Problem 13.8

Let $D$ and $T$ be random variables for disease and test, respectively.

\[
P(D = \text{true} | T = \text{true}) = \frac{P(T = \text{true} | D = \text{true}) P(D = \text{true})}{P(T = \text{true})}
\]

\[
= \frac{P(T = \text{true} | D = \text{true}) P(D = \text{true})}{P(T = \text{true} | D = \text{true}) P(D = \text{true}) + P(T = \text{true} | D = \text{false}) P(D = \text{false})}
\]

\[
= \frac{0.99 \cdot 0.0001}{0.99 \cdot 0.0001 + 0.01 \cdot 0.9999}
\]

\[
= \frac{1}{102}
\]

Since the disease is rare, the probability you have it is still small.

Problem 14.3

(a)

(i) Incorrect. Consider conditioning on $M_1$ and $M_2$. We have that $N$ is conditionally independent of $F_1$ and $F_2$ given $M_1$ and $M_2$. To see this, note that $N$ is not reachable from $F_1$ or $F_2$ when $M_1$ and $M_2$ are shaded (the Bayes’ ball algorithm). Hence, in this network $P(N|M_1, M_2, F_1, F_2) = P(N|M_1, M_2)$. But this is false. For example, $P(N = 1|M_1 = 1, M_2 = 1, F_1 = \text{true}, F_2 = \text{true}) = 0$ as both measurements have undercounted. But $P(N = 1|M_1 = 1, M_2 = 1)$ is nonzero.

By the chain rule, we can factor any probability distribution $P(A, B, C, D, ..., Z)$ as

\[
P(A, B, C, D, ..., Z) = P(A|B, C, D, ..., Z) P(B|C, D, ..., Z) P(C|D, ..., Z) \cdot ... \cdot P(Z)
\]

(ii) Correct. Consider the ordering of nodes $M_1, M_2, F_1, F_2, N$.

\[
P(M_1, M_2, F_1, F_2, N) = P(M_1|M_2, F_1, F_2, N) P(M_2|F_1, F_2, N) P(F_1|F_2, N) P(F_2|N) P(N)
\]

\[
= P(M_1|F_1, N) P(M_2|F_2, N) P(F_1) P(F_2) P(N)
\]
since $M_1$ depends only on $F_1$ and $N$ (and likewise for $M_2$), and $F_1$, $F_2$ are independent of $N$. This last line is the probability distribution represented by the Bayes’ net.

(iii) Correct. Consider the ordering of nodes $F_2, F_1, N, M_2, M_1$

$$P(F_2, F_1, N, M_2, M_1) = P(F_2|F_1, N, M_2, M_1)P(F_1|N, M_2, M_1)P(N|M_2, M_1)P(M_2|M_1)P(M_1)$$

$$= P(F_2|N, M_2)P(F_1|N, M_1)P(N|M_2, M_1)P(M_2|M_1)P(M_1)$$

This last line is the probability distribution represented by the Bayes’ net.

(b) Network (ii) is the best since it is the sparsest (most efficient) correct network.

(c) Network (ii) is the best since it is the sparsest (most efficient) correct network.

| $N$ | $M_1$ | $P(M_1|N)$ |
|-----|-------|------------|
| 1   | 0     | $e + f - ef$ |
| 1   | 1     | $(1 - 2e)(1 - f)$ |
| 1   | 2     | $e(1 - f)$ |
| 1   | 3     | 0 |
| 1   | 4     | 0 |
| 2   | 0     | $f$ |
| 2   | 1     | $e(1 - f)$ |
| 2   | 2     | $(1 - 2e)(1 - f)$ |
| 2   | 3     | $e(1 - f)$ |
| 2   | 4     | 0 |
| 3   | 0     | $f$ |
| 3   | 1     | 0 |
| 3   | 2     | $e(1 - f)$ |
| 3   | 3     | $(1 - 2e)(1 - f)$ |
| 3   | 4     | $e(1 - f)$ |

Note that the meaning of $e$ is slightly ambiguous in the problem. It could be that there is a probability $e$ of overcounting by one and a probability $e$ of undercounting by one. Or it could be the probability of each is $e/2$. Under this interpretation simply replace $e$ by $e/2$ in the above table.

(d) $M_1 = 1$ and $M_2 = 3$. $N >= 0$ as we cannot have a negative number of stars.

Can $N$ be $<= 1$? No, $M_2$ cannot overcount by 2 or more.

Can $N = 2$? Yes, if both $M_1$ and $M_2$ are off by one (errors occur).
Can \( N = 3? \) No, \( M_1 \) cannot undercount by exactly 2 (only 1 or more than 2).
Can \( N = 4? \) Yes, if \( F_1 = true \), and \( M_2 \) makes an error.
Can \( N = 5? \) No, \( M_2 \) cannot undercount by exactly 2.
Can \( N \) be \( >= 6? \) Yes, if both \( F_1 = F_2 = true \).
So \( N = 2, \) or \( N = 4, \) or \( N >= 6. \)

(e) \( P(N) \), the prior on the number of stars, is the additional information required to compute the most likely number.

**Problem 14.4**

\[
P(N = n|M_1 = 2, M_2 = 2) \propto P(N = n, M_1 = 2, M_2 = 2)
\]

\[
P(N = n, M_1 = 2, M_2 = 2)
= \sum_{f_1, f_2} P(N = n, M_1 = 2, M_2 = 2, F_1 = f_1, F_2 = f_2)
= \sum_{f_1, f_2} P(M_1 = 2|N = n, F_1 = f_1)P(M_2 = 2|N = n, F_2 = f_2)P(F_1 = f_1)P(F_2 = f_2)P(N = n)
= \sum_{f_1} \sum_{f_2} P(M_1 = 2|N = n, F_1 = f_1)P(M_2 = 2|N = n, F_2 = f_2)P(F_1 = f_1)P(F_2 = f_2)P(N = n)
= \sum_{f_1} \sum_{f_2} P(N = n)P(M_1 = 2|N = n, F_1 = f_1)P(M_2 = 2|N = n, F_2 = f_2)P(F_1 = f_1)P(F_2 = f_2)
= P(N = n)\sum_{f_1} P(M_1 = 2|N = n, F_1 = f_1)P(F_1 = f_1)\sum_{f_2} P(M_2 = 2|N = n, F_2 = f_2)P(F_2 = f_2)
= P(N = n)\sum_{f_1} P(M_1 = 2|N = n, F_1 = f_1)P(F_1 = f_1)P(M_2 = 2|N = n)
= P(N = n)P(M_1 = 2|N = n)P(M_2 = 2|N = n)
\]

Using the table from part (c), we have:
\[
P(N = 1, M_1 = 2, M_2 = 2) = [e(1 - f)]^2 \cdot P(N = 1)
P(N = 2, M_1 = 2, M_2 = 2) = [(1 - 2e)(1 - f)]^2 \cdot P(N = 2)
P(N = 3, M_1 = 2, M_2 = 2) = [e(1 - f)]^2 \cdot P(N = 3)
\]

Defining \( Z = P(M_1 = 2, M_2 = 2) = \sum_{n \in \{1, 2, 3\}} P(N = n, M_1 = 2, M_2 = 2) \) we have:

\[
P(N = n|M_1 = 2, M_2 = 2) = P(N = n, M_1 = 2, M_2 = 2)/Z \text{ for } n \in \{1, 2, 3\}.
\]

Note that since the question did not specify \( P(N) \), it is ok if you assumed a uniform prior \((P(N = 1) = P(N = 2) = P(N = 3) = \frac{1}{3})\).