probablility

Adam Okulicz-Kozaryn

adam.okulicz.kozaryn@gmail.com

this version: Wednesday 30th September, 2020 10:19

outline

why bother? and intuition

computing probability

conditional probability and independence

what is it?

- the numerical measure of the likelihood that the event will occur (the proportion of times the outcome would occur)
- it ranges from 0 to 1
- 0 means impossible
  - never 0, almost nothing is impossible eg $\frac{1\ (crashes)}{10m\ (flights)}$
- 1 means certain
  - also almost never 1; almost nothing is certain
- probability is really the basis of all statistics

outline

why bother? and intuition

computing probability

conditional probability and independence

making right decisions

- people and orgs make mistakes bc miscalculate probab
  - eg gambling and lotteries
- smoking (hundreds or thousands of % increased risks!)
- flying v driving, etc
  - 9/11 killed additional thousands because people chose to drive (Wheelan, 2013, p.72-3)
- think about probabilities when making a decision
- the easiest (but already informative and helpful):
  - $\frac{occurences}{total}$; eg : $\frac{cancers}{smokers}$; $\frac{crashes}{miles}$; $\frac{crashes}{hours\ travelled}$
  - so depends how you measure ! car v flight: hours v miles !

why is probabilty relevant to MPA student?

- probability may be confusing !
- but probability improves thinking/decision making
  - it is everywhere !
- usually don't realize it, but we calculate prob all the time!
  - and it helps our lives enormously if can do it better!
- how likely that this will be on exam–should i study it?
- how likely i will get caught if i am speeding?
- how likely a student will drop out?
- how likely that graduation rate is above some value?

important for organizations!

- eg identify teachers who cheat or doctors who overcharge
- (just identify outliers, and unlikely events (nontypical))
- (or overall changes/trends (for everyone) that are suspect (collusion?))
- public policy makers determining dangerous items
  eg swimming pools kill more kids than guns
  - (just see how many there are and how many people die from them!–simple des stat)
- for more interesting examples see Levit's Freakonomics
  `http://www.freakonomics.com/`

evolutionary and counterintuitive!

- evolution made us to survive in an environment that is long time gone!
  - so is our cognitive function and probability calcualtion off!
- and so we need statistics to help us think better!
- eg: overestimate prob of memorable/flashy events
  - terrorist attack and airplane crash are similar to mistaking stick for snake etc–better be extra careful and see even if it's not there
- we understimate, on the other hand
  - much more deadly effects of sugar and fat (which were always rare and desirable)

i used to have a policy about undocumented emergencies

- and i do not have it anymore! thank god for probability!
- there were 7 students and 4 of them had their grandmas die
- the reported probability was too high to be plausible!
- so what's the prob of grandma dying during this semester?
  - say avg grandma is expected to last at least (if not more) about 10 years, or 40 three-month periods in a year (about semster long)
  - so about $1/40$, so for 10 students class: every 4 semsters one grandama dead

outline

why bother? and intuition

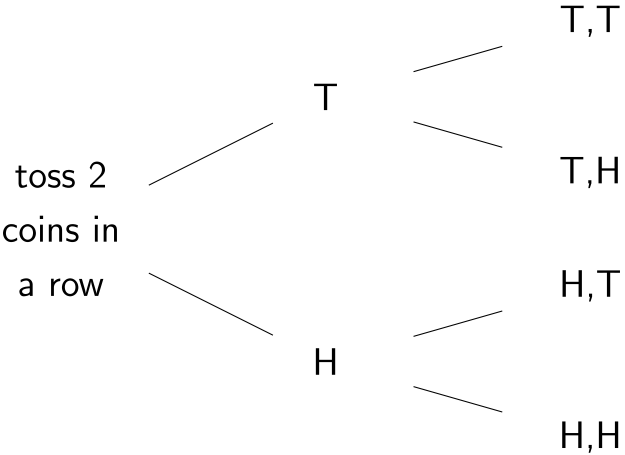
computing probability

conditional probability and independence

concepts

- event (A), (B), etc: a set of outcomes, eg $A = \{T, T\}$
- sample space (S): the finest grain, mutually exclusive, collectively exhaustive listing of all possible outcomes
- mutually exclusive (disjoint):
  - 2 outcomes can't occur at the same time, eg T and H
- collectively exhaustive:
  - 1 outcome must occur: eg eg T or H
- eg toss a coin twice
- $S = \{H, H\}, \{H, T\}, \{T, H\}, \{T, T\}$
- $A = \{T, T\}$

tree



table

HH	HT
TH	TT
- 1st row H in first flip
- 2nd row T in first flip
- 1st column H in second flip
- 2nd column T in second flip

exercises...

