Bayesian Model Details

Demetri Pananos

August 27, 2020

For a single diagnostic test, denote the sensitivity and specificity as δ and γ accordingly. In a population where the true prevalence is π, the estimated prevalence, p, obtained from the test is

$$p = P(T^+) = P(T^+|D^+)P(D^+) + P(T^+|D^-)|P(D^-)$$
$$= \delta\pi + (1-\gamma)(1-\pi)$$

Here T^+ is the event that the test is positive, and D^+ is the event that the patient is disease positive. When a patient is given two diagnostic tests, there are four possible outcomes:

1. Both tests are positive,

2. The first is negative, the second test is positive,

3. The first test is positive, the second test is negative, and

4. Both tests are negative

The results can typically be organized into a 2x2 table

		First Test	
		Positive	Negative
Second Test	Positive	y_1	y_2
	Negative	y_3	y_4

With some effort, the cell probabilities for this table can also be calculated. Let T_i denote the outcome of the ith test. Under the assumption that the results of the tests are independent, the cell probability for two positive tests is:

$$P(T_1^+ \cap T_2^+) = P(T_1^+ \cap T_2^+|D^+)P(D^+) + P(T_1^+ \cap T_2^+|D-)P(D^-) \tag{1}$$
$$= P(T_1^+|D^+)P(T_2^+|D+)P(D^+) + P(T_1^+|D^-)P(T_2^+|D^-)P(D^-) \tag{2}$$
$$= \delta_1\delta_2\pi + (1-\gamma_1)(1-\gamma_2)(1-\pi) \tag{3}$$

Sensitivity and specificity are probabilities. The expression above is merely a statement that the probability both tests positive is the weighted average of the probability of positive tests, weighted by the population prevalence for disease positive and disease negative cases. Through similar approaches, the remaining three cell probabilites (rowwise, left to right) are:

$$P(T_1^- \cap T_2^+) = (1 - \delta_1)\delta_2\pi + \gamma_1(1 - \gamma_2)(1 - \pi) \tag{4}$$

$$P(T_1^+ \cap T_2^-) = \delta_1(1 - \delta_2)\pi + (1 - \gamma_1)\gamma_2(1 - \pi) \tag{5}$$

$$P(T_1^- \cap T_2^-) = (1 - \delta_1)(1 - \delta_2)\pi + \gamma_1\gamma_2(1 - \pi) \tag{6}$$

With these probabilities in hand, we can describe our model's data generating process.

$$\delta_1 \sim \text{Trunc-Normal}(\mu_{\delta_1}, \sigma_{\delta_1})$$
$$\delta_2 \sim \text{Trunc-Normal}(\mu_{\delta_2}, \sigma_{\delta_2})$$

$$\gamma_1 \sim \text{Trunc-Normal}(\mu_{\gamma_1}, \sigma_{\gamma_1})$$
$$\gamma_2 \sim \text{Trunc-Normal}(\mu_{\gamma_2}, \sigma_{\gamma_2})$$

$$\mathbf{y}|\delta_1, \delta_2, \gamma_1, \gamma_2 \sim \text{Multinomial}(\boldsymbol{\pi}; N)$$

Here, Trunc-Normal is a truncated normal distribution. We truncate our priors at 0 and 1 since the sensitivities and specificities are probabilities. $\mathbf{y}$ is a vector in which the elements are cell counts from the two by two table rowwise from left to right. $\boldsymbol{\pi}$ is a vector of the probabilities of equations $(3) - (6)$.