NOSE analysis

Scott Coggeshall

6/12/2020

There are two *unobserved* 2x2 tables containing the true classifications of the standard of care (SoC) swab and the Alterntive (Alt) swab.

		Diseased	
		0	1
SoC	0	p_{00}	p_{01}
	1	p_{10}	p_{11}

		Diseased	
		0	1
Alt	0	q_{00}	q_{01}
	1	q_{10}	q_{11}

The *observed* data consists of the 2x2 cross-classification of the SoC swab and the Alt swab

		SoC	
		0	1
Alt	0	r_{00}	r_{01}
	1	r_{10}	r_{11}

We now express these observed quantities in terms of the underlying disease prevalence and the swabs' diagnostic characteristics. Doing so will allow us to better understand how our assumptions about things like varying disease prevalence affect the observed data.

$$
\begin{aligned}
r_{00} &= P(Alt = 0, SoC = 0) \\
&= P(Alt = 0, SoC = 0 \mid Disease = 1)P(Disease = 1) + P(Alt = 0, SoC = 0 \mid Disease = 0)P(Disease = 0) \\
&= P(Alt = 0 \mid SoC = 0, Disease = 1)P(SoC = 0 \mid Disease = 1) + P(Alt = 0 \mid SoC = 0, Disease = 0)P(SoC = 0 \mid Disea \\
r_{11} &= P(Alt = 1, SoC = 1) \\
&= P(Alt = 1, SoC = 1 \mid Disease = 1)P(Disease = 1) + P(Alt = 1, SoC = 1 \mid Disease = 0)P(Disease = 0) \\
&= P(Alt = 1 \mid SoC = 1, Disease = 1)P(SoC = 1 \mid Disease = 1)P(Disease = 1) + P(Alt = 1 \mid SoC = 1, Disease = 0)P(
\end{aligned}
$$

The first table can be parameterized in terms of the quantities

$$
\begin{aligned}
\pi_d &= P(Diseased) \\
p_1 &= P(SoC = 1) \\
FP &= P(SoC = 1 \mid Diseased = 0)
\end{aligned}
$$

since this fully defines the probabilities

$$
\begin{aligned}
p_{00} &= P(SoC = 0, Diseased = 0) \\
&= P(SoC = 0 \mid Diseased = 0)P(Diseased = 0) \\
&= (1 - P(SoC = 1 \mid Diseased = 0)(1 - P(Diseased = 1))) \\
&= (1 - FP)(1 - \pi_d) \\
p_{10} &= FP \\
p_{11} &= P(SoC = 1, Diseased = 1) \\
&= P(SoC = 1 \mid Diseased = 1)P(Diseased = 1) \\
&= p_1 - FP(1 - \pi_d))
\end{aligned}
$$