THE SURE THING PRINCIPLE (STP) AND SIMPSON'S PARADOX (SP)

Z. ŠIKIĆ

(UNIV. OF ZAGREB)

L. T. SAVAGE (FoS, 54)

"IF PERSON PREFERES f TO g, EITHER KNOWING THAT B OBTAINED, OR KNOWING THAT NON-B OBTAINED, THEN HE SHOULD PREFER f TO g IF HE KNOWS NOTHING ABOUT B"

IT IS NEITHER LOGICAL NOR PROBABILISTIC PRINCIPLE!

NON-EPISTEMIC VERSION:

- IF C INCREASES THE PROBABILITY OF E IN EACH SUBPOPULATION THEN IT ALSO INCREASES IT IN THE POPULATION AS A WHOLE.

(C = CAUSE
E = EFFECT)

G.R. BLYTH (ON S.P. & S.T.P. 72)

S.P. CONTRADICTS S.T.P.

	E	$\overline{E}$
C	20	20
$\overline{C}$	16	24

CAUSE = TREATMENT
EFFECT = RECOVERY

$pr(E \mid C) = 20/40 = 50\%$

$pr(E \mid \overline{C}) = 16/40 = 40\%$

WE SHOULD PREFERE C !

BUT WHAT IF FOR MALE AND FEMALE PATIENS (m:f) WE HAVE:

	E	$\bar{E}$
C	18 : 2	12 : 8
$\bar{C}$	7 : 9	3 : 21

$$pr(E \mid C, m) = 18/30 = 60\%$$

$$pr(E \mid \bar{C}, m) = 7/10 = 70\%$$

FOR MALES WE SHOULD PREFERE $\bar{C}$

$$pr(E \mid C, f) = 2/10 = 20\%$$

$$pr(E \mid \bar{C}, f) = 9/30 = 30\%$$

FOR FEMALES WE SHOULD PREFERE $\bar{C}$

	$E \& C$		$E \& \bar{C}$
m	18/30=60%	<	7/10=70%
f	2/10=20%	<	9/30=30%
all	20/40=50%	>	16/40=40%

CONTRARY TO S.T.P. $\bar{C}$ INCREASES THE PROBABILITY OF E IN m AND f SUBPOPULATION, BUT NOT IN THE POPULATION OF ALL PATIENTS

