Observational Astrophysics — Exercise 7: the Corbet diagram of HMXBs in the SMC and the Milky Way

Lecture 10 (correlations, two-sample tests). Data: `smc_hmxb_20231004.fits` (SMC HMXB catalogue, updated from Haberl & Sturm 2016, A&A 586, A81) and `HMXBwebcat_latest.fits` (Galactic HMXB catalogue, Neumann et al. 2023, A&A 677, A134). Read them with `astropy.table.Table.read`.

Physics in brief. In a Be/X-ray binary the neutron star accretes from the decretion disc of the Be star, mostly near periastron. The pulsar spins up or down until it reaches an *equilibrium* where the corotation radius equals the magnetospheric radius,

$$P_{\rm eq} \propto B^{6/7}\,\dot{M}^{-3/7} \quad \text{(for fixed NS mass and radius)},$$

and a wider orbit means a lower time-averaged $\dot{M}$, so $P_{\rm spin}$ should grow with $P_{\rm orb}$ (Corbet 1984, A&A 141, 91; 1986, MNRAS 220, 1047). Supergiant systems behave differently: wind-fed supergiants show long spin periods with no clear trend, Roche-lobe-overflow systems short $P_{\rm spin}$ and short $P_{\rm orb}$. The $P_{\rm spin}$ distribution of Be/XRBs has been reported to be *bimodal*, with peaks near $\sim$10 s and $\sim$200 s (Knigge, Coe & Podsiadlowski 2011, Nature 479, 372).

1 — Data (be explicit about every cut)

Identify the columns with orbital period, spin period and companion type/class. Keep systems with a measured $P_{\rm spin}$ (and, for task 2, $P_{\rm orb}$). Label each system Be, supergiant (SG) or other/unknown. Report the numbers in each class and population. Work in $\log P$.

2 — Correlation: the Corbet diagram

(a) Plot $\log P_{\rm spin}$ vs $\log P_{\rm orb}$ for both galaxies, with different symbols for Be and SG systems.

(b) Compute Pearson r, Spearman ρ_s and Kendall τ for: (i) all systems, (ii) Be systems only — separately for the SMC and the Milky Way. Give p-values by permutation and 68% intervals by bootstrap.

(c) Should SG systems be excluded? Argue from the physics above, then show what including them does to the coefficients.

(d) Fit $\log P_{\rm spin} = a + b \log P_{\rm orb}$ to the Be systems with OLS $(y|x)$, OLS $(x|y)$ and the bisector. Which slope would you quote, and why?

3 — Is the spin-period distribution bimodal? Do the two galaxies differ?

(a) Histogram (try several bin widths) and kernel density of $\log P_{\rm spin}$ for the Be systems of each galaxy.

(b) Fit one and two Gaussians to the unbinned $\log P_{\rm spin}$ (maximum likelihood, e.g. `sklearn.mixture.GaussianMixture`). Compare with BIC; calibrate $\Delta \ln \mathcal{L}$ by simulating from the one-Gaussian fit (why can't you use Wilks' theorem here?). Is the bimodality present in both populations?

(c) Compare the SMC and Galactic $\log P_{\rm spin}$ (and $\log P_{\rm orb}$) distributions with KS, Anderson–Darling and Mann–Whitney; give permutation p-values. Add the Dirichlet Bayes factor from the lecture (averaged over the number of bins). Do the tests agree? Which aspect of the distributions differs?

4 — Systematics (short written answer, ~10 lines)

Which selection effects and hidden variables could produce or hide the correlation and the differences? Consider at least: how pulsations and orbits were found (long $P_{\rm spin}$ needs long exposures, short $P_{\rm spin}$ needs time resolution; $P_{\rm orb}$ from optical photometry or X-ray timing, aliases); homogeneous SMC monitoring vs a heterogeneous Galactic sample (absorption, distance); spin changes during outbursts; metallicity and age of the populations; the B distribution; the boundary between Be and SG classes.

Hand in: code (notebook or script) and a one-page report: the Corbet diagram, a table of correlation coefficients with p and intervals, the bimodality and two-sample results, and the systematics paragraph.