2.4 Bayes rule n law of total probability

WK 3 - 1

LAW OF TOTAL PROBABILITY

assuming that B_i + B_j are mutually exclusive + exhaustive which means $B_i \cap B_j = \emptyset$, for $i \neq j$, $B_1 \cup B_2 \cup B_3 \dots \cup B_4 = \Omega$ then for any event A,

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

example for $k=3$, $P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$

* only works if B_i + B_j are mutually exclusive + exhaustive

TWO SETS ARE **EXHASTIVE** WHEN $A \cup B = \Omega$

BUT WHAT IS EXHAUSTIVE?

ex: roll a die once → $\Omega = \{1,2,3,4,5,6\}$

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

then A + B are disjoint + exhaustive → exhaustive because $A \cup B = \Omega$

EXAMPLE OF LAW OF TOTAL PROBABILITY:

ex: scenario → drawing 2 cards from a standard deck of cards

we want A (probability of heart 2 is a heart)

$P(H_2) = P(H_1)P(H_2|H_1) + P(\bar{H}_1)P(H_2|\bar{H}_1)$

$H_2 = (H_2 \cap H_1) \cup (H_2 \cap \bar{H}_1)$

there are only 2 scenarios for H_1, either H_1 or $\bar{H}_1$

ex: under good weather conditions, 80% of flights arrive on time. During bad weather, only 30% of flights arrive on time.

Tomorrow, chance of good weather is 60%. → so chance of bad weather = 40%

① What is the probability your flight will arrive on time?

* good + bad weather are mutually exclusive

* there are 2 things involved, good + bad weather / flight on time or not

to solve 4 on time:

A_1 = arriving on time

B_1 = good weather

B_2 = bad weather

$P(A_1) = \underbrace{P(B_1)P(A_1|B_1)}_{\text{good weather, flight arrive}} + P(B_2)P(A_1|B_2)$

$= (0.6 \cdot 0.8) + (0.4 \cdot 0.3) = 0.6$

② probability flight will not arrive on time

A_2 = not arriving on time

$P(A_2) = P(B_1)P(A_2|B_1) + P(B_2)P(A_2|B_2)$

$= (0.6 \cdot 0.2) + (0.4 \cdot 0.7)$

$= .12 + .28$

$= .4$ ②

③ probability of bad weather?

.4 ①

BAYE'S RULE

Flipping the rule of total probability w/ the same concept as the formula on conditional probability

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{k} P(A|B_j)P(B_j)}, \quad \text{for } i = 1, 2, 3, \dots, k$$

ex: for $k=3$, $i=2$ it looks like

$$P(B_2|A) = \frac{P(A|B_2)P(B_2)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)}$$

EX: Under good weather conditions, 80% of flights arrive on time. During bad weather, only 30% of flights arrive on time. The chance of good weather is 60%.

Given a flight has arrived on time, what's the probability that the weather is good.

$P(B_1)$ = good weather $P(B_2)$ = bad weather

$P(A)$ = arrive on time

$P(\bar{A})$ = not arrive on time

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2)}$$

$$= \frac{(0.6)(0.8)}{(0.6 \cdot 0.8) + (0.4 \cdot 0.3)} = \frac{0.48}{0.48 + 0.12} = \frac{0.48}{0.6} = 0.8$$

weather + time in this case are dependent

IN THIS CASE ↓

the information that the person has arrived on time, increased the chance of there being good weather

EX: A computer maker recieved parts from three suppliers, S1, S2, + S3. 50% come from S1, 20% comefrom S2, 30% from S3. Among all S_1 parts, 5% are defective. Among S2 parts, 3% are defective. Among S3 parts, 6% are defective. A customer complains his part is defective, what is the probability that it was supplied by S1.