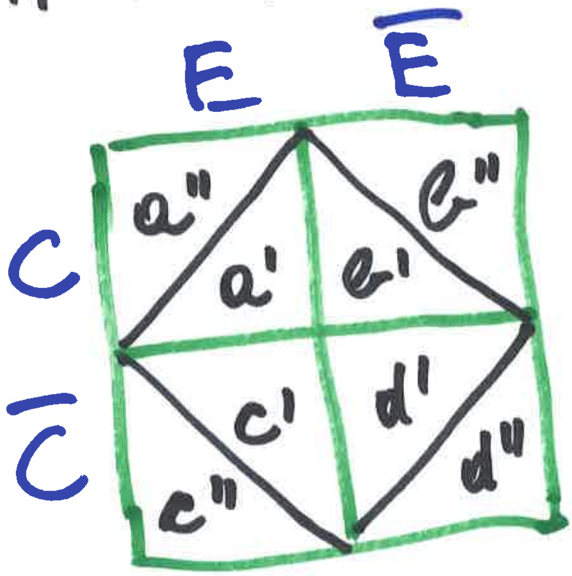
	E	$\bar{E}$
C	a	b
$\bar{C}$	c	d

$\boxed{\text{Cause } \underline{\text{is}} \text{ Effective}} \equiv$

$$p(E|C) > p(E|\bar{C}) \equiv \frac{a}{a+b} > \frac{c}{c+d}$$

$$\equiv \boxed{ad > bc}$$

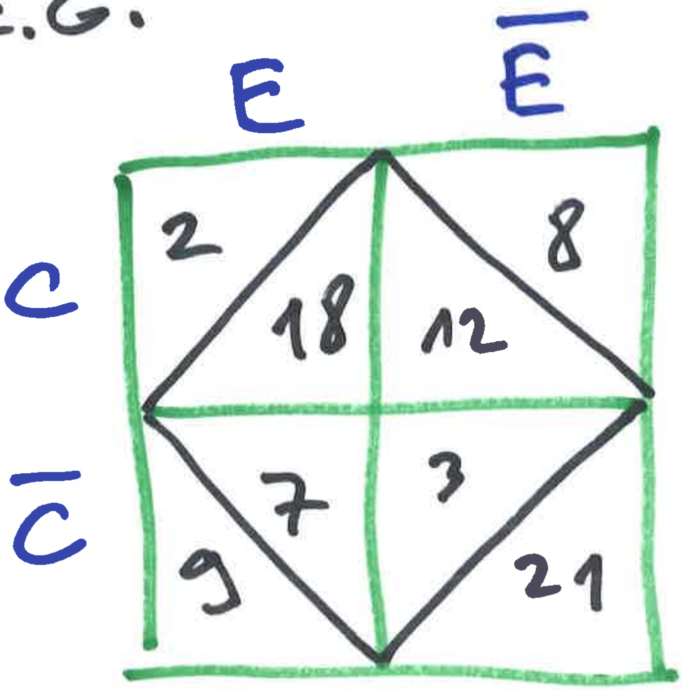
IF WE INTRODUCE PARTITION $P, \bar{P}$



$$a'd' > b'c' \ \& \ a''d'' > b''c''$$

$$\not\Rightarrow \ ad > bc$$

E.G.



$18 \cdot 3 < 7 \cdot 12$

$2 \cdot 21 < 9 \cdot 8$

BUT

$20 \cdot 24 > 16 \cdot 20$

WE SHOULD PRESCRIBE THE TREATEMENT WHEN WE DO NOT KNOW THE GENDER OF A PATIENT.

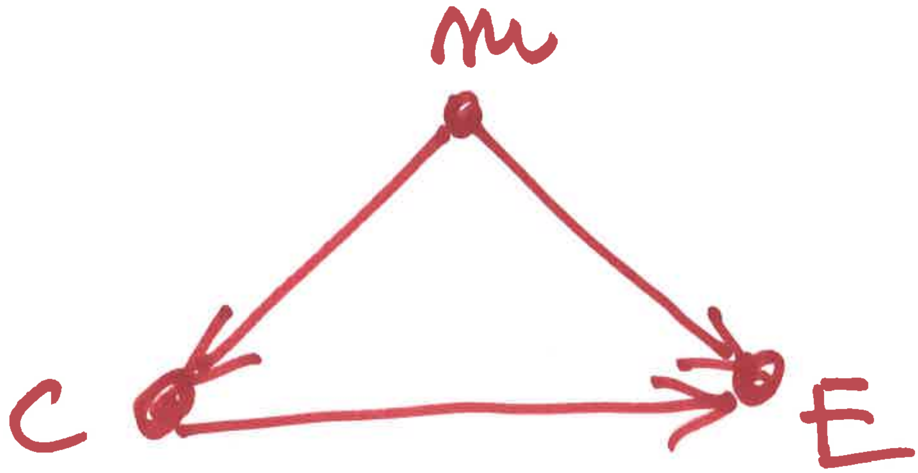
WE SHOULD NOT PRESCRIBE THE TREATEMENT WHEN WE KNOW THE GENDER OF A PATIENT.

NUNSENSE !!!

WHERE DOES THE TRUTH LIE?

IN SUBPOPULATIONS OR IN POPULATION?

