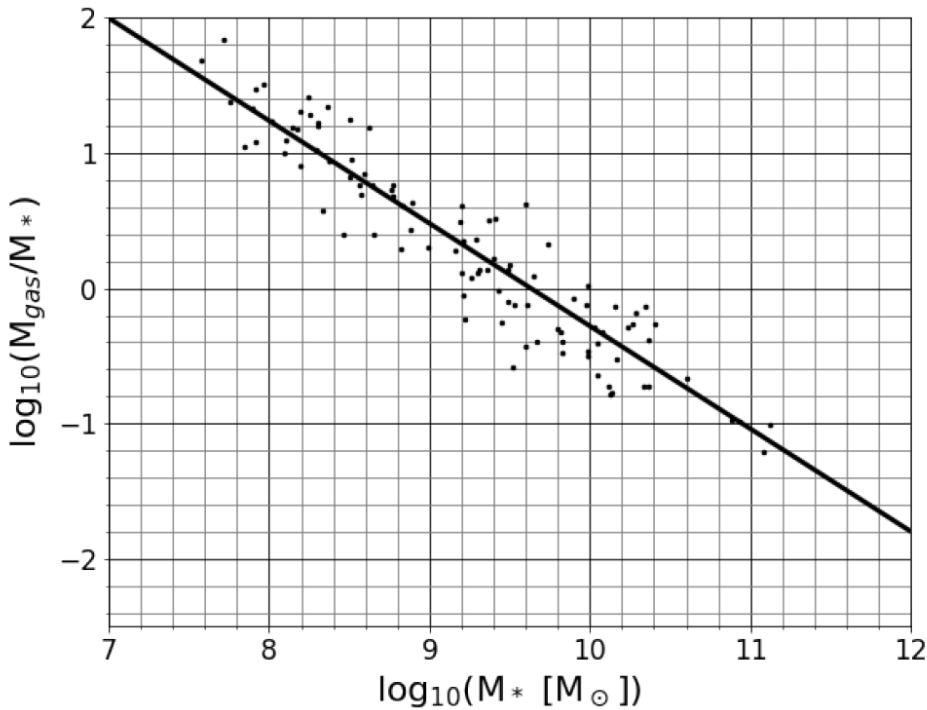


图3: 气体分数与恒星质量的关系.

另一个有趣的趋势显示在图3中: 具有较大恒星质量(M_*)的星盘往往具有较小的气体比例(M_{gas}/M_*).

在下面的问题中, 你将被要求利用刚才介绍的比例关系来提取有关星系的物理信息. 请考虑以下准则.

- 假设 V_{max} 是在所有星系的相同半径(R_{max})处测量的, 位于旋转曲线的平坦部分, 远远超过恒星盘的末端.
- 用 M_{dm} 来表示达到 R_{max} 处的暗物质质量, 用 M_{tot} 表示所有成分(气体、恒星和暗物质)的总和.
- 假设所有星系都有相同的星族¹, 并假设气体成分不与恒星光相互作用.
- 星系团很远. 它的距离比星团的大小大得多.
- 在球形对称的质量分布中, 要推断离中心有 r 距离的粒子的引力作用, 只需考虑到该半径以内的总质量 $M(\leq r)$, 就好像它被放在分布的最中心.

¹术语星族指的是一个星系中存在的恒星的类型, 以及每种类型相对于恒星总数的数量.

第一部分

1.1 根据对图3的分析, 找出下列关系中合适的常数: $M_{\text{gas}} = a \times M_*^b$

$$a = ?$$

$$b = ?$$

1.2 在Tully-Fisher关系图中, 有5个突出的点. 这5个星系的数据在下表中给出. 使用这个数据集, 使用最小二乘法进行线性拟合, 为表下的TF关系找到适当的常数.

注: 在线性拟合中, 将 $\log_{10}(V_{\text{max}})$ 作为 x 变量, 将 K 作为 y 变量.

| V_{max} [km/s] | K [mag] |
|-------------------------|-----------|
| 79.4 | -16.8 |
| 100.1 | -19.2 |
| 158.5 | -21.3 |
| 251.2 | -21.4 |
| 316.2 | -24.0 |

$$K = c \times \log_{10}(V_{\max}) + d$$

$$c = ?$$

$$d = ?$$

第二部分

对于星团中的两个星系, G_1 和 G_2 , 记录的视星等是:

$$k_1 = 19.2 \quad ; \quad k_2 = 25.2$$

利用这些信息和第一部分中标定的关系, 在下列方程中找到正确的指数.

