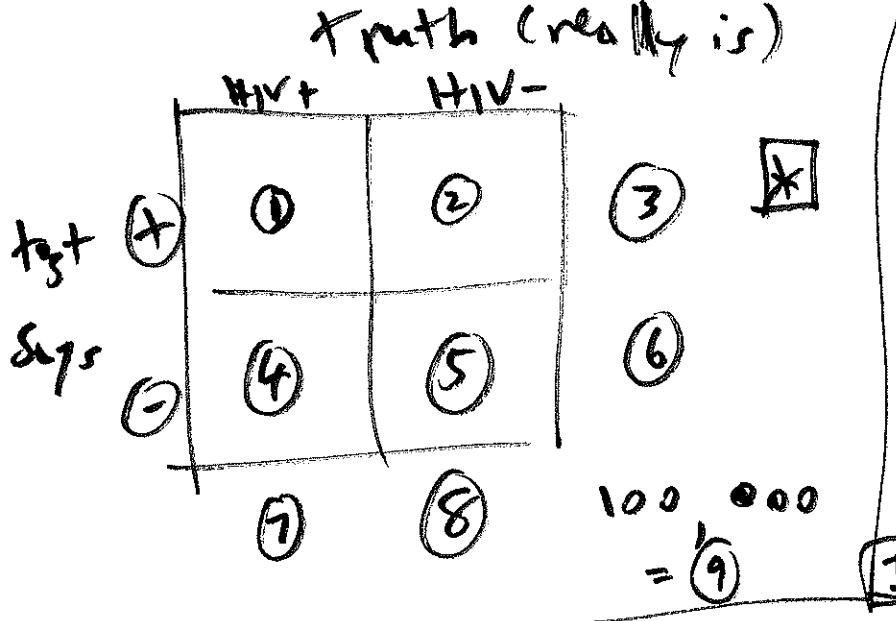
$P(Y_1 = 1 \mid B)$ Bayesians sometimes refer to this as the annoying normalizing constant because it's often the hardest ingredient in Bayes's Theorem to pin down. I will now show you 3 different ways to estimate this probability, all of them instructive: (8 Jan 19)

Method 1: 2×2 contingency table

Imagine randomly sampling 100,000 people from the population P = {all people similar to Bob in all relevant ways} (i.e., people with the same B) and running all the blood samples through "Determine HIV"; cross-tabulate (truth) against (what blood test says), by which I mean what we expect to happen:

blood



what we know ⑬ so far:

I $P(\text{really is HIV+} \mid B) = .01$

II $P(\text{test says HIV+} \mid \text{really is HIV+}) = .999$

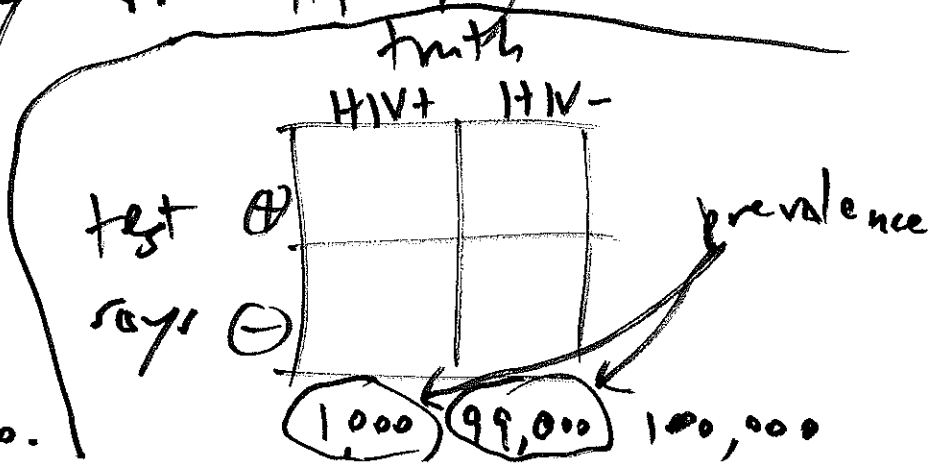
Fact I is called the prevalence of HIV in Bob's population P; let's use it first.