IF MALES GET THE TREATEMENT MORE OFTEN & SPONTANEOUSLY RECOVER MORE OFTEN:

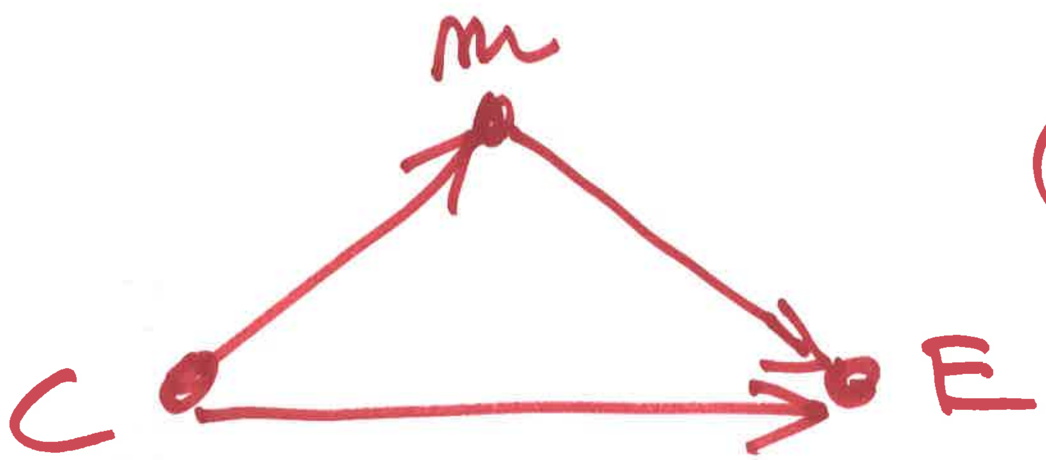


(CBN)

THEN THE TREATEMENT LOOKS BENEFICIAL ALTHOUGH IT IS NOT I.E.

"THE TRUTH IS IN SUBPOPULATIONS"

IF THE TREATEMENT STA-
BILIZES m (E.G. Bl. Pr.)
& THAT LEEDS TO RECOVERY:



(CDN)

THEN SUBPOPULATIONS CON-
FOUND INFLUENCES OF C
AND m I.E.

"THE TRUTH IS IN THE PO-
PULATION AS A WHOLE"

PEARL : IT HAPPENS BECAUSE SUBPOPULATIONS ARE NOT CAUSALY INDEPENDENT OF C.

cf.

WILL ROGERS :

WHEN THE OKIES MOVED TO CALIFORNIA, THEY RAISED THE AVERAGE INTELLIGENCE IN BOTH STATES.

CONTRARY TO S.T.P. THE MOVEMENT DID NOT CHANGE THE AVERAGE INTELLIGENCE IN THE COMBINED STATES.

THE TRUE VERSION OF S.T.P.

IF C INCREASES THE PROBABILITY OF E IN EACH SUB-POPULATION THEN IT ALSO INCREASES IT IN THE POPULATION AS A WHOLE PROVIDED THAT C DOES NOT CHANGE THE SUBPOPULATIONS.

$\Longrightarrow$

$$pr(m|c) = pr(m|\bar{c}) = pr(m)$$

$$pr(\bar{m}|c) = pr(\bar{m}|\bar{c}) = pr(\bar{m})$$

THE TRUE VERSION IS A PROBABILISTIC PRINCIPLE

$$
\begin{array}{ccccc}
pr(E|c) & = & pr(E|c,m)\,pr(m) & + & pr(E|c,\bar{m})\,pr(\bar{m}) \\
\vee & & \vee & & \vee \\
pr(E|\bar{c}) & = & pr(E|\bar{c},m)\,pr(m) & + & pr(E|c,\bar{m})\,pr(\bar{m})
\end{array}
$$

$$\boxed{\text{IF} > \text{AND} > \text{THEN} >}$$

SO, WHERE IS THE PARADOX?

