Signal Detection Theory

Professor David Heeger

The starting point for signal detection theory is that nearly all reasoning and decision making takes place in the presence of some uncertainty. Signal detection theory provides a precise language and graphic notation for analyzing decision making in the presence of uncertainty. The general approach of signal detection theory has direct application for us in terms of sensory experiments. But it also offers a way to analyze many different kinds of decision problems.

What you should know from this lecture

- Information acquisition
- Criterion
- Internal response and external noise
- Probability of occurrence curves
- Receiver operating characteristic (ROC curve)
- Discriminability index (d')

Information and Criterion

I begin here with medical scenario. Imagine that a radiologist is examining a CT scan, looking for evidence of a tumor. Interpreting CT images is hard and it takes a lot of training. Because the task is so hard, there is always some uncertainty as to what is there or not. Either there is a tumor (signal present) or there is not (signal absent). Either the doctor sees a tumor (they respond "yes") or does not (they respond "no"). There are four possible outcomes: hit (tumor present and doctor says "yes"), miss (tumor present and doctor says "no"), false alarm (tumor absent and doctor says "yes"), and correct rejection (tumor absent and doctor says "no"). Hits and correct rejections are good. False alarms and misses are bad.

There are two main components to the decision-making process: information acquisition and criterion.

Information acquisition: First, there is information in the CT scan. For example, healthy lungs have a characteristic shape. The presence of a tumor might distort that shape. Tumors may have different image characteristics: brighter or darker, different texture, etc. With proper training a doctor learns what kinds of things to look for, so with more practice/training they will be able to acquire more (and more reliable) information. Running another test (e.g., MRI) can also be used to acquire more information. Regardless, acquiring more information is good. The effect of information is to increase the

likelihood of getting either a hit or a correct rejection, while reducing the likelihood of an outcome in the two error boxes.

Criterion: The second component of the decision process is quite different. For, in addition to relying on technology/testing to provide information, the medical profession allows doctors to use their own judgement. Different doctors may feel that the different types of errors are not equal. For example, some doctors may feel that missing an opportunity for early diagnosis may mean the difference between life and death. A false alarm, on the other hand, may result only in a routine biopsy operation. They may choose to err toward "yes" (tumor present) decisions. Other doctors, however, may feel that unnecessary surgeries (even routine ones) are very bad (expensive, stress, etc.). They may chose to be more conservative and say "no" (no turmor) more often. They will miss more tumors, but they will be doing their part to reduce unnecessary surgeries. And they may feel that a tumor, if there really is one, will be picked up at the next check-up. These arguments are *not about information*. Two doctors, with equally good training, looking at the same CT scan, will have the same information. But they may have a different *bias/criterion*.

Internal Response and Internal Noise

Detecting a tumor is hard and there will always be some amount of uncertainty. There are two kinds of noise factors that contribute to the uncertainty: *internal noise* and *external noise*.

External noise: There are many possible sources of external noise. There can be noise factors that are part of the photographic process, a smudge, or a bad spot on the film. Or something in the person's lung that is fine but just looks a bit like a tumor. All of these are to be examples of *external noise*. While the doctor makes every effort possible to reduce the external noise, there is little or nothing that they can do to reduce internal noise.

Internal noise: Internal noise refers to the fact that neural responses are noisy. To make this example really concrete, let's suppose that our doctor has a set of tumor detector neurons and that they monitor the response of one of these neurons to determine the likelihood that there is a tumor in the image (if we could find these neurons then perhaps we could publish and article entitled ``What the radiologist's eye tells the radiologist's brain"). These hypothetical tumor detectors will give noisy and variable responses. After one glance at a scan of a healthy lung, our hypothetical tumor detectors might fire 10 spikes per second. After a different glance at the same scan and under the same conditions, these neurons might fire 40 spikes per second.

Internal response: Now I do not really believe that there are tumor detector neurons in a radiologist's brain. But there is some internal state, reflected by neural activity somewhere in the brain, that determines the doctor's impression about whether or not a tumor is present. This is a fundamental issue; the state of your mind is reflected by neural activity somewhere in your brain. This neural activity might be concentrated in just a few neurons or it might be distributed across a large number of neurons. Since we do not know much about where/when this neural activity is, let's simply refer to it as the doctor's *internal response*.

This internal response is inherently noisy. Even when there is no tumor present (no-signal trials) there will be some internal response (sometimes more,

sometimes less) in the doctor's sensory system.

Probability of Occurrence Curves

Figure 1 shows a graph of two hypothetical internal response curves. The curve on the left is for the *noise-alone* (healthy lung) trials, and the curve on the right is for the *signal-plus-noise* (tumor present) trials. The horizontal axis is labeled *internal response* and the vertical axis is labeled *probability*. The height of each curve represents how often that level of internal response will occur. For example, on noise-alone trials, there will generally be about 10 units of internal response: very little. However, there will be some trials with more (or less) internal response because of the internal noise.