- what is the probability of getting 2 T in 2 flips ?
- we just showed with tree and table that there are 4 possible events, and only one outcome with 2 T, so $P = \frac{1}{4}$
- how about at least 1 T ?
- $P = \frac{3}{4}$
- how about exactly 1 T ?
- $P = \frac{2}{4}$

some useful properties and useful language

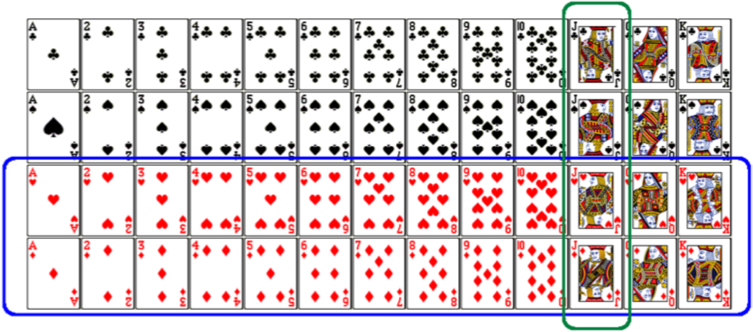
- intersection (and) $\cap$ (both have to happen)
  - multiply prob ($=< 1$), so less likely to have both (frownie)
- union (or) $\cup$ (at least one has to happen)
  - add prob, so more likely to have both (smiley)
- eg $A = H \;\; B = H$
  - $A \cap B = \{H, H\}$
  - $A \cup B = \{H, H\}, \{H, T\}, \{T, H\}$
- if mutually exclusive: $P(A \cup B) = P(A) + P(B)$
- if mutually exclusive and collectively exhaustive: $P(A) + P(B) = 1$
  - eg $A = tail \;\; B = head$; $A = male \;\; B = female$

cards examples

- $P(heart) = \frac{1}{4}$
- mutually exclusive
  $P(A \cup B) = P(A) + P(B)$
  - $P(ace\ or\ king) = P(ace) + P(king) = \frac{1}{13} + \frac{1}{13} = 2/13$
- not mutually exclusive
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - $P(ace\ or\ black) = P(ace) + P(black) - P(ace\ and\ black) = \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = 7/13$
- $P(heart|red) = 1/2$ because sample space is reduced to 26 red cards (will get back to it at the end!)

Union of non-disjoint events

What is the probability of drawing a jack or a red card from a well shuffled full deck?



$$P(\text{jack or red}) = P(\text{jack}) + P(\text{red}) - \color{red}{P(\text{jack and red})}$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

Figure from http://www.milefoot.com/math/discrete/counting/cardfreq.htm

- **What is the probability that a randomly sampled student thinks marijuana should be legalized <u>or</u> they agree with their parents' political views?**

	Share Parents' Politics		
Legalize MJ	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(a) (40 + 36 - 78) / 165

(b) (114 + 118 - 78) / 165

(c) 78 / 165

(d) 78 / 188

(e) 11 / 47

Product rule for independent events

P(A and B) = P(A) x P(B)

Or more generally, P(A1 and … and Ak) = P(A1) x … x P(Ak)

You toss a coin twice, what is the probability of getting two tails in a row?

P(T on the first toss) x P(T on the second toss)

= (1 / 2) x (1 / 2) = 1 / 4

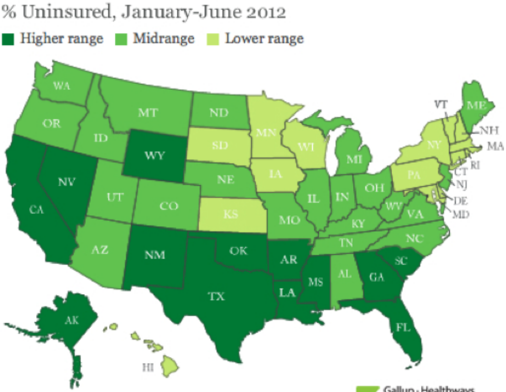
A recent Gallup poll suggests that 25.5% of Texans do not have health insurance as of June 2012. Assuming that the uninsured rate stayed constant, what is the probability that two randomly selected Texans are both uninsured?