PEARL:
SIMPSON'S REVERSALS SEEMS PARADOXICAL BECAUSE OF

(1) OUR CAUSAL INTUITIONS (S.T.P.)
(THERE IS NO MIRACLE TREATMENT THAT IS HARMFUL TO BOTH MALES & FEMALES AND IS SIMULTANEOUSLY BENEFICIAL TO EVERYBODY)

(2) THE ASSUMPTION THAT CAUSAL RELATIONS ARE GOVERNED BY PROBABILITY LAWS
(THERE IS SUCH A MIRACLE TREATMENT)

+ THE TREATMENT HAS NO EFFECT ON GENDER !!!

BUT!!! THE TRUE VERSION OF S.T.P. IS PROBABILISTIC

THE MOST COMMON ARGUMENT FOR INDEPENDENCE OF C AND M IS CAUSAL BUT THE PROOF OF T.V. S.T.P. USES ONLY INDEPENDENCE.

I SUPPOSE PEARL WOULD SAY:

THE "PARADOX" IS PSYCHOLOGICAL AND PSYCHOLOGICALY WE THINK IN CAUSAL TERMS

BANDYOAPDHYAY, NELSON, BRITTAN, GREENWOOD, BERWALD (2010) "THE LOGIC OF S.P." SYNTHESE

SIMPSON'S REVERSALS ARE PARADOXICAL AND OPONENTS "MUST ARGUE THAT THEY ARE NOT SURPRISING"

LET ME ARGUE THAT THEY ARE NOT SURPR-ISING (IN FEW STEPS); AFTER FEW MORE EXAMPLES

	1. STORE	2. STORE
A BRAND	$20 \cdot \$3 = \$60$	$80 \cdot \$4 = \$320$
B BRAND	$80 \cdot \$5 = \$400$	$20 \cdot \$6 = \$120$
AVERAGE	$\$4.6$	$\$4.4$

BERKELEY

	WOMEN		MEN
S. DEPART.	106/133 = 80%	>	864/1385 = 60%
H. DEPART.	451/1702 = 26%	>	334/1306 = 25%
ALL	557/1835 = 30%	<	1198/2691 = 45%

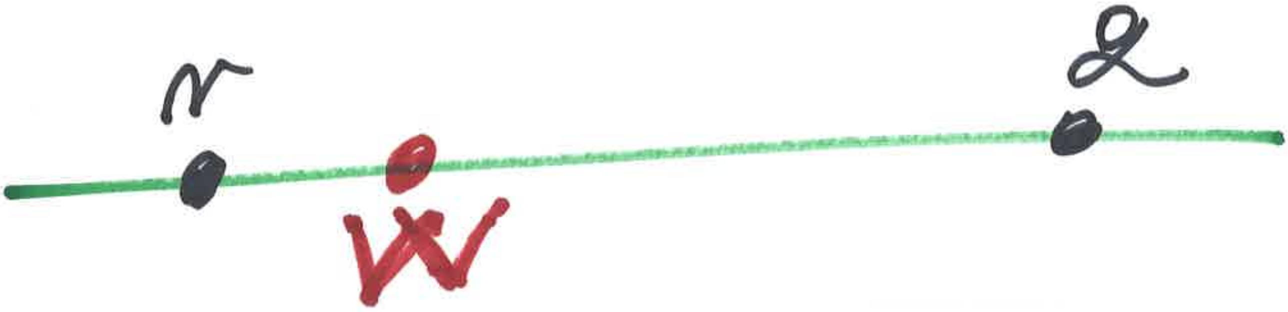
	H.S. STUDENTS		
	EXCELENT		AVERAGE
5*-UNIV.	1/9 = 11%	>	0/1 = 0%
1*-UNIV.	1/1 = 100%	=	99/99 = 100%
ALL	2/10 = 20%	<	99/100 ≈ 100%

IS IT BETTER TO BE AVERAGE?