Notice that we never lose the noise. The internal response for the signal-plus-noise case is generally greater but there is still a distribution (a spread) of possible responses. Notice also that the curves overlap, that is, the internal response for a noise-alone trial may exceed the internal response for a signal-plus-noise trial.
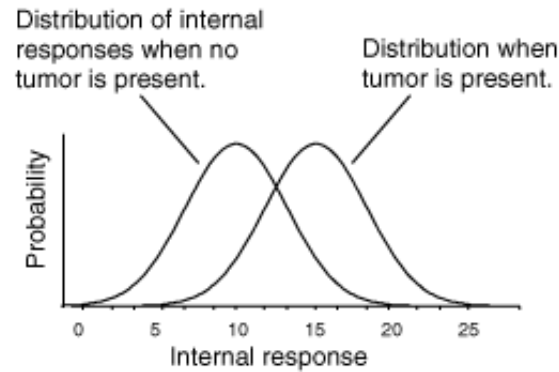

Figure 1: Internal response probability of occurrence curves for noise-alone and for signal-plus-noise trials.

Just to be really concrete, we could mark the horizontal axis in units of firing rate (10, 20, 30,..., etc. spikes per second). This would mean that on a noise-alone (no tumor) trial, it is most likely that the internal response would be 10 spikes per second. It is also rather likely that the internal response would be 5 or 15 spikes per second. But it is very unlikely that the internal response would be 25 spikes per second when no tumor is present. Because I want to remain noncommittal about what and where in the brain the internal response is, I did not label the horizontal axis in terms of firing rates. The internal response is in some unknown, but quantifiable, units.

The role of the criterion: Perhaps the simplest strategy that the doctor can adopt is to pick a *criterion* location along the internal response axis. Whenever the internal response is greater than this criterion they respond "yes". Whenever the internal response is less than this criterion they respond "no".

An example criterion is indicated by the vertical lines in Figure 2. The criterion line divides the graph into four sections that correspond to: hits, misses, false alarms, and correct rejections. On both hits and false alarms, the internal response is greater than the criterion, because the doctor is responding "yes". Hits correspond to signal-plus-noise trials when the internal response is greater than criterion, as indicated in the figure. False alarms correspond to noise-alone trials when the internal response is greater than criterion, as indicated in the figure.
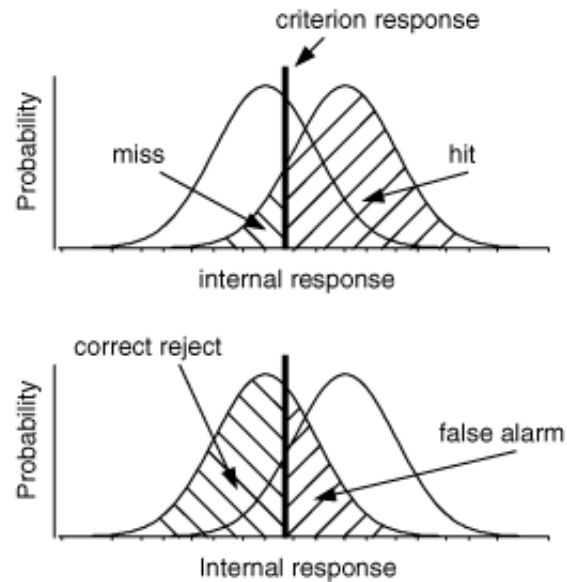

Figure 2: Internal response probability of occurrence curves for noise-alone and signal-plus-noise trials. Since the curves overlap, the internal response for a noise-alone trial may exceed the internal response for a signal-plus-noise trial. Vertical lines correspond to the criterion response.

Suppose that the doctor chooses a low criterion (Figure 3, top), so that they respond "yes" to almost everything. Then they will never miss a tumor when it is present and they will therefore have a very high hit rate. On the other hand, saying "yes" to almost everything will greatly increase the number of false alarms (potentially leading to unnecessary surgeries). Thus, there is a clear cost to increasing the number of hits, and that cost is paid in terms of false alarms. If the doctor chooses a high criterion (Figure 3, bottom) then they respond "no" to almost everything. They will rarely make a false alarm, but they will also miss many real tumors.
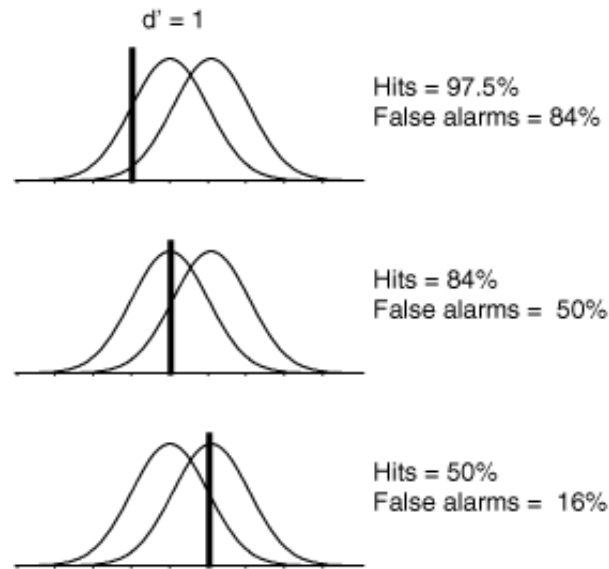
Figure 3: Effect of shifting the criterion

