

Bayesian analysis using Gibbs sampler



2 POPULATIONS **2 INDEPENDENT TESTS**

2 populations 2 independent tests



2 populations 2 independent tests



- The problem is very similar to that of 1 population 2 independent tests
- Main differences:
 - The two populations are assumed to be characterized by two different values of prevalence of infection
 - Usually, the two tests are assumed to have the same performance in both populations (=same sensitivity and same specificity)

2 populations 2 independent tests



- The last assumption might not always be true:
 - Concerning **sensitivity**:
 - ✦ A test may have different sensitivities according to the time elapsed from the infection of the animals
 - ✦ e.g., PCR becomes positive very soon after the infection, but usually the positivity stays for a few months only; conversely, serological tests are slower to become positive, but the positivity may remain for several years
 - ✦ This may be a relevant feature in small populations, such as the animals of individual herds
 - Concerning **specificity**:
 - ✦ Depending on the frequency of cross-reacting agents in the population of interest, specificity may vary from a population to another

2 populations 2 independent tests



- Now we will consider the case in which the two tests have the same performance in both populations
- The case in which the two tests have different Se and Sp according with the population is coincident with the case of 1 population 2 tests
- The case in which either the Se or the Sp in one or both tests vary with the population of interest is a straightforward derivation of the case we are now describing

Observed and latent data



Observed data for two tests in the absence of a gold standard



Population 1		Test 1		
		Positive	Negative	Totals
Test 2	Positive	T_{1PP}	T_{1NP}	$(T_{1PP} + T_{1NP})$
	Negative	T_{1PN}	T_{1NN}	$(T_{1PN} + T_{1NN})$
	Totals	$(T_{1PP} + T_{1PN})$	$(T_{1NP} + T_{1NN})$	N_1

Observed data for two tests in the absence of a gold standard



Population 2		Test 1		
		Positive	Negative	Totals
Test 2	Positive	$T_{2_{PP}}$	$T_{2_{NP}}$	$(T_{2_{PP}} + T_{2_{NP}})$
	Negative	$T_{2_{PN}}$	$T_{2_{NN}}$	$(T_{2_{PN}} + T_{2_{NN}})$
	Totals	$(T_{2_{PP}} + T_{2_{PN}})$	$(T_{2_{NP}} + T_{2_{NN}})$	N_2

Latent data for two tests in the absence of a gold standard



Population 1 Infected		Test 1		
		Positive	Negative	Totals
Test 2	Positive	D_{1PP}	D_{1NP}	$(D_{1PP} + D_{1NP})$
	Negative	D_{1PN}	D_{1NN}	$(D_{1PN} + D_{1NN})$
	Totals	$(D_{1PP} + D_{1PN})$	$(D_{1NP} + D_{1NN})$	$D_{1PP} + D_{1NP} + D_{1PN} + D_{1NN}$

Population 1 Non-Infected		Test 1		
		Positive	Negative	Totals
Test 2	Positive	$T_{1PP} - D_{1PP}$	$T_{1NP} - D_{1NP}$	$(T_{1PP} + T_{1NP} - D_{1PP} - D_{1NP})$
	Negative	$T_{1PN} - D_{1PN}$	$T_{1NN} - D_{1NN}$	$(T_{1PN} + T_{1NN} - D_{1PN} - D_{1NN})$
	Totals	$(T_{1PP} + T_{1PN} - D_{1PP} - D_{1PN})$	$(T_{1NP} + T_{1NN} - D_{1NP} - D_{1NN})$	$N_1 - D_{1PP} - D_{1PN} - D_{1NP} - D_{1NN}$

Latent data for two tests in the absence of a gold standard



Population 2 Infected		Test 1		
		Positive	Negative	Totals
Test 2	Positive	D_{2PP}	D_{2NP}	$(D_{2PP} + D_{2NP})$
	Negative	D_{2PN}	D_{2NN}	$(D_{2PN} + D_{2NN})$
	Totals	$(D_{2PP} + D_{2PN})$	$(D_{2NP} + D_{2NN})$	$D_{2PP} + D_{2NP} + D_{2PN} + D_{2NN}$

Population 2 Non-Infected		Test 1		
		Positive	Negative	Totals
Test 2	Positive	$T_{2PP} - D_{2PP}$	$T_{2NP} - D_{2NP}$	$(T_{2PP} + T_{2NP} - D_{2PP} - D_{2NP})$
	Negative	$T_{2PN} - D_{2PN}$	$T_{2NN} - D_{2NN}$	$(T_{2PN} + T_{2NN} - D_{2PN} - D_{2NN})$
	Totals	$(T_{2PP} + T_{2PN} - D_{2PP} - D_{2PN})$	$(T_{2NP} + T_{2NN} - D_{2NP} - D_{2NN})$	$N_2 - D_{2PP} - D_{2PN} - D_{2NP} - D_{2NN}$

Likelihood contribution for all latent and observed data



Population 1

Truth = Infected		Test 1	
		Positive	Negative
Test 2	Positive	$D_{1_{PP}} = \pi_1 * Se_1 * Se_2$	$D_{1_{NP}} = \pi_1 * (1 - Se_1) * Se_2$
	Negative	$D_{1_{PN}} = \pi_1 * Se_1 * (1 - Se_2)$	$D_{1_{NN}} = \pi_1 * (1 - Se_1) * (1 - Se_2)$