	CITY HOSPITAL		STATE HOSPIT.
TYP. PAT.	870/900 = 97%	<	590/600 = 98%
ATYP. PAT	30/100 = 30%	<	210/400 = 55%
ALL	900/1000 = 90%	>	800/1000 = 80%

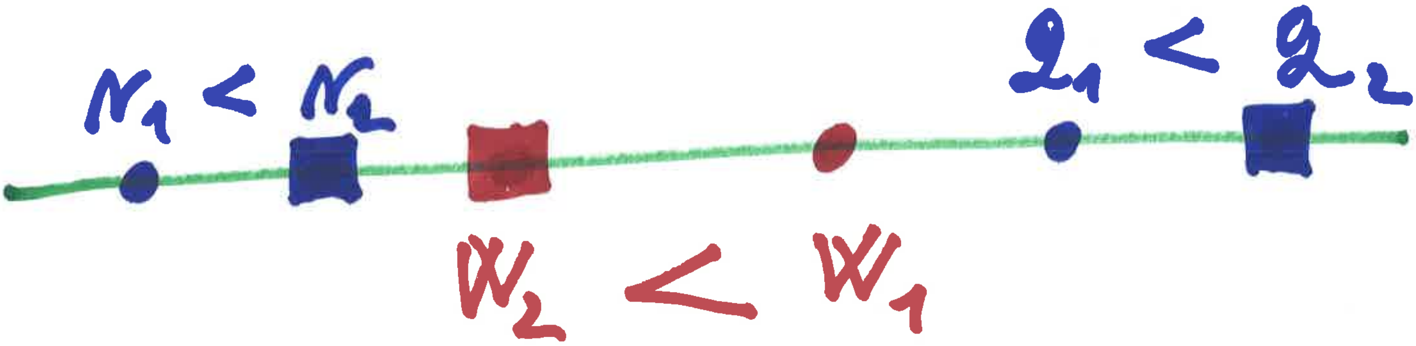
W IS A WEIGHTED AVERAGE OF $r < q$ $:\Longleftrightarrow$ $r < w < q$

i.e. $\boxed{w = \alpha r + \beta q \quad \alpha, \beta > 0 \quad \alpha + \beta = 1}$



I FIRST ARGUE THAT IT IS NOT SURPRISING THAT:

$r_1 < r_2$, $q_1 < q_2$ & $w_1 > w_2$



m IS A MEDIANT OF $r < g$

$$:\equiv \; r = \frac{a}{A}, \; g = \frac{b}{B} \; \& \; m = \frac{a+b}{A+B}$$

WE USE MEDIANTS IN S. REVERSALS:

A = CARDINALITY OF A-SUBPOPULATION

B = CARDINALITY OF B-SUBPOPULAT.

$A+B$ = CARDINALITY OF POPULATION

r = INCIDENCE OF SMTH IN A-SUBP.

