Discussion Section 5

p. R. 56

AMS 7 7 May 18

①

G = card really is good (truth)

B = card really is bad

+ = system says card is bad

− = ———— good

$P(B) = 1\%$ (prevalence)

$P(-\mid G) = 97\%$ (specificity) ~~[illegible]~~

$P(+\mid B) = 98\%$ (sensitivity) ~~[illegible]~~

$P(B \mid +) \doteq 25\%$ (show) (Bayes's Theorem AMS 13)

sensitivity — truth — specificity

what system says	B	G	
+	98	297	395
−	2	9,603	9,605
	100	9,900	10,000

(100 and 9,900: prevalence)

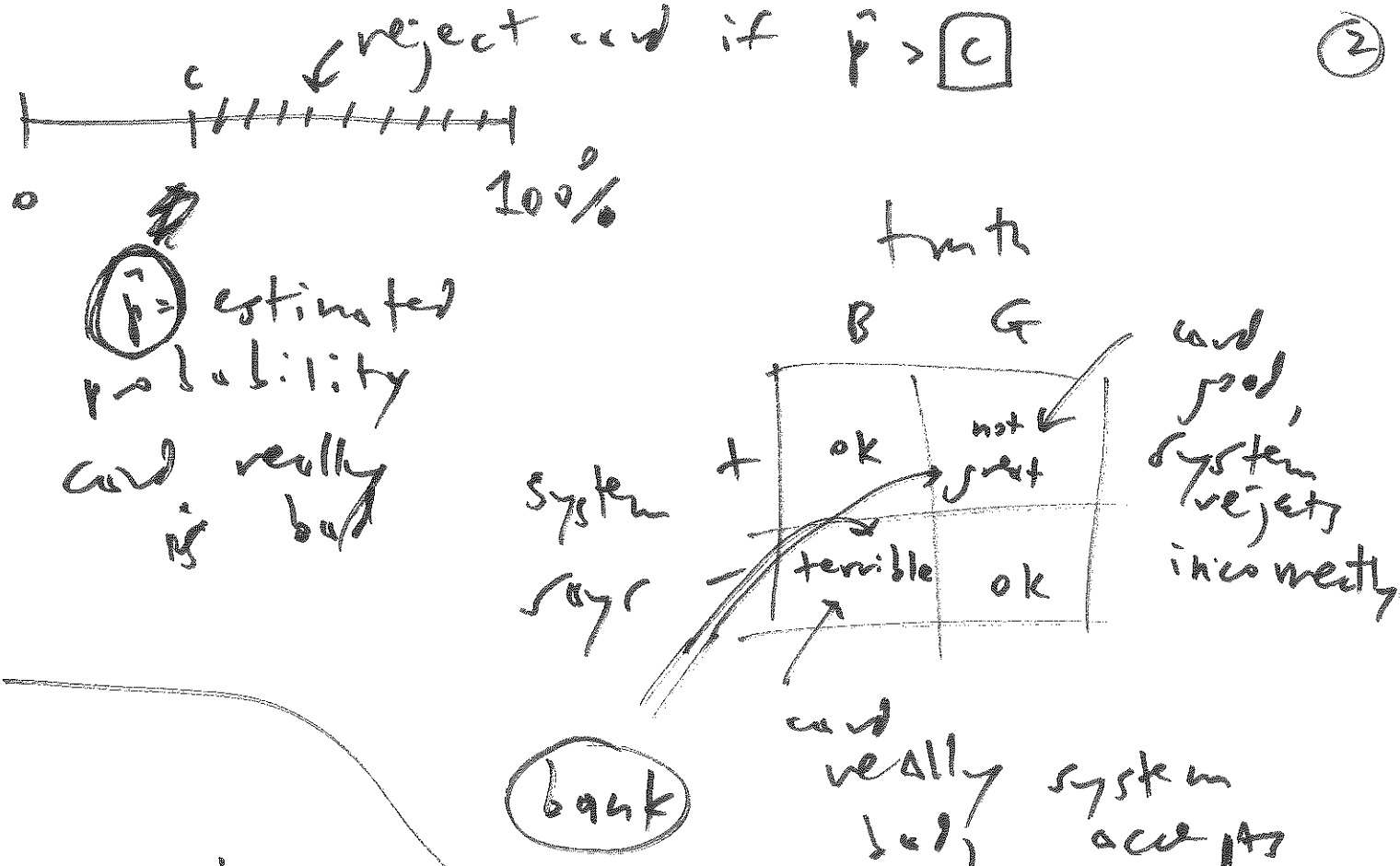
why?

① low prevalence

$$P(B \mid +) = \frac{98}{395} \doteq 25\%$$

② $98 + \boxed{297}$ (!)

specificity ideally should have been ↑



$$P(G \mid -) = \frac{9603}{9,605} = 100\%$$

$1 - P(B \mid +)$ = false positive rate

$= 75\%$ (!)

③

pop — all adult London underground riders

sample — the observed people

imag. data

weight

$N = ?$

100,000

actual ~~likely~~ IID

weight (lb.)

$n = 192$

sum $S = ?$ (ex. 31,100)

31,100

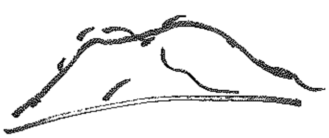
32,580

⋮

as S, $n \to \infty$

mean $\mu = 158$ lb.

SD $\sigma = 33$ lb.



pop. hist

hyp.

$n = 192$

sum $S = ?$ (ex. 32,580 lb)

long run mean

long run SD

long run hist

⋮

P(overload) =

$P(S > 31{,}400 \text{ lb.})$