Truth = Non-Infected		Test 1	
		Positive	Negative
Test 2	Positive	$T_{1_{PP}} - D_{1_{PP}} = (1 - \pi_1) * (1 - Sp_1) * (1 - Sp_2)$	$T_{1_{NP}} - D_{1_{NP}} = (1 - \pi_1) * Sp_1 * (1 - Sp_2)$
	Negative	$T_{1_{PN}} - D_{1_{PN}} = (1 - \pi_1) * (1 - Sp_1) * Sp_2$	$T_{1_{NN}} - D_{1_{NN}} = (1 - \pi_1) * Sp_1 * Sp_2$

Likelihood contribution for all latent and observed data



Population 2

Truth = Infected		Test 1	
		Positive	Negative
Test 2	Positive	$D_{2_{PP}} = \pi_2 * Se_1 * Se_2$	$D_{2_{NP}} = \pi_2 * (1 - Se_1) * Se_2$
	Negative	$D_{2_{PN}} = \pi_2 * Se_1 * (1 - Se_2)$	$D_{2_{NN}} = \pi_2 * (1 - Se_1) * (1 - Se_2)$

Truth = Non-Infected		Test 1	
		Positive	Negative
Test 2	Positive	$T_{2_{PP}} - D_{2_{PP}} = (1 - \pi_2) * (1 - Sp_1) * (1 - Sp_2)$	$T_{2_{NP}} - D_{2_{NP}} = (1 - \pi_2) * Sp_1 * (1 - Sp_2)$
	Negative	$T_{2_{PN}} - D_{2_{PN}} = (1 - \pi_2) * (1 - Sp_1) * Sp_2$	$T_{2_{NN}} - D_{2_{NN}} = (1 - \pi_2) * Sp_1 * Sp_2$

The Markov chain – Step 1



The step 1 of Markov chain



- For each of the two populations (i)
- We can use the contribution to likelihood of each latent class to extract the D_{iPP} , D_{iPN} , D_{iNP} and D_{iNN} values from the proper binomial distributions:

[the formulas are the same as for 1 population 2 tests, just considering the values of the variables relevant for the population under study]

$$D_{iPP} = \text{Binomial} \left[T_{iPP}, \frac{\pi_i * Se_1 * Se_2}{\pi_i * Se_1 * Se_2 + (1 - \pi_i) * (1 - Sp_1) * (1 - Sp_2)} \right]$$

$$D_{iNP} = \text{Binomial} \left[T_{iNP}, \frac{\pi_i * (1 - Se_1) * Se_2}{\pi_i * (1 - Se_1) * Se_2 + (1 - \pi_i) * Sp_1 * (1 - Sp_2)} \right]$$

The step 1 of Markov chain



- For each of the two populations (i)
- We can use the contribution to likelihood of each latent class to extract the D_{iPP} , D_{iPN} , D_{iNP} and D_{iNN} values from the proper binomial distributions: :

[the formulas are the same as for 1 population 2 tests, just considering the values of the variables relevant for the population under study]

$$D_{iPN} = \text{Binomial} \left[T_{iPN}, \frac{\pi_i * Se_1 * (1 - Se_2)}{\pi_i * Se_1 * (1 - Se_2) + (1 - \pi_i) * (1 - Sp_1) * Sp_2} \right]$$

$$D_{iNN} = \text{Binomial} \left[T_{iNN}, \frac{\pi_i * (1 - Se_1) * (1 - Se_2)}{\pi_i * (1 - Se_1) * (1 - Se_2) + (1 - \pi_i) * Sp_1 * Sp_2} \right]$$

The Markov chain – Step 2



The step 2 of Markov chain



- When the latent values for the infected and non-infected animals are known, they can be used to estimate the values of π_i , Se1, Se2, Sp1, Sp2
- The procedure is similar to what we have already seen for the case of 1 population and 2 tests

The step 2 of Markov chain



- The posterior probability distributions of π_i , Se_1 , Se_2 , Sp_1 , Sp_2 are Beta distributions:

$$\pi_i = \text{Beta} \left[(D_{iPP} + D_{iPN} + D_{iNP} + D_{iNN}) + \alpha_{\text{prior}}, N_i - (D_{iPP} + D_{iPN} + D_{iNP} + D_{iNN}) + \beta_{\text{prior}} \right]$$

$$Se_1 = \text{Beta} \left[\sum_i (D_{iPP} + D_{iPN}) + \alpha_{\text{prior}}, \sum_i (D_{iNP} + D_{iNN}) + \beta_{\text{prior}} \right]$$

$$Se_2 = \text{Beta} \left[\sum_i (D_{iPP} + D_{iNP}) + \alpha_{\text{prior}}, \sum_i (D_{iPN} + D_{iNN}) + \beta_{\text{prior}} \right]$$

The step 2 of Markov chain



- The posterior probability distributions of π_i , Se1, Se2, Sp1, Sp2 are Beta distributions:

$$Sp_1 = Beta \left\{ \sum_i [(T_{iNP} - D_{iNP}) + (T_{iNN} - D_{iNN})] + \alpha_{prior}, \sum_i [(T_{iPP} - D_{iPP}) + (T_{iPN} - D_{iPN})] + \beta_{prior} \right\}$$

$$Sp_2 = Beta \left\{ \sum_i [(T_{iPN} - D_{iPN}) + (T_{iNN} - D_{iNN})] + \alpha_{prior}, \sum_i [(T_{iPP} - D_{iPP}) + (T_{iNP} - D_{iNP})] + \beta_{prior} \right\}$$

The step 2 of Markov chain



- The process is repeated feeding the results in the first set of equations
- And the values of π_i , Se1, Se2, Sp1 and Sp2 are stored and their frequency distributions will be used to approximate the posterior probability distribution of the parameters



**And now we are
going to consider
the case of 2
populations and
2 conditionally
dependent tests**