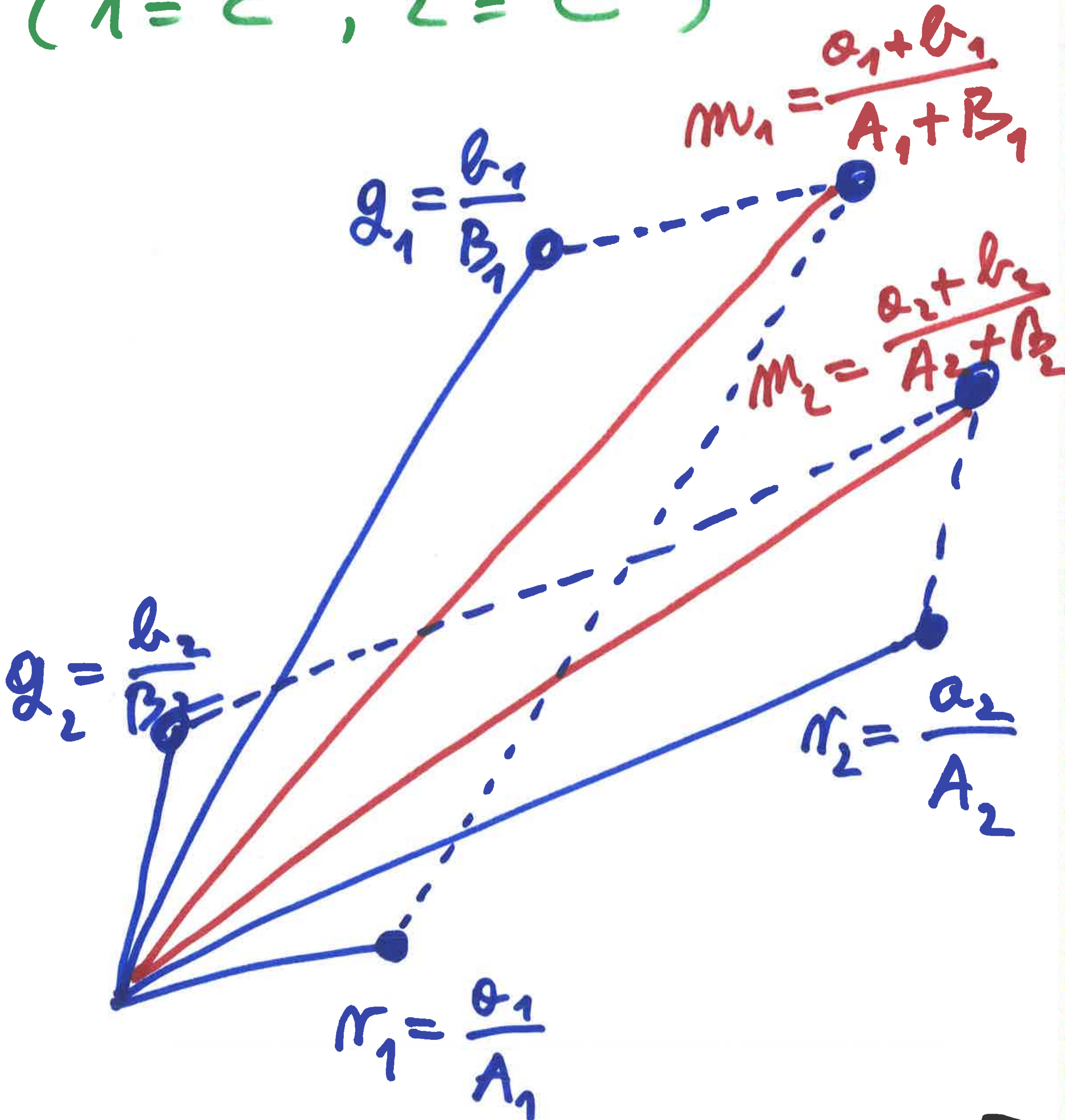
g = INCIDENCE OF SMTH IN B-SUBP.

m = INCIDENCE OF SMTH IN POPUL.

SOME FIND SURPRISING:

$r_1 < r_2$, $g_1 < g_2$ & $m_1 > m_2$

$(1 = c \;,\; 2 = \bar{c})$



D.H. FOWLER (A.H.E.S. 1991):

"THE GREEKS WHO REASONED IN PROPORTIONS MIGHT HAVE DIFFICULTY REASONING IN DECIMAL NOTATIONS; WE, IN TURN, SEEM TO HAVE DIFFICULTY REASONING IN PROPORTIONS. THIS MAY BE SO INGRAINED THAT WE MAY ACTUALLY FIND SOME OF THIS REASONING PARADOXICAL."