2.1

$$\frac{M_{*1}}{M_{*2}} = 10^e \quad ; \quad e = ?$$

2.2

$$\frac{M_{\text{gas}1}}{M_{\text{gas}2}} = 10^f \quad ; \quad f = ?$$

2.3

$$\frac{M_{\text{tot}1}}{M_{\text{tot}2}} = 10^g \quad ; \quad g = ?$$

第三部分

3.1

| 星系 | 视星等 k | M_{gas} [M_{\odot}] | M_{*} [M_{\odot}] | M_{dm} [M_{\odot}] | M_{tot} [M_{\odot}] |
|-------|---------|----------------------------------|-------------------------|---------------------------------|----------------------------------|
| G_1 | 19.2 | | | | 4.39×10^{11} |

利用以下事实填补表中缺失的数值: 对于星系 G_1 , 在 R_{\max} 之内, 暗物质与重子的质量比为6.82.

第四部分

4.1 考虑到由于CCD的校准误差, 每个视星等的系统误差为 $\sigma_{\text{sys}} = \pm 0.2$. 那么 k_1 必须读作 $k_1 = 19.2 \pm 0.2$, 也就是说, 我们唯一知道的是, k_1 很可能位于[19.0, 19.4]的区间内. k_2 的情况也是如此.

重新计算比例关系 $\frac{M_{*1}}{M_{*2}} = 10^e$ 中的指数(可在2.1中找到), 将 e 表示为通过考虑 k_1 和 k_2 的极端可能变化而估计的区间.

$$e \in [?, ?]$$

4.2 现在我们考虑, 任何关系附近的数据总是有一个自然分布. 例如, 对于给定的 K 星等, TF关系给出的单一值为 $\log_{10}(V_{\max})$, 但更现实的是报告一个可信值的区间, 该区间来自TF关系平均值附近数据的自然扩散. 我们把这称为统计误差, σ_{stat} .

使用问题1.2中的TF关系从 K 中推断出 $\log_{10}(V_{\max})$, 估计统计误差. 为此, 考虑每一点的 $\log_{10}(V_{\max})$ 值与实际测量的 $\log_{10}(V_{\max})$ 之间的差异, 并将 σ_{stat} 作为这些差值的均方根(RMS)的两倍[†].

$$\sigma_{\text{stat}} = ?$$

[†]一组数值的均方根是这些数值平方的算术平均值的平方根.

4.3 重新计算比例关系 $\frac{M_{\text{tot1}}}{M_{\text{tot2}}} = 10^g$ 中的指数, 将 g 表示为通过考虑系统和统计误差引起的极端可能变化而估计的区间.

$$g \in [?, ?]$$

数据分析策 恒星和系外行星

在开始这个问题之前, 请阅读单独信封中的一般说明.

在这个问题中, 我们将探索系外行星的物理特性和它们的宿主恒星之间的联系, 并将利用观测数据尽可能多地发现这些系统. 你可以忽略星际消光.

第一部分

| 行星名 | 恒星名 | T_{eff} (K) | g (m s ⁻²) | m_v (mag) | 视差(毫角秒) |
|---------|-----------|----------------------|--------------------------|-------------|---------|
| Gorgona | HD 209458 | 5980 | 347 | 7.63 | 20.67 |

表1: Gorgona系外行星及其母星HD 209458的观测数据

有效温度(T_{eff})和恒星表面的重力加速度(g)可以通过光谱的形状和吸收线来测量. 目视星等(m_v)和视差则分别通过光度测量和天体测量来测量.

此外, 据观察, 每隔3.52天, 这颗恒星的亮度就会因为它前面的行星的凌星而下降, 就像这条光曲线中所表示的那样:

1.1

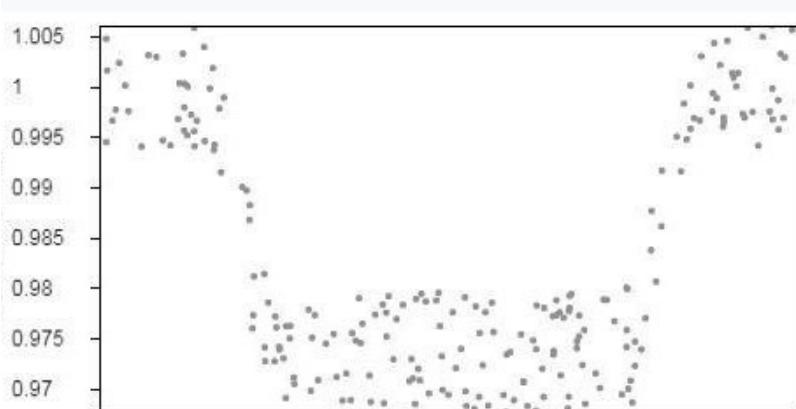


图1: 母星HD 209458的归一化通量(y轴)与时间(x轴)的关系

使用给定的信息来计算HD 209458系统的下列量.

| 恒星光度 | 恒星半径 | 恒星质量 | 行星轨道 平均半径 | 行星半径, 以木星半径为单位 |
|------------------------|------------------------|------------------------|--------------|--------------------|
| L_* [L_\odot] | R_* [R_\odot] | M_* [M_\odot] | a [au] | R_p [R_J] |
| | | | | |

注: 假设所有F型和G型恒星的测光校正是相同的.