People call * a 2 by 2 contingency table; the counts (integers) in the body of the table (①, ②, ④, ⑤) record how many blood samples have the indicated row attribute (what the test says) and the indicated column attribute (the truth) (e.g., ④ = # blood samples for which test says ⊖ and truth is HIV+). The counts

at the margins of the table (③, ⑥, ⑤, ⑭, ⑧) record how many blood samples have the indicated row (or column) attribute, adding up over the column (or row) attribute, so that ① + ② = ③, ④ + ⑤ = ⑥, ① + ④ = ⑤, ② + ⑤ = ⑧ (e.g. ⑥ = # blood samples for which test says ⊖).

Q1: How can we use the prevalence [1] to begin filling in the table?

A1: 1% of 100,000 is 1,000, so ⑦ = 1,000 (because the first column represents all the blood samples that really are HIV+) and therefore ⑧ has to be 100,000 − 1,000 = 99,000.

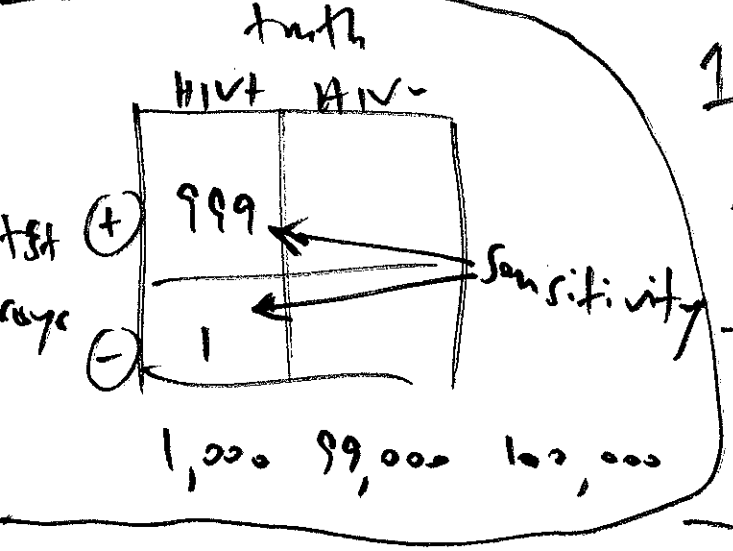


Q2: How can we bring the sensitivity to bear on the table! **A2:** well,

sensitivity = P(test says HIV⊕ | really is HIV+) = .999 ;

Q3 where are we in the table when we know that a blood sample really is HIV+? **A3**

In the first column. So, in 1,000 truly HIV+ samples, the test gets it right $(0.999) \cdot 1000 = 999$ times, & only makes 1 mistake, leading to this table. **Q4** How can we fill in the rest of the table? **A4** we need new information on how good the test is:



The missing analogue to sensitivity is called the specificity P(test says HIV⊖ | really is HIV⊖) of the test.

It turns out that "Determine HIV"'s specificity is almost as good as its sensitivity:

$P(\text{test says HIV}\ominus \mid \text{really is HIV}\ominus) = 0.994$.

Now we can complete the table, by using the specificity on the second column:

$(0.994)\cdot(99{,}000) = 98{,}406$ truly HIV$\ominus$ samples are correctly marked as HIV$\ominus$ by the test, leaving $(99{,}000 - 98{,}406) \to 594$ mistakes; and the row totals are now available & complete the table.

	truth: HIV+	truth: HIV−	
test ⊕	999	594	1,593
test ⊖	1	98,406	98,407
	1,000	99,000	100,000

specificity

We can now compute $P(\theta = 1 \mid y_1 = 1, B)$, and the answer is surprising: [given that the test says HIV⊕ ($y_1 = 1$)] puts us in the first row of the table

so $P(\theta = 1 \mid y_1 = 1, B)$ disappointingly low $= \frac{999}{1{,}593} \doteq 0.627$

(terrific) sensitivity = 0.999^2, (terrific) specificity = 0.994^2

A related & also relevant probability:

$P(\theta=0 \mid y_1=1, B)$ = false positive rate

$= \frac{594}{1593} \approx 0.373 = 1 - P(\theta=1 \mid y_1=1, B)$

with sensitivity & specificity values so close to 1, a false positive rate of 37% is disappointingly high; why has this happened? Q5