↳ defective or not is <u>dependent</u> on supplier

→ SHOULD GIVE FRACTION EXACT ANSWER INSTEAD OF DECIMAL APPROXIMATIONS

S_1 $P(D|S_1) = .05$

S_2 $P(D|S_2) = .03$

S_3 $P(D|S_3) = .06$

D = defective

$$P(S_1) = \frac{P(S_1)P(D|S_1)}{P(S_1)P(D|S_1) + P(S_2)P(D|S_2) + P(S_3)P(D|S_3)}$$

$$= \frac{(.5)(.05)}{(.5\cdot.05) + (.2\cdot.03) + (.3\cdot.06)}$$

$$= \frac{0.025}{0.049} = \underbrace{\frac{25}{49}}_{\text{EXACT ANSWER}} \approx \underbrace{0.5}_{\text{approximate}}$$

For HW PROBLEM:

B_1 = has disease (1%) B_2 = don't have disease (99%)

A = Positive $\bar{A}$ = Negative result

↳ within disease testing, there are more to the positive + negative

∘ in + there are only a % of people who have the disease. same w/ negative

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)}$$

$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2)$

$$= \frac{0.01\cdot0.95}{0.01\cdot0.95 + 0.99\cdot0.06} = \frac{0.0095}{0.0689} = \frac{95}{689}$$

given a person w/ a positive test, the % the person <u>actually</u> has the disease is.

if 0.01 was much larger, like .2 then the probability would be much higher

more of the pop. has the disease

$P(A|B_2)$ is also important as it can greatly alter the denominator

is not conditional, $\cap$ is conditional

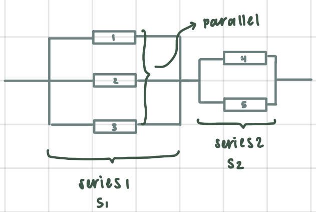
cond. problems will have more than 2 numbers.

new term!

RELIABILITY:

PROBABILITY THAT THE SYSTEM WORKS

EX: In the system



each component fails with probability 0.3 independently of other components. Compute the system's reliability.

∘ all need to work in a series for the system to work

∘ only one in parallel needs to work

reliability = P(system works) = $P(S_1 \cap S_2)$

$P(S_1) = P(1 \cup 2 \cup 3) = 1 - \underbrace{P(\bar{1} \cap \bar{2} \cap \bar{3})}_{\text{complement}}$

$= 1 - 0.3^3 = 1 - 0.027 = 0.973$

$P(S_2) = P(4 \cup 5) = 1 - P(\bar{4} \cap \bar{5}) = 1 - .3^2 = 1 - .09 = .91$

$P(S_1 \cap S_2) = P(S_1)\,P(S_2|S_1)$ ↳ the 2 parallel series are actually independent of each other

$= 0.973\,(.91)$

$= 0.88543$

Ch.3 DISCRETE RANDOM VARIABLES n DISTRIBUTIONS

wk 3-2

RANDOM VARIABLE (RV)

VARIABLE THAT ASSUMES VALUE BY CHANCE

PROPERTIES OF DISCRETE RV
$0 < P(X) < 1$
summation of all probabilities = 1

DISCRETE RV

POSSIBLE VALUES ARE LISTED IN SEQUENCE

EX: how many siblings? A: 1, 2, 3 ... or ... # ← only integers
like smth we can count

CONTINUOUS RV

POSSIBLE VALUES FROM AN INTERVAL

EX. salary, height, weight A: can be 0.000 – 1,000,000.000 ← all dec. #s in between

PROBABILITY MASS FUNCTION (PMF)

X = # of spots that appear when a fair die is thrown

probability distribution of X can be:

X	1	2	3	4	5	6
P(x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

a) $P(X=1) = P(1) = \frac{1}{6}$

b) $P(X \leq 2) = P(1) + P(2) = \frac{1}{3}$

c) $P(X<4) = P(1) + P(2) + P(3) = \frac{1}{2}$

only true for discrete

d) $P(X=-1) = 0$ → not part of solution set

e) $P(X \leq 2.5) = P(1) + P(2) = \frac{1}{3}$
highest # is 2

f) $P(X \geq 4) = P(4) + P(5) + P(6)$

↳ $P(X<4) = P(X \leq 3)$ → ARE THE SAME ONLY WHEN DISCRETE

ex 2 using ^ P(X)