第二部分

宜居带被定义为一颗行星表面可能有液态水的区域. 这主要与从宿主恒星接收的辐射量有关, 辐射量必须在一定范围内, 以确保行星表面温度的有利范围.

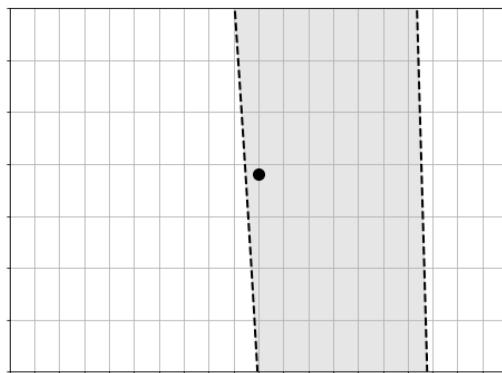
我们将行星接收的有效通量定义为: $S_{\text{eff}} = \frac{L}{a^2}$, 其中 L 为太阳单位的恒星光度, a 为平均轨道半径 (au). 宜居带的最小通量可近似为 $S_{\text{min}} = S_{\text{eff}_{\odot}} + n \cdot T_{\star} + b \cdot T_{\star}^2 + c \cdot T_{\star}^3 + d \cdot T_{\star}^4$. 其中 $T_{\star} = (T_{\text{eff}} - T_{\text{eff}_{\odot}})$, $S_{\text{eff}_{\odot}}$ 是太阳情况下的等效通量, 它与系数 n 、 b 、 c 、 d 一起在下表中给出. 宜居的最大通量, S_{max} , 用同样的公式但不同的常数可算出.

| 常数 | S_{max} | S_{min} |
|--------------------------|---------------------------|---------------------------|
| $S_{\text{eff}_{\odot}}$ | 1.0512 | 0.3438 |
| n | 1.3242×10^{-4} | 5.8942×10^{-5} |
| b | 1.5418×10^{-8} | 1.6558×10^{-9} |
| c | -7.9895×10^{-12} | -3.0045×10^{-12} |
| d | -1.8328×10^{-15} | -5.2983×10^{-16} |

下表给出了7个不同恒星-行星系统的真实数据. 而行星的名称已经被改变, 以纪念哥伦比亚的一些自然保护区.

| 恒星参数 | | 行星参数 | |
|----------------------|-------------|-----------|----------|
| T_{eff} [K] | M_V [mag] | 名字 | a [au] |
| 6180 | 3.68 | Tayrona | 0.04 |
| 5730 | 3.87 | Iguaque | 0.04 |
| 5980 | 4.21 | Gorgona | 0.04 |
| 5480 | 6.04 | Amacayacu | 0.08 |
| 5770 | 3.48 | Malpelo | 0.05 |
| 6130 | 3.07 | Pisba | 0.03 |
| 6140 | 3.85 | Tatamá | 0.06 |

2.1 在下图中, 纵轴代表恒星的有效温度, 横轴代表行星接收的有效通量. 图中标记的点代表地球, 虚线标志着宜居带的界限.



在两个轴上的刻度线位置标上数字. 在同一张图上画出Gorgona和Amacayacu的确切位置, 假设如果它们也位于距离相应恒星1 au的位置.

2.2 现在, 考虑表格中给出的每颗行星的实际轨道半径, 用“YES”或“NO”来表示哪些行星位于宜居带. 在Working Sheet上展示你的定量推理.

| 行星名 | 是否在宜居带内? YES/NO |
|-----------|--------------------|
| Tayrona | |
| Iguaque | |
| Gorgona | |
| Amacayacu | |
| Malpelo | |
| Pisba | |
| Tatamá | |

第三部分

在最后一页，你会发现一份38颗系外行星的名单，目标是找出低质量系外行星(LME)和高质量系外行星(HME)是否倾向于围绕具有不同特征的恒星运行。

3.1 为了得到一个可信的低质量子样本，我们可以应用一种叫做“迭代 σ 排除”的方法。其思路是计算质量的平均值(μ)和标准差(σ)，并从样本中排除那些质量高于 $\mu + \sigma$ 的行星。然后对剩余的子样本再重复同样的步骤两次。我们会说，最后的子样本中的行星是低质量的，而那些在迭代过程中被排除的行星是高质量的。将你在这一过程中发现的数字填入下表。