Notice that there is no way that the doctor can set their criterion to achieve only hits and no false alarms. The message that you should be taking home from this is that it is *inevitable* that some mistakes will be made. Because of the noise it is simply a true, undeniable, fact that the internal responses on noise-alone trials may exceed the internal responses on signal-plus-noise trials, in some instances. Thus a doctor cannot always be right. They can adjust the *kind* of errors that they make by manipulating their criterion, the one part of this diagram that is under their control.

The Receiver Operating Characteristic

We can describe the full range of the doctor's options in a single curve, called an *ROC curve,* which stands for receiver-operating characteristic. The ROC curve captures, in a single graph, the various alternatives that are available to the doctor as they move their criterion to higher and lower levels.

ROC curves (Figure 4) are plotted with the false alarm rate on the horizontal axis and the hit rate on the vertical axis. The figure shows several different ROC curves, each corresponding to a different signal strengths. Just pay attention to one of them (the curve labeled d'=1) for the time being. We already know that if the criterion is very high, then both the false alarm rate and the hit rate will be very low, putting you somewhere near the lower left corner of the ROC graph. If the criterion is very low, then both the hit rate and the false alarm rate will be very high, putting you somewhere near the upper right corner of the graph. For an intermediate choice of criterion, the hit rate and false alarm rate will take on intermediate values. The ROC curve characterizes the choices available to the doctor. They may set the criterion anywhere, but any choice that they make will land them with a hit and false alarm rate somewhere on the ROC curve. Notice also that for any reasonable choice of criterion, the hit rate is always larger than the false alarm rate, so the ROC curve is bowed upward.
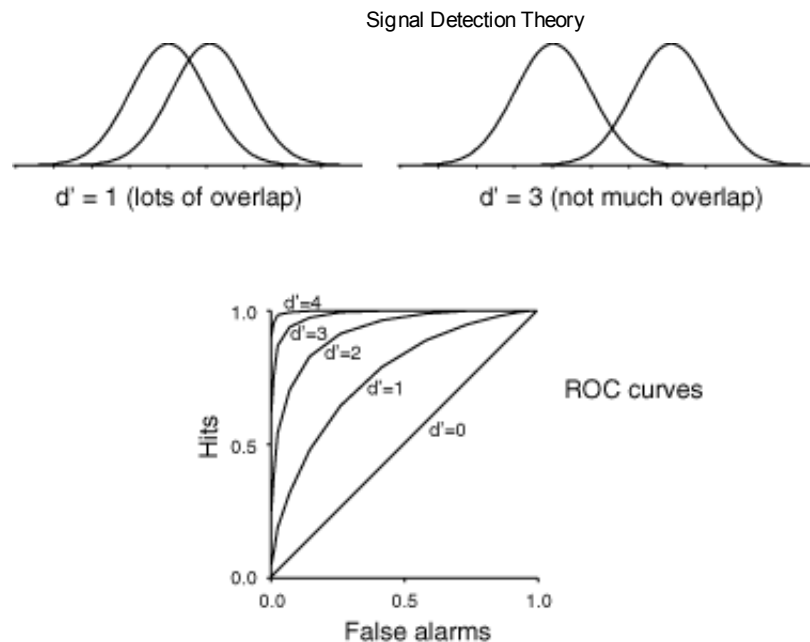
Figure 4: Internal response probability of occurrence curves and ROC curves for different signal strengths. When the signal is stronger there is less overlap in the probability of occurrence curves, and the ROC curve becomes more bowed.

The role of information: Acquiring more information makes the decision easier. Running another test (e.g., MRI) can be used to acquire more information about the presence or absence of a tumor. Unfortunately, the radiologist does not have much control over how much information is available.

In a controlled perception experiment the experimenter has complete control over how much information is provided. Having this control allows for quite a different sort of outcome. If the experimenter chooses to present a stronger stimulus, then the subject's internal response strength will, on the average, be stronger. Pictorially, this will have the effect of shifting the probability of occurrence curve for signal-plus-noise trials to the right, a bit further away from the noise-alone probability of occurrence curve.

