Observational Astrophysics (MSc, NKUA) **Homework 3b**

Is it a Be/X-ray binary? Reasoning with clues

All numbers are fictional, chosen to keep the arithmetic simple.

For every question below, solve the problem in **three ways** and check that they agree:

(i) **logic**: count sources, using a table or a frequency tree;

(ii) **Bayes' theorem**: $p(H \mid D, C) = \dfrac{p(H \mid C)\, p(D \mid H, C)}{p(H \mid C)\, p(D \mid H, C) + p(\bar{H} \mid C)\, p(D \mid \bar{H}, C)}$;

(iii) **odds**: posterior odds = prior odds × likelihood ratio, then convert back with $p = \text{odds}/(1 + \text{odds})$.

A. A first clue: pulsations

You are working with an X-ray catalogue C of the Small Magellanic Cloud. It contains 1000 point sources. From earlier optical and X-ray studies, 100 of them are known to be Be/X-ray binaries (BeXRBs); the rest are mostly background AGN, foreground stars and a few other types of X-ray binaries.

Many BeXRBs host an accreting neutron star whose spin produces X-ray pulsations. With the exposure available for this catalogue, pulsations are detected in 60% of the BeXRBs. Pulsations are rare among the other sources: only 1% of them show a signal that passes the detection threshold (other neutron stars, or occasional false alarms).

A colleague runs a timing analysis on one source of the catalogue and finds pulsations.

Let $H \equiv$ "the source is a BeXRB", $\bar{H} \equiv$ "it is not", and $P \equiv$ "pulsations are detected".

A1. Write down $p(H \mid C)$, $p(P \mid H, C)$ and $p(P \mid \bar{H}, C)$.

A2. What is the probability that the source is a BeXRB, $p(H \mid P, C)$? Use all three methods.

A3. Your colleague says: "60% of BeXRBs pulsate, so there is a 60% chance this is a BeXRB." What is wrong with this argument?

B. A second clue: an early-type counterpart

A different source in the same catalogue has no timing information. Instead, optical follow-up finds an early-type (B) star inside its X-ray error circle. Every BeXRB has a B-type companion, and with the astrometry of this catalogue it falls inside the error circle in 95% of cases. But the SMC is crowded with young, massive stars: for a source that is *not* a BeXRB, there is still a 20% chance that an unrelated B star lies in the error circle by coincidence.

Let $E \equiv$ "an early-type star is found in the error circle".

B1. Write down $p(E \mid H, C)$ and $p(E \mid \bar{H}, C)$.

B2. What is $p(H \mid E, C)$? Use all three methods.

B3. "There is only a 20% chance of a coincidence, so the source is 80% likely to be a BeXRB." Is this right? Explain.

B4. Complete the table, and say which clue is more informative and how you can tell from the table alone:

Clue	$p(\text{clue} \mid H, C)$	$p(\text{clue} \mid \bar{H}, C)$	likelihood ratio	$p(H \mid \text{clue}, C)$
P: pulsations detected				
E: early-type counterpart				

C. Both clues, assumed independent

A third source shows pulsations *and* has a B star in its error circle. Assume that, for a given source type, the two clues are independent:

$$p(E \mid P, H, C) = p(E \mid H, C), \qquad p(E \mid P, \bar{H}, C) = p(E \mid \bar{H}, C).$$

C1. Extend your frequency tree by one level and compute $p(H \mid P, E, C)$ by counting.

C2. Write $p(P, E \mid H, C)$ and $p(P, E \mid \bar{H}, C)$ with the product rule, and compute $p(H \mid P, E, C)$ with Bayes' theorem.

C3. Compute it with odds, applying the clues one at a time. Then apply them in the *opposite* order. Does the order matter? Why?

D. Looking inside the "other" HMXB sources

The 900 non-BeXRBs are not one kind of object. A closer look at the catalogue shows that 10 of them are a different type of high-mass X-ray binaries (e.g. supergiant systems). Like BeXRBs, they contain neutron stars that often pulsate, and their companion is itself an early-type star, but we can only seperate Sg from Be via optical spectroscopy. The remaining 890 are background AGN and foreground stars: they pulsate only through rare false alarms, and they have a B star in the error circle only by chance coincidence.

Non-BeXRB population	number	pulsating	B star in error circle	pulsating *and* B star
Sg HMXBs	10	6	10	6
AGN, foreground stars, ...	890	3 (false alarms)	170 (chance)	1
All non-BeXRBs	900			

D1. Complete the last row. Check that it reproduces $p(P \mid \bar{H}, C) = 0.01$ and $p(E \mid \bar{H}, C) = 0.20$ used in A and B.

D2. Among the *pulsating* non-BeXRBs, what fraction has a B star, $p(E \mid P, \bar{H}, C)$? Compare with $p(E \mid \bar{H}, C)$. Was the independence assumption of part C justified? Why not?

D3. Recompute $p(H \mid P, E, C)$ with all three methods, using the product rule $p(P, E \mid \bar{H}, C) = p(P \mid \bar{H}, C)\, p(E \mid P, \bar{H}, C)$. Compare with your answer in C. Which number in your frequency tree changed?

E. Reflection

In a few sentences reflect on what you learned: CAn we use logic and frequency tries in this type of problems, is it efficient? What does the odds form need instead?

Submission. Hand-written or typed solutions; a short Python script is welcome but not required.