1. Derive the expression for multiple propositions in terms of probabilities of single propositions:

$$p(XYZ|I) = p(X|I) \cdot p(Y|XI) \cdot p(Z|XYI) \quad (4)$$

$P(X,Y,Z|I) = P(X|I)\,P(YZ|X,I) = P(X|I)\,P(Y|XI)\,P(Z|XYI)$

3. As a corollary of Eq. (2), prove Bayes' theorem:

$$p(Y|X,I) = \frac{p(X|Y,I) \cdot p(Y|I)}{p(X|I)} \quad (9)$$

Given 3 statements A, B and I, we know that $p(AB|I) = p(BA|I)$. $\Leftrightarrow$

$\Leftrightarrow P(A|B,I) \cdot P(B,I) = P(B|A,I) \cdot P(A|I) \Leftrightarrow P(A|B,I) = \dfrac{P(B|A,I) \cdot P(A|I)}{P(B|I)}$

2. Manipulating Eqs. (1-2), obtain the *marginalization* rule in the case of a complete discrete set of alternatives:

$$p(X|I) = \sum_k p(XY_k|I) \quad (5)$$

where a *complete set* is a set of events which respect the normalization condition:

$$\sum_k p(Y_k|X,I) = 1 \quad (6)$$

for a given Y_k we have: $P(x,Y_k|I) = P(Y_k|x,I)\,P(x|I)$

$\Rightarrow \sum_k P(x,Y_k|I) = \sum_k P(Y_k|x,I)\,P(x|I) = P(x|I)\underbrace{\sum_k P(Y_k|x,I)}_{=1}$

$\Leftrightarrow \boxed{P(x|I) = \sum_k P(x,Y_k|I)}$

integration over all possible $\vec{\theta}$

4. (a) Use Bayes' theorem to derive the connection between the evidence, the likelihood and the prior:

$$p(D|H,I) = \int d\vec{\theta}\, p(\vec{\theta}|H,I) \cdot p(D|\vec{\theta},H,I) \qquad (12)$$

(b) Re-derive the equation above, but now using the marginalisation rule.

a) $\int p(\vec{\theta}|DHI)\, d\vec{\theta} = 1 \iff \int \frac{P(\vec{\theta}|H,I)\, P(D|\vec{\theta}HI)}{P(D|HI)}\, d\vec{\theta} = 1$

$$\iff \boxed{\int p(\vec{\theta}|H,I)\, P(D|\vec{\theta}H,I)\, d\vec{\theta} = P(D|H,I)}$$

b) Marginalization rule $p(D|H,I) = \int d\vec{\theta}\, P(\vec{\theta},D|H,I)$

we have

$p(\vec{\theta},D|H,I) = p(D|\vec{\theta},H,I)\, p(\vec{\theta}|H,I)$

$$\therefore \boxed{p(D|H,I) = \int d\vec{\theta}\, p(\vec{\theta}|H,I)\, P(D|\vec{\theta},H,I)}$$

5. (a) Derive the posterior probability for multiple datasets (D_1, D_2):

$$p(H|D_1,D_2,I) = p(D_2|H,D_1,I) \cdot p(H|D_1,I) \qquad (13)$$

(b) What does this equation tell us on an inference process which is performed sequentially (i.e. first analysing D_1, then D_2) on multiple datasets, compared to analysing all dataset at the same time?

(c) How does the expression above change when the two datasets are independent?

a) $P(H,D_1,D_2|I) = P(H|D_1,D_2,I)\, P(D_1,D_2|I)$

However this is equivalent to $P(H,D_1,D_2|I) =$

$= P(D_2|HD_1I)\, P(HD_1|I) = P(D_2|HD_1I) \cdot P(H|D_1I) \cdot P(D_1|I)$

$\Rightarrow P(H|D_1,D_2,I) = \frac{P(H,D_1,D_2|I)}{P(D_1,D_2|I)} = \frac{P(D_2|HD_1I)\, P(H|D_1I)\, P(\cancel{D_1}|I)}{P(D_2|D_1I)\, P(\cancel{D_1}|I)}$

$$= \frac{P(D_2 | H D_1 I)\, P(H | D_1 I)}{P(D_2 | D_1 I)}$$

b) Note that P(H|D1,I) is the prior on the first dataset. Thus, we can use the posterior of the so far analyzed datasets, as a prior for the new dataset. If the datasets D1 and D2 are independent, then P(D2|H,D1,I) = P(D2|H,I) and the above expression becomes:

c)

$$P(H | D_1, D_2, I) = \frac{P(D_2 | H, I) \cdot P(H | D_1, I)}{P(D_2 | I)}$$

6. As a corollary to both Eq.s(1,2), prove the expression for the sum ("or"):

$$p(X+Y|I) = p(X|I) + p(Y|I) - p(XY|I) \tag{14}$$

Hint – Think how to relate "X+Y" to products of the X, Y propositions and their negations, which can be then manipulated using the rules of logic.

From a logical perspective, the phrase "X or Y" is equivalent to the phrace
"X and Y or X and not Y or not X and Y". Or expressed mathematically: $X + Y = XY + X\bar{Y} + \bar{X}Y$
Now each of the products of Y and X are mutually exclusive and can thus be manipulated by the product rule:

$$P(X+Y|I) = P(XY|I) + P(\bar{X}Y|I) + P(X\bar{Y}|I) =$$

$$= P(X|Y,I)\,P(Y|I) + P(\bar{X}|Y,I)\,P(Y|I) + P(X|\bar{Y},I)\,P(\bar{Y}|I)$$

$$= P(Y|I)\underbrace{\left[P(X|Y,I) + P(\bar{X}|Y,I)\right]}_{=1} + P(X|\bar{Y},I)\,P(\bar{Y}|I)$$

$$= P(Y|I) + P(X|\bar{Y},I)\,P(\bar{Y}|I) \quad \text{Commutativity}$$

$$= P(Y|I) + P(\bar{Y}|X,I)\,P(X|I)$$

$$= P(Y|I) + \underbrace{\left(1 - P(Y|X,I)\right)}_{P(Y|X,I) + P(\bar{Y}|X,I) = 1} P(X|I)$$

$$= P(Y|I) + P(X|I) - \underbrace{P(Y|X,I)\,P(X|I)}_{P(X,Y|I)}$$

$$= \boxed{P(Y|I) + P(X|I) - P(X,Y|I)}$$