

Example 3

You are told: (1) 24.8% of adults have HIV

$$\Rightarrow \text{prevalence} = \Pr(\text{HIV}+) = 0.248, \Pr(\text{HIV}-) = 1 - 0.248 = 0.752$$

(2) Of 10,000 HIV+ people, expect 9980 to test positive

$$\Rightarrow \text{sensitivity} = \Pr(T+ | \text{HIV}+) = \frac{9980}{10,000} = 0.998$$

(3) Of 10,000 HIV- people, expect 1 to test positive and 9999 to test negative

$$\Rightarrow \text{specificity} = \Pr(T- | \text{HIV}-) = \frac{9999}{10,000} = 0.9999$$
$$\Pr(T+ | \text{HIV}-) = 1 - \Pr(T- | \text{HIV}-) = 0.0001$$

Asked for positive predictive value = $\Pr(\text{HIV}+ | T+)$

Method 1: formula - use Bayes rule

$$\Pr(\text{HIV}+ | T+) = \frac{\Pr(T+ | \text{HIV}+) \Pr(\text{HIV}+)}{\Pr(T+)}$$

$$= \frac{\Pr(T+ | \text{HIV}+) \Pr(\text{HIV}+)}{\underbrace{\Pr(T+ | \text{HIV}+) \Pr(\text{HIV}+)}_{\Pr(T+ \cap \text{HIV}+)} + \underbrace{\Pr(T+ | \text{HIV}-) \Pr(\text{HIV}-)}_{\Pr(T+ \cap \text{HIV}-)}}$$

$$= \frac{0.998(0.248)}{0.998(0.248) + 0.0001(0.752)}$$

$$= 0.9996963$$

Example 3, Method 2 (Table)

Steps are numbered in the order I used.

	Test		
	+	-	Total
HIV+	④ 247,504	⑤ 496	② 248,000
HIV-	⑦ 75.2	⑥ 751,924.8	③ 752,000
Total	⑧ 247,579.2	⑨ 752,420.8	① 1,000,000

Pick any #. I chose a million.

- ① Pick a hypothetical sample size. I picked 1,000,000.
- ② 24.8% of people have HIV (prevalence), so we have $1,000,000 \times 0.248 = 248,000$ people with HIV
- ③ Rest of people are HIV- : $1,000,000 - 248,000 = 752,000$
- ④ Use sensitivity to figure out how many HIV+ people will be detected by the test: $248,000 \times 0.998 = 247,504$
- ⑤ Rest with HIV go undetected: $248,000 - 247,504 = 496$ 😞
- ⑥ Use specificity to figure out how many HIV- people are told by test they are disease-free: $752,000 \times 0.9999 = 751,924.8$
- ⑦ Rest with ~~no~~ no HIV are false positives: $752,000 - 751,924.8 = 75.2$ 😞
- ⑧ Calculate total who test positive: $247,504 + 75.2 = 247,579.2$
- ⑨ Calculate total who test negative: $751,924.8 + 496 = 752,420.8$
- ⑩ Be sure column totals sum to overall total: $247,579.2 + 752,420.8 = 1,000,000$ (if not you made a mistake!)

$$Pr(HIV+ | T+) = \frac{\# \text{ who test positive AND have HIV}}{\# \text{ who test positive}} = \frac{247,504}{247,579.2} = \underline{\underline{0.996963!}}$$

Example 4

You are told: (1) Prevalence is $0.043 = \Pr(D+)$

(2) Sensitivity $0.83 = \Pr(T+|D+)$

(3) Specificity $0.86 = \Pr(T-|D-)$

You are asked to give $\Pr(\text{High grade lesions} | \text{Test } +)$

Method 1: Bayes rule

$$\Pr(D+|T+) = \frac{\Pr(T+|D+) \Pr(D+)}{\Pr(T+)}$$

You can test positive 2 ways:
(a) and have disease
(b) and be disease-free

$$= \frac{\Pr(T+|D+) \Pr(D+)}{\Pr(T+|D+) \Pr(D+) + \Pr(T+|D-) \Pr(D-)}$$

$$= \frac{0.83(0.043)}{0.83(0.043) + [1 - \Pr(T-|D-)] [1 - \Pr(D+)]}$$

$$= \frac{0.83(0.043)}{0.83(0.043) + 0.14(0.957)} = \frac{0.03589}{0.03589 + 0.13398} = \frac{0.03589}{0.16987} \approx 0.21$$