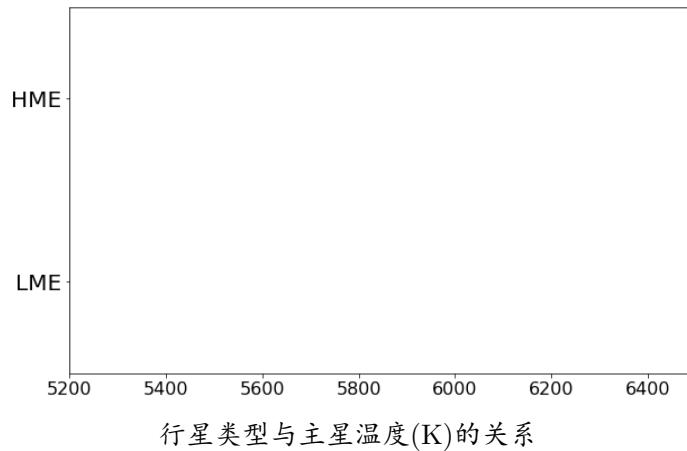
| 低质量样本 | 样本容量 | μ | σ | $\mu + \sigma$ | 排除的行星数 |
|-----------|------|-------|----------|----------------|--------|
| 全部/原始 | 38 | | | | |
| 第一次迭代后子样本 | | | | | |
| 第二次迭代后子样本 | | | | | |
| 最终迭代后子样本 | | — | — | — | — |

3.2 用X轴表示列表中行星的序列号(1, 2, 3,), Y轴表示行星的质量，做一个图。在你在迭代中发现的 $\mu + \sigma$ 阈值处标记3条横线。

3.3 让我们研究一下这两组宿主星的有效温度可能存在的差异，计算一些描述性的统计数据。

| | 最小值 | 第一四分位数 | 中位数 第二四分位数 | 第三四分位数 | 最大值 |
|-----|-----|--------|---------------|--------|-----|
| LME | | | | | |
| HME | | | | | |

3.4 画出总结你刚才计算的数字的箱型图。你是否看到低质量和高质量行星所围绕的恒星的温度有明显的不同？写下“是”或“否”。



| 序号 | 行星名 | 行星质量 [M_J] | 恒星 T_{eff} |
|----|--------------|-------------------|---------------------|
| 1 | KEPLER-37 b | 0.01 | 5520 |
| 2 | KEPLER-21 b | 0.02 | 6256 |
| 3 | HD 97658 b | 0.02 | 5468 |
| 4 | HD 46375 b | 0.23 | 5345 |
| 5 | HD 219134 h | 0.28 | 5209 |
| 6 | HD 88133 b | 0.30 | 5582 |
| 7 | HD33283 b | 0.33 | 5877 |
| 8 | HD 149026 b | 0.36 | 6096 |
| 9 | BD-10 3166 b | 0.46 | 5578 |
| 10 | HD 75289 b | 0.47 | 6196 |
| 11 | HD 217014 b | 0.47 | 5755 |
| 12 | HD 2638 b | 0.48 | 5564 |
| 13 | WASP-13 b | 0.49 | 6025 |
| 14 | WASP-34 b | 0.59 | 5771 |
| 15 | HD 209458 b | 0.69 | 5988 |
| 16 | HAT-P-30 b | 0.71 | 6177 |
| 17 | WASP-76 b | 0.92 | 6133 |
| 18 | WASP-74 b | 0.97 | 5727 |
| 19 | HAT-P-6 b | 1.06 | 6442 |
| 20 | HD189733 b | 1.14 | 5374 |
| 21 | WASP-82 b | 1.24 | 6257 |
| 22 | KELT-7 b | 1.29 | 6460 |
| 23 | HD 149143 b | 1.33 | 6067 |
| 24 | KELT-3 b | 1.42 | 6404 |
| 25 | KELT-2A b | 1.49 | 6164 |
| 26 | HD86081 b | 1.50 | 6015 |
| 27 | HAT-P-7 b | 1.74 | 6270 |
| 28 | HD 118203 b | 2.14 | 5847 |
| 29 | HAT-P-14 b | 2.20 | 6490 |
| 30 | WASP-38 b | 2.71 | 6178 |
| 31 | HD17156 b | 3.20 | 5985 |
| 32 | KELT-6 c | 3.71 | 6176 |
| 33 | HD 75732 d | 3.86 | 5548 |
| 34 | HD 115383 b | 4.00 | 5891 |
| 35 | HD 120136 b | 5.84 | 6210 |
| 36 | WASP-14 b | 7.34 | 6195 |
| 37 | HAT-P-2 b | 8.74 | 6439 |
| 38 | XO-3 b | 11.79 | 6281 |