MATH OLYMPIAD 2009/10:

4 GLASSES ARE GIVEN WITH APPLE J., PEACH J., GRAPEFRUIT J. AND CARROT J.

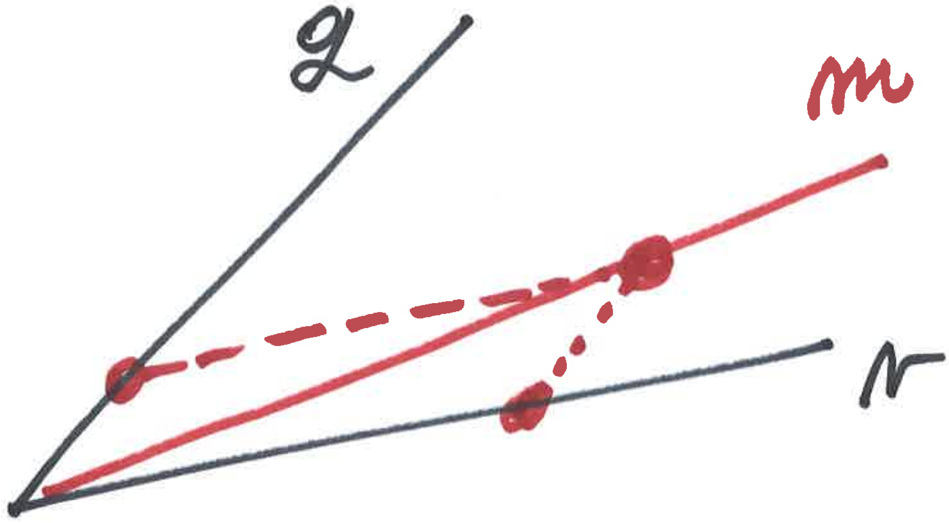
A IS SWEETER THAN G
P IS SWEETER THAN C

IS IT TRUE THAT:
THE MIXTURE OF A & P
IS SWEETER THEN
THE MIXTURE OF G & C?

ANSWER:

NOT NECESSARILY SO!!!

FOR NON GREEKS A REDUCTION OF MEDIANTS:



TO WEIGHTED AVERAGES:

COULD BE HELPFUL

CF. p. 15. VS. p. 13 !!

IT IS SIMPLE:

$$r = \frac{R'}{R} \ \& \ g = \frac{Q'}{Q} \implies m = \frac{R'+Q'}{R+Q}$$

$$\implies m = \frac{R'}{R+Q} + \frac{Q'}{R+Q} =$$

$$\frac{R'}{R} \underbrace{\frac{R}{R+Q}}_{\alpha} + \frac{Q'}{Q} \underbrace{\frac{Q}{R+Q}}_{\beta} = r\alpha + g\beta$$

$$r = \frac{R'}{R} = \frac{xR'}{xR} \qquad g = \frac{Q'}{Q} = \frac{yQ'}{yQ}$$

$$\implies R \in (0,\infty) \ Q \in (0,\infty)$$

i.e. $\alpha \in (0,1) \ \beta \in (0,1)$

IT IS SIMPLE:

$$r = \frac{xr}{x} < \frac{xr + yq}{\underbrace{x+y}_{m}} < \frac{yq}{y} = q$$

$$r < \underbrace{\alpha r + \beta q}_{w} < q$$

$$\boxed{\begin{array}{ll} r < \dfrac{r + zq}{1+z} < q & \quad z = \dfrac{y}{x} \\ & \\ r < \alpha r + (1-\alpha) q < q & \quad 1-\alpha = \beta \end{array}}$$

$$\boxed{\alpha = \frac{1}{1+z} \iff z = \frac{1-\alpha}{\alpha}}$$

$$\alpha \in (0,1) \iff z \in (\infty, 0)$$

BACK TO MATH OLYMPIAD:

