0.1 Trinary hidden random variables

As mentioned in the beginning, the reason for developing a probabilistic model of lymphatic tumor progression in the first place is that modern imaging modalities cannot detect tumor cells directly, but only the (macroscopic) changes they exert on their region of growth, e.g. when they cause lymph nodes to swell. In contrast, a histopathological examination of a resected malignancy or biopsy sample uses various staining techniques in combination with microscopes to detect carcinoma cells directly.

That raises the question whether clinical imaging and pathological examinations can both be modelled as observations of the same hidden random variable. The current resolution of the former is magnitudes away from being able to directly detect cancer cells and yet we consider the chance of detecting a microscopic involvement to be given by the sensitivity of magnetic resonance imaging (MRI), computed tomography (CT), etc., the same one even as the chance for detecting macroscopic metastases. In other words: The sensitivity for detecting lesions far smaller that the voxel resolution of imaging is zero.

To account for this issue, one could categorize lymphatic involvement a little more finely: Instead of treating the true state of a lymph node level (LNL) as a binary random variable representing a health node and a metastaic node respectively, we could cosider micro- and macroscopic involvement separately. The hidden states were then modelled as *trinary* hidden variables (see **??** for comparison):

$$
\begin{aligned}
\text{hidden} \quad & \mathbf{X} = (X_v) \to \{0, \mu, M\}^V \\
\text{observed} \quad & \mathbf{Z} = (Z_v) \to \{0, 1\}^V
\end{aligned}
\tag{1}
$$

where μ and M now respectively represent *microscopic* and *macroscopic* involvement separately.

Diagnoses can still only be binary and in order to decide on a treatment, i.e. whether to irradiate the LNL in question, we also still only care about the distinction *cancerous* – meaning there are malign cells present – vs *healthy*. However, the trinary hidden state represents the underlying reality much more precisely: To an imaging modality the hidden states 0 and μ are healthy states, since it cannot detect microscopic disease, and the respective probabilities for true and false negative observations is governed by the specificity of the modality. To a pathologist however, it is the other way around: The hidden states μ and M appear as truly involved and the probability to correctly identify these states as involved is given by the sensitivity (which we usually assume to be one for pathology).

Now, it seems that with binary observations of an underlying trinary state we cannot expect to correctly infer the true state of the LNL. But imagine that we have data on diagnostic CTs and/or MRIs taken before a neck dissection. After the surgery, the resected LNLs are examined by a pathologist. The outcome is as follows:

$$
\begin{aligned}
Z_v^{\text{MRI}} &= 0 \\
Z_v^{\text{path}} &= 1
\end{aligned}
\tag{2}
$$

From the second observation we can immediately infer that the hidden state must be either μ or M, because it means tumor cells have been found in the LNL. The

MRI diagnosis gives us the additional information that the probability for state μ is $P(X_v = \mu) = s_P$, i.e. the specificity. For state M it is $p(X_v = M) = 1 - s_N$, where s_N is the sensitivity. Obviously, we can infer the likely true hidden state by combining observations of different kinds of modalities: Imaging on the one hand and fine needle aspiration (FNA) and pathology on the other hand.

The inference above might seem unnecessary at first: Could the pathologist not simply distinguish the three hidden states? But the distinction between micro- and macroscopic metastases is actually done by the imaging modality. Any occult disease that cannot be detected by them is considered a microscopic disease. This ability depends on many factors, like presence of necrotic tissue and also characterisitics of the machine used to obtain the scans, that make the distinction fuzzy. Clearly, pathologists cannot routinely think about whether the disease they see would already have been visible on a scan or not and consequently, the distinction between micro and macro is not done in pathology.