A5 Let's look at $P(\theta=1 \mid y_1=1, B) \doteq 63\%$

Q6 When is a fraction small? A6 When its numerator is small, or its denominator is large, or both. $63\% = \frac{999}{1,593}$ (9) Given the low prevalence in Bob's population P, 999 is almost as large as it possibly could be; its biggest possible value is only $999+1 = 1,000$

(b) The denominator, 1,593 = 999 + 594, got large in 2 ways: the terrific sensitivity yielded 999, and the prevalence was so low that the second column total was 99,000; even with a specificity of 0.994, (0.006)(99,000) = 594 is big.

	truth HIV+	HIV−	
test +	999	594	1,593
says −	1	98,406	98,407
	1,000	99,000	100,000

Note that this test's false negative rate is $P(\theta = 1 \mid y_1 = 0, B) = \frac{1}{98,407} \doteq 0.\text{oooo}1$, a fantastically low value. Q1 Evidently the developers of this blood test were much more worried about false negatives than false positives; is this sensible? A1 Answering this involves two additional questions: Q2 The test will be used for what purpose? Q3 What are the

real-world consequences of each type of error?

A8 In CS1 the test will be used by you (the doctor) to advise your patient Bob on his HIV status (other uses would include screening blood donated to a blood bank). **Note:** Something interesting just happened: before Q5, we were doing *science* (the acquisition of knowledge for its own sake); starting with Q5, we're thinking about *decision theory* (using scientific knowledge (behavioral) to make a *choice*: what should you (as Bob's doctor) do if "Determine HIV" comes back positive?)

A9 False positive = (test says ⊕ but really ⊖)

This is bad: if your treatment plan for Bob is based only on the "Determine HIV" result, you will now begin treating Bob for HIV when in fact this treatment is unnecessary, & Bob will think he has a serious disease which in fact he doesn't.

False negative = (test says ⊖ but really ⊕)

This is also bad, but for a different reason: You will tell Bob he's HIV negative when he's actually HIV positive; he will go untreated, & he may unintentionally infect other people.

We can't make any additional progress without making a value judgment: which type of bad is worse, or are they equally bad? Ans Statisticians & economists have developed the concept of a utility function

to make value judgments like this one precise: $u_{Bob}(a, \theta^* | B)$ = the numerical value
↑ utility function
Bob places on action (a) if the unknown (θ) has the value θ^*. This requires specifying $(a|B)_{Bob}$ ← the action space = {all feasible actions judged by Bob to be relevant (worthy of consideration)}. Suppose (for example) that Bob chooses to put his treatment plan completely in your hands; then $(a_{Bob}|B) = (a_{You}|B)$. Suppose further (for example) that You choose $(a_{You}|B) = \{a_1\}$ = {treat Bob for HIV if "determine HIV" returns a positive result & prescribe no treatment for Bob if test result is ⊖}

$U_{Bob}(a, \theta | B)$ + truth

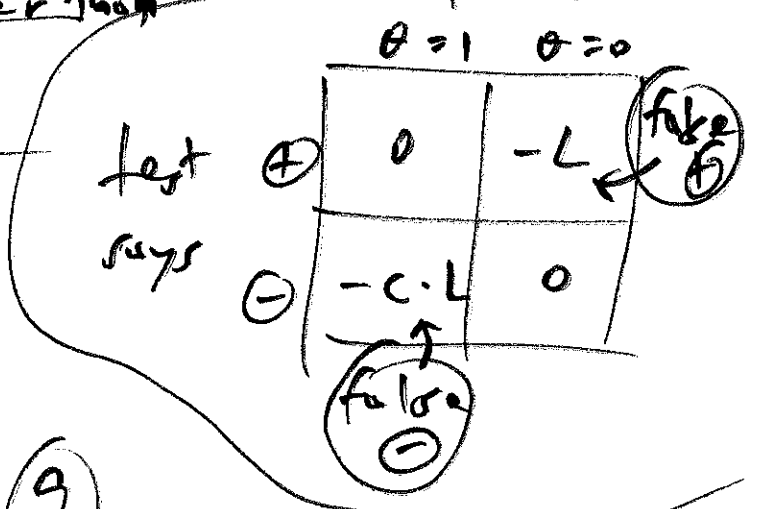
	$\theta=1$	$\theta=0$
test ⊕	u_{11}	u_{12}
says ⊖	u_{21}	u_{22}