$y = 2X$ + $z = X_1 + X_2$ → same probabilities as X, but X_1 + X_2 are independent

we construct P(y) + P(z)

$y = 2X$

y	2	4	6	8	10	12
P(y)	1/6	1/6	1/6	1/6	1/6	1/6

because $y = 2X$, $y = 2 = X = 1$
↳ same event = same probability

$z = X_1 + X_2$

z	2	3	4	5	6	7	8	9	10	11	12
P(z)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$P(z=2) = P(X_1 = 1 \cap X_2 = 1) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

$P(z=3) = P(X_1 = 1 \cap X_2 = 2) \overset{\text{"or"}}{+} P(X_1 = 2 \cap X_2 = 1) = \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{36} = \frac{1}{18}$

these 2 are mutually exclusive

mult. because they are independent

what about V? $V = X_1 - X_2$

V	-5	-4	-3	-2	-1	0	1	...
P(V)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	

same concept!

CUMULATIVE DISTRIBUTION FUNCTION (CDF)

F(x) of a discrete random variable X with pmf p(x) is defined for every real number x by $F(x) = P(X \leq x) = \sum_{y \leq x} p(y)$

fixed #

is a probability so $0 < F(x) < 1$

ex if $x = 2$ ↴
$F(2) = P(X \leq 2) = P(1) + P(2) = \frac{1}{3}$

EXAMPLE A fair die is rolled once. If it lands six, you win \$4. However, you lose \$1 if it does not land six.

Let m = amount of money you win, find distribution (PMF) for m

→ TO FIND DISTRIBUTION:
- MUST LIST POSSIBLE VALUES (THE RANGE)
- MUST LIST PROBABILITIES ASSOCIATED W/ EACH VALUE

PMF

M	\$4	-\$1
P(M)	1/6	5/6

CDF

$$F(x) = \begin{cases} 0 & x < 1 \\ 5/6 & -1 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

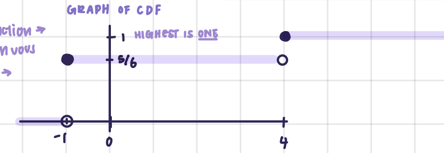
$F(-2) = P(X \leq -2) = 0$

$F(-1) = P(X \leq -1) = P(X=-1) = \frac{5}{6}$

$F(4) = P(X \leq 4) = P(X=-1) + P(X=4) = 1$

GRAPH OF CDF

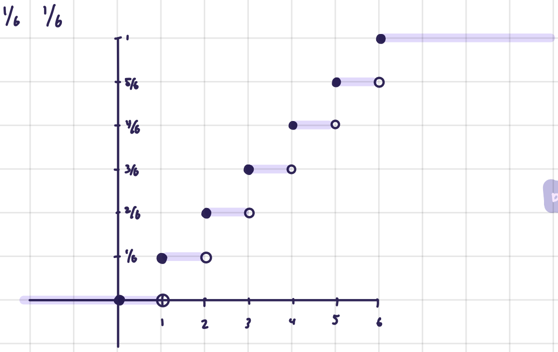
not dec. function → right continuous function →



EXTRA CDF EXAMPLE ON DIE

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

$$F(X) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \leq x < 2 \\ 2/6 & 2 \leq x < 3 \\ 3/6 & 3 \leq x < 4 \\ 4/6 & 4 \leq x < 5 \\ 5/6 & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$



PROPERTY OF DISCRETE DISTRIBUTION

1) $0 < P(X) < 1$ each prob. is between 0 + 1

2) $\sum P(X) = 1$ sum of all = 1