Monday, August 31, 2015 12:05 PM

A blood test is 95% accurate (detects a sick person as sick), but has 1% false positive (detects a healthy person as sick). We know 0.5% population are sick.

Q: if a person is tested positive, what is the prob. she is really sick?

E: { tested positive }

S: { the person is sick }

$p(E|S) = 0.95$

$p(E|\bar{S}) = 0.01$

$p(S) = 0.005$

Q: $p(S|E)$?

$$p(S|E) = \frac{p(SE)}{p(E)} = \frac{p(E|S)\cdot p(S)}{p(E)} \approx 0.323$$

$$p(E) = p(E|S)\cdot p(S) + p(E|\bar{S})\cdot p(\bar{S})$$

$$= 0.95 \times 0.005 + 0.01 \times 0.995$$

1000 persons $\Rightarrow$ 5 sick, 995 healthy

5 positive

10 positive

① Model:

email contains

$E = \{\text{dollar}, \text{cheap}\}$

Monday, August 31, 2015 12:27 PM

From training data, we know that a spam email has prob. 0.2 to contain 'dollar', 0.5 to contain 'cheap', ….; a normal email has prob. 0.05 to contain 'dollar', 0.01 to contain 'cheap',….
Among all received emails by our email server, 10% are spam and 90% are normal emails

$p(S|E)$?

$p(H|E)$?

$p(\text{dollar}|S) = 0.2$

$p(\text{cheap}|S) = 0.5$

$p(\text{dollar}|H) = 0.05$

$p(\text{cheap}|H) = 0.01$

$p(S) = 0.1$

$p(H) = 0.9$

$$p(S|E) = \frac{p(SE)}{p(E)} = \frac{p(E|S) \cdot p(S)}{p(E)} = 0.95?$$

$$p(E|S) = p(\text{dollar}|S) \cdot p(\text{cheap}|S) = 0.1$$

$$p(E) = p(E|S) \cdot p(S) + \underbrace{p(E|H)}_{0.05 \times 0.01} \cdot p(H) =$$

$f_X(1) = p$ $\quad f_X(0) = 1-p$ $\quad p(AB) = p(A) \cdot p(B)$

Monday, August 31, 2015 12:46 PM

$p(m, h) = p(m) \cdot p(h) = (1-p) \cdot p$

$p = 0.1$, $n = 35$ $\quad$ r.v. X : # of active users

$p(\text{congestion}) = p(>10 \text{ users active at the same time})$

$= p(X > 10) = p(X=11) + \cdots p(X=35)$

$= 1 - p(X \leq 10) = 1 - [p(X=0) + p(X=1) + \cdots p(X=10)]$