Lecture 6: Independence

Ex

COVID has prevalence rate of 1%

we test for COVID and get a + or −

↳ test accurately report + 95% (sensitivity)

↳ 〃 〃 − 99% (specificity)

Q: I get a +, what's prob I have COVID?

D = have COVID $\quad P(D) = .01$

D^c = don't have COVID $\quad P(D^c) = .99$

+ or − $= +^c$

$P(+ \mid D) = .95 \Rightarrow P(- \mid D) = .05$

$P(- \mid D^c) = .99 \quad \boxed{\Rightarrow P(+ \mid D^c) = .01}$

Want: $P(D \mid +)$

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)}$$

$$= \frac{(.95)(.01)}{(.95)(.01) + (.01)(.99)}$$

$$\approx .49$$

Independence

Laymen's defn:

→ things don't affect each other

→ knowing if A occurred doesn't affect prob. of B occuring.

Defn: Independence

We say A and B are independent

denoted $A \perp B$

if $P(AB) = P(A)P(B)$.

→ distributive law for intersections

→ justifies product notation for $\cap$

Theorem: If $A \perp B$ then

$$P(A \mid B) = P(A).$$

pf. $$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)\cancel{P(B)}}{\cancel{P(B)}} = P(A).$$

Ex. Roll two dice independently.

$P(\text{at least one } 6)$

$= 1 - P(\text{no } 6\text{s})$

$A_1 =$ no 6 on first roll

$A_2 =$ " second

$= 1 - P(A_1 A_2)$

$= 1 - P(A_1)P(A_2)$ [independence]

$= 1 - (5/6)(5/6)$

$= 11/36$.

Counting persp.

Sampling twice ($r=2$) from $\{1, \dots, 6\}$
($n=6$)

$\rightarrow$ w/ repl.

Ordered: $|S| = 6^2 = 36$

$E =$ "at least one 6"

$= \{(1,6), (2,6), (3,6), (4,6), (5,6), (6,6)$
$(6,1), (6,2), (6,3), (6,4), (6,5)\}$

So $P(E) = \frac{|E|}{|S|} = \frac{11}{36}$.

Unordered:

$$|S| = \binom{n+r-1}{r} = \binom{6+2-1}{2} = \binom{7}{2} = 21$$

$$E = \{\{1,6\},\{2,6\},\{3,6\},\{4,6\},\{5,6\},\{6,6\}\}$$

$$P(E) = \frac{|E|}{|S|} = 6/21.$$

Ex. Roll two dice indeply.

$E = \{$ 1 or 2 on first, 3, 4, 5 on second $\}$

Ordered: $|S| = n^r = 6^2 = 6 \cdot 6$

$$E = \{1,2\} \times \{3,4,5\}$$

$$|E| = 2 \cdot 3$$

$$P(E) = \frac{|E|}{|S|} = \frac{2 \cdot 3}{6 \cdot 6} = \left(\frac{2}{6}\right)\left(\frac{3}{6}\right)$$

prob. 1 or 2 on first

prob. 3, 4, 5 on second.

Theorem:

If $A \perp B$ then

1. $A \perp B^c$
2. $A^c \perp B$
3. $A^c \perp B^c$.

Pf. $P(AB^c)$

$$= P(A) - P(AB)$$
$$= P(A) - P(A)P(B)$$
$$= P(A)(1 - P(B))$$
$$= P(A)P(B^c).$$

Defn: (Mutual) Independence

Generalize to multiple events

If (A_i) is a seq of events we say they are (mutually) independent if for all subsequences $A_{i_1}, \dots, A_{i_k}$

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

Q: Do I really need to check all subsequences?

Can I just check?

$$P(A_1 A_2 A_3 \cdots A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n)$$

No.

Ex. Roll two dice

$A =$ "doubles" $= \{(1,1), (2,2), \ldots, (6,6)\}$

$|A| = 6$

$B =$ "sum between 7 and 10"

$= \{(1,6), (2,5), (3,4), \ldots \}$

$|B| = 18$

$C =$ "Sum is 2, 7 or 8"

$= \{(1,1), (2,5), \ldots, (2,6) \ldots \}$

$|C| = 12$

Are A, B, C mutually indep?
Check all subseqs.

$$P(ABC) = P(A)P(B)P(C)$$

$$\frac{1}{36} = \left(\frac{6}{36}\right)\left(\frac{18}{36}\right)\left(\frac{12}{36}\right)$$

$$= (1/6)\left(\frac{1}{2}\right)(1/3)$$

$$= 1/36$$

Consider B and C:

$$P(BC) = \frac{11}{36} \stackrel{?}{=} P(B)P(C) = \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)$$

$$= 1/6$$

Sum is 7 or 8

Not mutually indep.

Can we have: $A \perp A$?

$$P(A) = P(AA) = P(A)P(A) = P(A)^2$$

Works if $P(A) = 0$ or 1.

Ex. JWST had ~400 points of failure.

$$\boxed{T1} \to \boxed{T2} \to \boxed{T3} \to \cdots \to \boxed{T400}$$

JWST fails if any task fails.

W_i = i^{th} task succeeds

W_i^c = i^{th} task fails

Assume they are independent.

$$P(W_i^c) = 1/1000$$

P(JWST works)

$$= P(W_1 W_2 W_3 \cdots W_{400})$$

$$= P(W_1)P(W_2) \cdots P(W_{400})$$

$$= (1-P(W_1^c))(1-P(W_2^c)) \cdots (1-P(W_{400}^c))$$

$$= (1-1/1000)(1-1/1000) \cdots (1-1/1000)$$

$$= (1-1/1000)^{400}$$

$\cong 67$

↑ EXAM 1 ↑

Ex. Flip a coin 3 times.

X = # heads

$s \in S$	$X(s)$
HHH	3
THH	2
HTH	2
TTH	1
HHT	2
THT	1
HTT	1
TTT	0

Defn: Random Variable

A random variable (RV) X is a function

$$X : S \to \mathbb{R}$$

also called: random variate,
real-valued RV,
univariate RV.

Ex. (1) toss two dice

X = sum of dice

(2) toss a coin 25 times,

X = length of longest consecutive chain of Hs