Uses of probability in Epidemiology

- Frequency probability models used in epidemiology and other applications
  - Sensitivity, Specificity, Predictive value
  - Relative risk
  - Interpretation of other rates

- Sensitivity = $Pr(T^+|D^+)$
- Specificity = $Pr(T^-|D^-)$

	Disease (D+)	No disease (D-)	Total
Test +	a	b	a+b
Test -	c	d	c+d
	a+c	b+d	N=a+b+c+d

- Sensitivity = $Pr(T+|D+) = \frac{a}{a+c}$
- Specificity = $Pr(T-|D-) = \frac{d}{b+d}$

Example: Screening for Prostate Cancer

	Ca	No Ca	
PSA +	57	1,712	1,769
PSA -	21	13,126	13,147
	78	14,838	14,916

- Sensitivity $= \frac{57}{78} = 0.73$
- Specificity $= \frac{13126}{14838} = 0.88$

Predictive value

- Condition on test results
- Useful for physician planning, etc.
  - Of the patients who test positive, what percentage actually have the disease
- Measures the practical utility of implementing a test in a large setting.

- Predictive value of a positive test: PV+ = Pr(D+|T+)
- Predictive value of a negative test: PV- = Pr(D-|T-)

Prostate cancer

- PV+ $= \frac{57}{1769} = 0.032$
- PV- $= \frac{13126}{13147} = 0.998$

For rare diseases, we often evaluate a test on a sample of cases and and equal size sample of non-cases (controls).

	Disease	No disease	
Test +	a	b	a+b
Test -	c	d	c+d
	n	n	N=2n

- Sensitivity = $\frac{a}{n}$
- Specificity = $\frac{b}{n}$

But, cannot estimate PV+ or PV- because Pr(D+) artificially set to 0.5

Use Bayes's theorem

- If disease prevalence is known, PV+ can be computed from sensitivity, specificity and disease prevalence.

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

HIV Example

- The sensitivity and specificity of the HIV screening test are very high.
- Should the general public be screened for HIV?
- What is the predictive value of the HIV test?

9/1/02 13

	HIV +	HIV -	Total
Test +	99	4	103
Test -	1	96	97
	100	100	200

9/1/02 14

- Sensitivity = 0.99
- Specificity = 0.96
- Disease prevalence is about 2%

9/1/02 15

Predictive value by Bayes's Theorem

$$PV+ = \Pr(D+ \mid T+)$$

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

$$= \frac{0.99*0.02}{0.99*0.02 + (1-0.96)*(1-0.02)}$$

$$= \frac{0.0198}{0.0198 + 0.0392} = 0.34$$

9/1/02 16