Now, to approach this formally, you would need in this table to elicit the u_{ij} values from Bob (since you're acting as his (proxy) agent, so to speak). A simple set of reasonable u_{ij} values can be obtained by (without loss of generality) (a) adopting the convention that large utility values are preferred over small values (better → / utility ↑) and (b) regarding utility as negative loss: $U_{Bob}(a, \theta^* | B) = -L_{Bob}(a, \theta^* | B)$, where $L_{Bob}(a, \theta^* | B)$ is the loss Bob incurs if action a is chosen & the unknown θ takes the value θ^*. From this point of view the diagonal entries in the table above are arguably 0, yielding the new table (more subtle choices are possible)

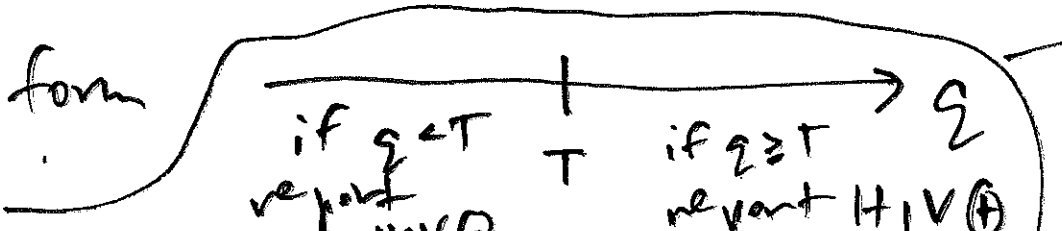
$U_{Bob}(a, \theta | B)$ + truth

	$\theta=1$	$\theta=0$
test ⊕	0	$-l_{12}$
says ⊖	$-l_{21}$	0

$\ell_{12} > 0$ is the loss incurred by Bob if a false positive occurs, and $\ell_{21} > 0$ is Bob's loss under a false negative. It helps to set $\ell_{12} = L$, $\ell_{21} = c \cdot L$ for some $c > 0$; then Bob only needs to think about c. Almost everyone would agree that false negatives are (a lot) worse than false positives here, so that $c \gg 1$. (is a lot bigger than) Fact:



all tests that try to detect HIV attempt to estimate a quantity q (e.g., concentration of antibodies against the virus) & use a decision rule of the form:



T is a tuning constant that

trades off false ⊕ against false ⊖ errors: as T ↑ it gets harder to report HIV⊕, which lowers the false ⊕ rate but increases the false ⊖ rate, and vice versa as T ↓. 45 You can see that qualitatively the designers of "determine HIV" have their utility function aligned with Bob's: the test is terrified of false negatives, & inevitably as a result it makes a lot of false positive mistakes (i.e., the designers aren't stupid).

(continued)

		truth: HIV+	HIV−	
test says	⊕	999	594	1,593
	⊖	1	98,406	98,407
		1,000	99,000	100,000

Method 2 (creative laziness; Bayes's Theorem in odds form) [CS1]

$$P(\theta=1 \mid y_1=1, B) = \frac{P(\theta=1 \mid B)\, P(y_1=1 \mid \theta=1, B)}{P(y_1=1 \mid B)}$$

The denominator in the right-hand side is annoying; how shall we cope with this? [Q10] [Advice]

If solving a problem directly is difficult, try to be creatively lazy, by solving it indirectly. [A10] Application of this advice:

The opposite possibility to $P(\theta=1 \mid y_1=1, B)$ is $P(\theta=0 \mid y_1=1, B)$, and crucially this has the same denominator; take the ratio of these two probabilities and the annoying denominator will cancel:

$$P(\theta=1 \mid y_1=1, B) = \frac{P(\theta=1 \mid B)\, P(y_1=1 \mid \theta=1, B)}{P(y_1=1 \mid B)}$$

$$P(\theta=0 \mid y_1=1, B) = \frac{P(\theta=0 \mid B)\, P(y_1=1 \mid \theta=0, B)}{P(y_1=1 \mid B)}$$