(a) 25.5^2

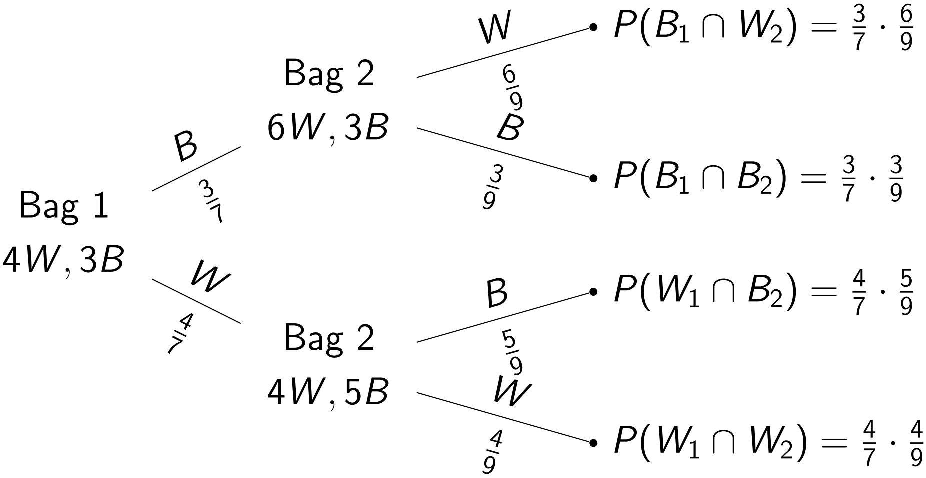
(b) 0.255^2

(c) 0.255×2

(d) $(1 - 0.255)^2$



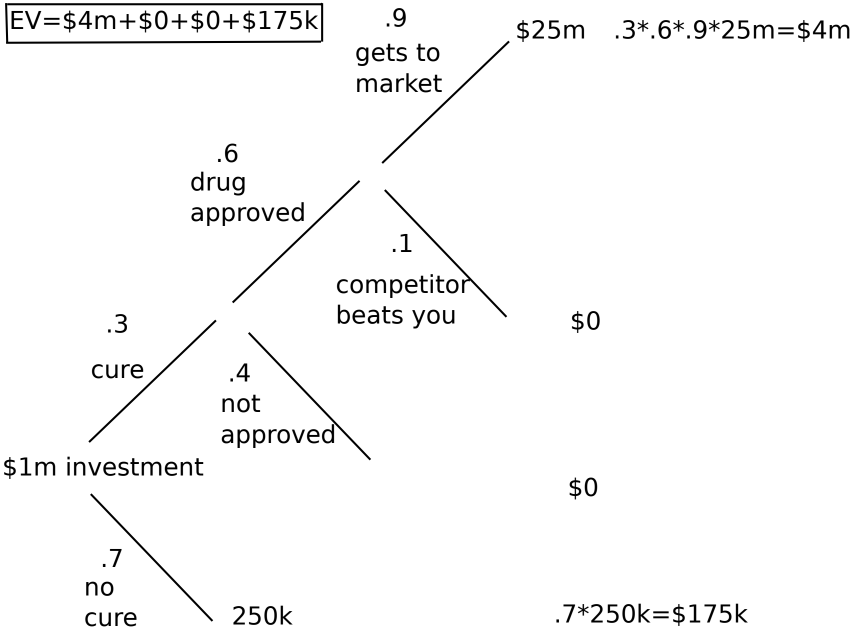
tree example



http://www.onemathematicalcat.org/Math/Algebra_II_obj/prob_tree_diagrams.htm

Expected Value (Wheelan, 2013, p83)

- just multiply value ($ amount) by associated probability(ies) AND add them up
- and this is how much you are expected to get on average


EV=$4m+$0+$0+$175k
.9
gets to
market
$25m
.3*.6*.9*25m=$4m
.6
drug
approved
.1
competitor
beats you
$0
.3
cure
.4
not
approved
$1m investment
$0
.7
no
cure
250k
.7*250k=$175k

outline

why bother? and intuition

computing probability

conditional probability and independence

conditional probability

- $\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$
- you have $P(A \cap B)$ in numerator because both A and ($\cap$) B need to happen to be conditional on B, if A happens but not B, then it cannot be conditional on B

table practice (all numbers in the body are "$\cap$")

	Type of Policy (%)			
Category	**Fire**	**Auto**	**Other**	**Total %**
Fraudulent	6	1	3	10
Nonfraudulent	14	29	47	90
Total	20	30	50	100

-
- $P(fire) = \frac{20}{100} = .2$

◦ $P(F|fire)$ is $6/20$ or $.06/.2$

◦ $P(fire|F)$ is $.06/.1$

strategy

- yes, probability can be confusing...but
- probability is not a rocket science
- if you think about it you can figure it out
- formulas may be more confusing than revealing
- you can just use tables or possibly best to use trees...
- i don't care what method you use
- we'll practice more next week!

Levitt, S. D. and S. J. Dubner (2010): Freakonomics, vol. 61, Sperling & Kupfer.

Wheelan, C. (2013): Naked statistics: stripping the dread from the data, WW Norton & Company.