

(see syllabus)

(see syllabus)

AMS 131  
18 April

Case Study

credit card screening

2 kinds of info:

(truth)

B = (card really is bad)

(system says)

G = (card really is good)

(+) = (system says bad)

1% = (prevalence)  
= P(B)

(-) = (system says good)



97% = P(- | G) (specificity)

98% = P(+ | B) (sensitivity)

2x2 contingency table

Cross-tabulation of (truth) vs. (system) transactions

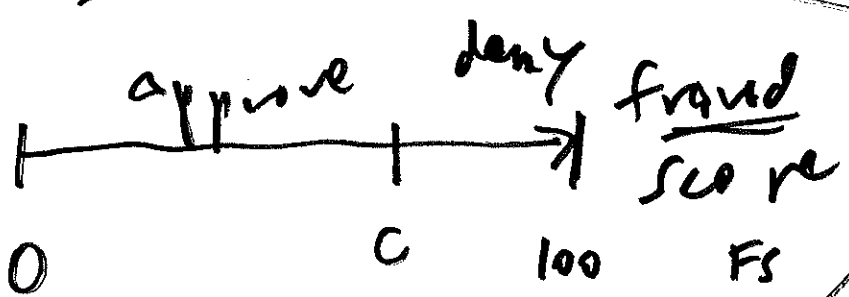
	truth		
	B	G	
system says	(+) 98	297	395
	(-) 2	9,603	9,605
prevalence	100	9,900	10,000

Labels for table: sensitivity (row 1), specificity (row 2), prevalence (row 3), truth (col 1), system (col 2), transactions (col 3)

we want  $P(B | \oplus) = \frac{98}{395} = 25\% \text{ (!)} \text{ (2)}$

(false positive rate)  $= P(G | \oplus) = 75\% \text{ (!)}$

(false negative rate)  $= P(B | \ominus) = \frac{2}{9605} = .00021 = 0.02\%$



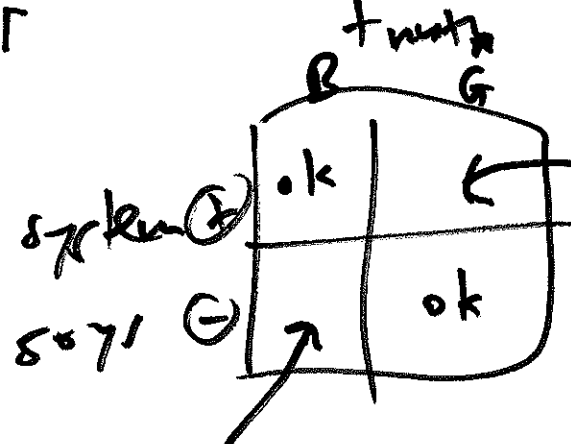
is the card bad?

fact (statistical inference)

should we deny the transaction?

choice

Bayesian decision theory



- agents:
- customer (3)
  - merchant
  - bank (building system)

bank suffers small loss (bad) but not really bad

card bad (but system says good) (bad)  
(much worse to bank)

Simulation - based approximate probabilities

- (1750) Button needle
- (1908) W. Gosset (Monte Carlo sampling)
- (1941) Metropolis & Ulam  
John von Neumann  
algorithms for pseudo-random generation

fact: All probabilities are

conditional; conditional on

assumptions

$P(H) = \text{undefined}$

$P(H | \text{fair coin tossing}) = \frac{1}{2}$

hard

$$P(A) = P(A \text{ and } B_1 \text{ or } A \text{ and } B_2 \text{ or } \dots \text{ or } A \text{ and } B_k)$$

(no-overlap)  
simple  
addition rule

$$\begin{aligned} \text{for } \textcircled{\text{or}} &= P(A \text{ and } B_1) + P(A \text{ and } B_2) \\ &+ \dots + P(A \text{ and } B_k) \\ &= \text{easy } P(B_1) \cdot \text{easy } P(A|B_1) + P(B_2) \cdot P(A|B_2) \\ &+ \dots + P(B_k) P(A|B_k) \end{aligned}$$

---


$$\underline{\underline{P(A)}} = \sum_{j=1}^k P(B_j) P(A|B_j) \quad (\text{LTP})$$