Method 2: Table

		Test		
		T+	T-	Total
High-grade lesions	D+	④ 35.69	⑤ 7.31	② 43
	D-	⑦ 133.98	⑥ 823.02	③ 957
	Total	⑧ 169.67	⑨ 830.33	⑩ 1000

- ① Pick hypothetical sample size, say 1000
 - ② Prevalence 0.043 means $1000(0.043) = 43$ are expected to be D+
 - ③ Rest are D- so $1000 - 43 = 957$
 - ④ Use sensitivity to figure out how many of the 43 diseased should test positive: $43 \times 0.83 = 35.69$
 - ⑤ Rest, test negative: $43 - 35.69 = 7.31$
of diseased
 - ⑥ Use specificity to determine how many disease-free people should test negative: $957 \times 0.86 = 823.02$
 - ⑦ Rest of disease-free will test positive: $957 - 823.02 = 133.98$
 - ⑧ $35.69 + 133.98 = 169.67$ will test positive in all
 - ⑨ $1000 - 169.67$ will test negative
"830.33
 - ⑩ Double check $7.31 + 823.02 = 830.33$
- } slightly different order from Example 3 - either order is fine!

$$Pr(D+ | T+) = \frac{\# \text{ who have disease AND } T+}{\# \text{ who } T+} = \frac{35.69}{169.67} = \boxed{0.21}$$

Example 5

You are told: (1) Sensitivity is $0.906 = \Pr(T+|D+)$

(2) Specificity is $0.999 = \Pr(T-|D-)$

(3) Prevalence of pregnancy among test users is 0.10
 $= \Pr(D+)$

You are asked to find the positive predictive value of the test,
 $\Pr(D+|T+)$.

Method 1: Bayes's Rule

$$\begin{aligned}\Pr(D+|T+) &= \frac{\Pr(T+|D+)\Pr(D+)}{\Pr(T+)} \\ &= \frac{\Pr(T+|D+)\Pr(D+)}{\Pr(T+|D+)\Pr(D+) + \Pr(T+|D-)\Pr(D-)}\end{aligned}$$

Noting $\Pr(D-) = 1 - \Pr(D+) = 1 - 0.10 = 0.90$

and $\Pr(T+|D-) = 1 - \Pr(T-|D-) = 1 - 0.999 = 0.001$

we have

$$\begin{aligned}\Pr(D+|T+) &= \frac{0.906(0.10)}{0.906(0.10) + 0.001(0.90)} \\ &= \boxed{0.99.}\end{aligned}$$

Example 5: Method 2 (Table)

	T+	T-	Total
Pregnant D+	④ 9060	⑤ 940	② 10,000
Not pregnant D-	⑦ 90	⑥ 89,910	⑨ 90,000
Total	⑧ 9150	⑩ 90,850	① 100,000

① Pick hypothetical sample size - say 100,000

② Use prevalence to get # who are pregnant:
 $100,000 \times 0.10 = 10,000$

③ So $100,000 - 10,000 = 90,000$ are not pregnant

④ Use sensitivity to estimate # of pregnant ♀ who test +
 $10,000 \times 0.906 = 9060$

⑤ Rest of pregnant ♀ are false negatives: $10,000 - 9060 = 940$

⑥ Use specificity to get # of non-pregnant ♀ who test negative:
 $90,000 \times 0.999 = 89,910$

⑦ Rest of non-pregnant ♀ are false positives: $90,000 - 89,910 = 90$

⑧ $9060 + 90 = 9150$ test positive

⑨ $940 + 89,910 = 90,850$ test negative

⑩ Verify $9150 + 90,850 = 100,000$ ✓

$$Pr(D+ | T+) = \frac{\# \text{ who test + and are pregnant}}{\# \text{ who test +}} = \frac{9060}{9150} = \boxed{0.99}$$

Q: WAIT! Example 1 and Example 2 in class were more confusing!

A: Yes, you're right. In Example 1, you have to calculate prevalence, sensitivity, and ~~2~~ specificity first on your own. That makes it a harder question.

—
People were confused by the way prevalence, sensitivity, + specificity were given. I think it is safest always to calculate these #'s first.