Divide these two equations to obtain (likelihood ratio)

$$\left[\frac{P(\theta=1 \mid y_1=1, B)}{P(\theta=0 \mid y_1=1, B)}\right] = \left[\frac{P(\theta=1 \mid B)}{P(\theta=0 \mid B)}\right] \cdot \left[\frac{P(y_1=1 \mid \theta=1, B)}{P(y_1=1 \mid \theta=0, B)}\right] \quad (*)$$

Definition The conditional odds ratio in favor of A given B is $O_{A|B} = \dfrac{P(A \mid B)}{P(\text{not } A \mid B)}$

With this definition, (*) becomes

$$\left(\begin{array}{c}\text{posterior odds} \\ \text{in favor of} \\ \theta=1 \text{ given} \\ y_1=1 \text{ \& } B\end{array}\right) = \left(\begin{array}{c}\text{prior odds} \\ \text{in favor} \\ \text{of } \theta=1 \\ \text{given } B\end{array}\right) \cdot \left(\begin{array}{c}\text{Bayes} \\ \text{factor} \\ \text{in favor} \\ \text{of } \theta=1 \\ \text{given } y_1=1 \text{ \& } B\end{array}\right)$$

(data odds)

(Bayes's Theorem in odds Form)

Note that we know (via the prevalence, sensitivity & specificity) all of the probabilities on the right-hand side of (*):

$$\left(\begin{array}{c}\text{prior odds in favor} \\ \text{of } \theta=1 \text{ given } B\end{array}\right) = \frac{.01}{.99} \approx \begin{array}{c}99 \text{ to } 1 \\ \text{odds against} \\ \theta=1\end{array}$$

Def. The conditional odds ratio against A given B is $O_{\text{not } A|B} = \frac{P(\text{not } A|B)}{P(A|B)} = \frac{1}{O_{A|B}}$

The sensitivity gives us the numerator of the Bayes factor: $P(y_1=1|\theta=1, \cancel{\beta}) = 0.999$

And the denominator is just 1 minus the specificity: $P(y_1=1|\theta=0, \cancel{\beta}) = 1 - P(y_1=0|\theta=0)$

$= 1 - 0.994 = 0.006.$ So the Bayes factor is

$$\frac{P(y_1=1|\theta=1)}{P(y_1=1|\theta=0)} = \frac{0.999}{0.006} = \frac{333}{2} = 166.5 \text{ to } 1$$

in favor of $\theta=1$

The prior information pushes away from ($\theta=1$) with a 99 to 1 strength on the odds scale, but the data information pushes toward ($\theta=1$) with a 166.5 to 1 strength on the same scale; Bayes's Theorem tells

us that the optimal way to combine these information sources is multiplicatively

the odds scale:

$$\frac{P(\theta=1 \mid y_1=1, B)}{P(\theta=0 \mid y_1=1, B)} = \left[\frac{P(\theta=1 \mid B)}{P(\theta=0 \mid B)}\right] \cdot \left[\frac{P(y_1=1 \mid \theta=1, B)}{P(y_1=1 \mid \theta=0, B)}\right]$$

$$= \frac{1}{99} \cdot \frac{333}{2} = \frac{37}{22}$$

so the posterior ratio of odds in favor of Bob being HIV+, given the blood test results and Bob's background information, is $\frac{37}{22}$.

Q: How are odds ratios and probabilities related?

A: $O_{A|B} = \frac{P(A \mid B)}{P(\text{not } A \mid B)} = \frac{P_{A|B}}{1 - P_{A|B}}$

solve this for $P_{A|B}$ in terms of $O_{A|B}$:

$$P_{A|B} = \frac{O_{A|B}}{1 + O_{A|B}}$$

Thus the posterior probability that Bob is HIV+, given the test results & Bob's background information, is $P(\theta=1 \mid y_1=1, B) = \frac{\frac{37}{22}}{1+\frac{37}{22}} = \frac{37}{59} = \frac{999}{1593} = 0.627$