Figure 4 shows two sets of probability of occurrence curves. When the signal is stronger there is more separation between the two probability of occurrence curves. When this happens the subject's choices are not so difficult as before. They can pick a criterion to get nearly a perfect hit rate with almost no false alarms. ROC curves for stronger signals bow out further than ROC curves for weaker signals. Ultimately, if the signal is really strong (lots of information), then the ROC curve goes all the way up to the upper left corner (all hits and no false alarms).

Varying the noise: There is another aspect of the probability of occurrence curves that also determines detectability: the amount of noise. Less noise reduces the *spread* of the curves. For example, consider the two probability of occurrence curves in Figure 5. The separation between the peaks is the same but the second set of curves are much skinnier. Clearly, the signal is much more discriminable when there is less spread (less noise) in the

probability of occurrence curves. So the subject would have an easier time setting their criterion in order to be right nearly all the time.
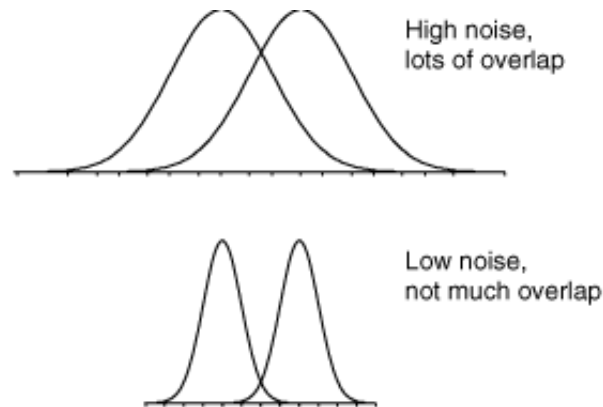

Figure 5: Internal response probability of occurrence curves for two different noise levels. When the noise is greater, the curves are wider (more spread) and there is more overlap.

In reality, we have no control over the amount of internal noise. But it is important to realize that decreasing the noise has the same effect as increasing the signal strength. Both reduce the overlap between the probability of occurrence curves.

Discriminability index (*d′*): Thus, the discriminability of a signal depends both on the separation and the spread of the noise-alone and signal-plus-noise curves. Discriminability is made easier either by increasing the separation (stronger signal) or by decreasing the spread (less noise). In either case, there is less overlap between the probability of occurrence curves. To write down a complete description of how discriminable the signal is from no-signal, we want a formula that captures both the separation and the spread. The most widely used measure is called d-prime (*d′*), and its formula is simply:

$$d′ = \text{separation} / \text{spread}$$

This number, *d′*, is an estimate of the strength of the signal. Its primary virtue, and the reason that it is so widely used, is that its value does not depend upon the criterion the subject is adopting, but instead it is a true measure of the internal response.

Estimating *d′*: To recap... Increasing the stimulus strength separates the two (noise-alone versus signal-plus-noise) probability of occurrence curves. This has the effect of increasing the hit and correct rejection rates. Shifting to a high criterion leads to fewer false alarms, fewer hits, and fewer surgical procedures. Shifting to a low criterion leads to more hits (lots of worthwhile surgeries), but many false alarms (unnecessary surgeries) as well. The discriminability index, *d′*, is a measure of the strength of the internal response that is independent of the criterion.

But how do we measure *d′*? The trick is that we have to measure both the hit rate and the false alarm rate, then we can read-off *d′* from an ROC

curve. Figure 4 shows a family of ROC curves. Each of these curves corresponds to a different d-prime value; $d'=0$, $d'=1$, etc. As the signal strength increases, the internal response increases, the ROC curve bows out more, and d' increases.

So let's say that we do a detection experiment; we ask our doctor to detect tumors in 1000 CT scans. Some of these patients truly had tumors and some of them didn't. We only use patients who have already had surgery (biopsies) so we know which of them truly had tumors. We count up the number of hits and false alarms. And that drops us somewhere on this plot, on one of the ROC curves. Then we simply read off the d' value corresponding to that ROC curve. Notice that we need to know both the hit rate *and* the false alarm rate to get the discriminability index, d'.

Medical Malpractice Example: A study of doctors' performance was performed in Boston. 10,000 cases were analyzed by a special commission. The commission decided which were handled negligently and which well. They found that 100 were handled very badly and there is good cause for a malpractice suit. Of these 100, only 20 cases were pursued. What should we conclude?

Ralph Nader and others concluded that doctors are not being sued enough. But this conclusion was based on only partial information (hits and misses). I did not tell you what happened in the other 9900 cases. How many law suits were there in those cases? What if there were many (e.g., 9000 out of 9900) false alarms? The AMA concluded that doctors are being sued too much.

For a more advanced treatment of signal detection theory, download my Signal Detection Theory handout (75 KB, pdf).

David Heeger