A, P, G, C QUANTITIES OF JUICE

a, p, g, c QUANTITIES OF SUGAR

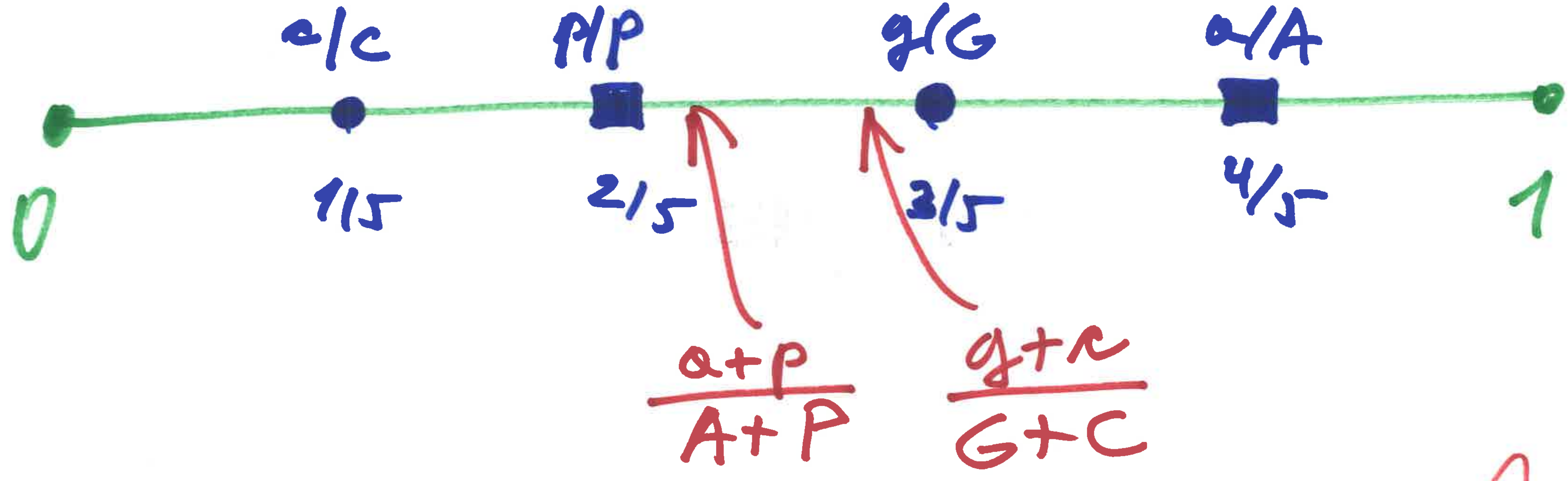
PROBLEM:

$$\frac{a}{A} > \frac{g}{G} \;\&\; \frac{p}{P} > \frac{c}{C} \overset{?}{\Rightarrow} \frac{a+p}{A+P} > \frac{g+c}{G+C}$$

WE KNOW:

$$\frac{a+p}{A+P} = \frac{a}{A}\,\frac{A}{A+P} + \frac{p}{P}\,\frac{P}{A+P}$$

$$\frac{g+c}{G+C} = \frac{g}{G}\,\frac{G}{G+C} + \frac{c}{C}\,\frac{C}{G+C}$$



$$\frac{a+p}{A+P} = \frac{4}{5}\frac{A}{A+P} + \frac{2}{5}\frac{P}{A+P} = \frac{4}{5}\frac{1}{1+19} + \frac{2}{5}\frac{19}{1+19} = \frac{42}{100} \approx \frac{40}{100} = \frac{p}{P}$$

$$\frac{g+c}{G+C} = \frac{3}{5}\frac{G}{G+C} + \frac{1}{5}\frac{C}{G+C} = \frac{3}{5}\frac{19}{1+19} + \frac{1}{5}\frac{1}{1+19} = \frac{58}{100} \approx \frac{60}{100} = \frac{g}{G}$$

$A, G = 1\ell \quad P = C = 19\ell$

19''

"LOGICAL S.T.P." IS VALID:

$$(CP \to E \;\&\; C\bar{P} \to E) \to (C \to E)$$

TAUTOLOGY; AN INSTANCE OF:

$$(Y \to X \;\&\; Z \to X) \to (Y \vee Z \to X)$$

$$(P \to (C \to E) \;\&\; \bar{P} \to (C \to E)) \to (P \vee \bar{P} \to (C \to E))$$

"DISJUNCTIVE SYLOGISM"

BUT S.T.P. IS NOT:

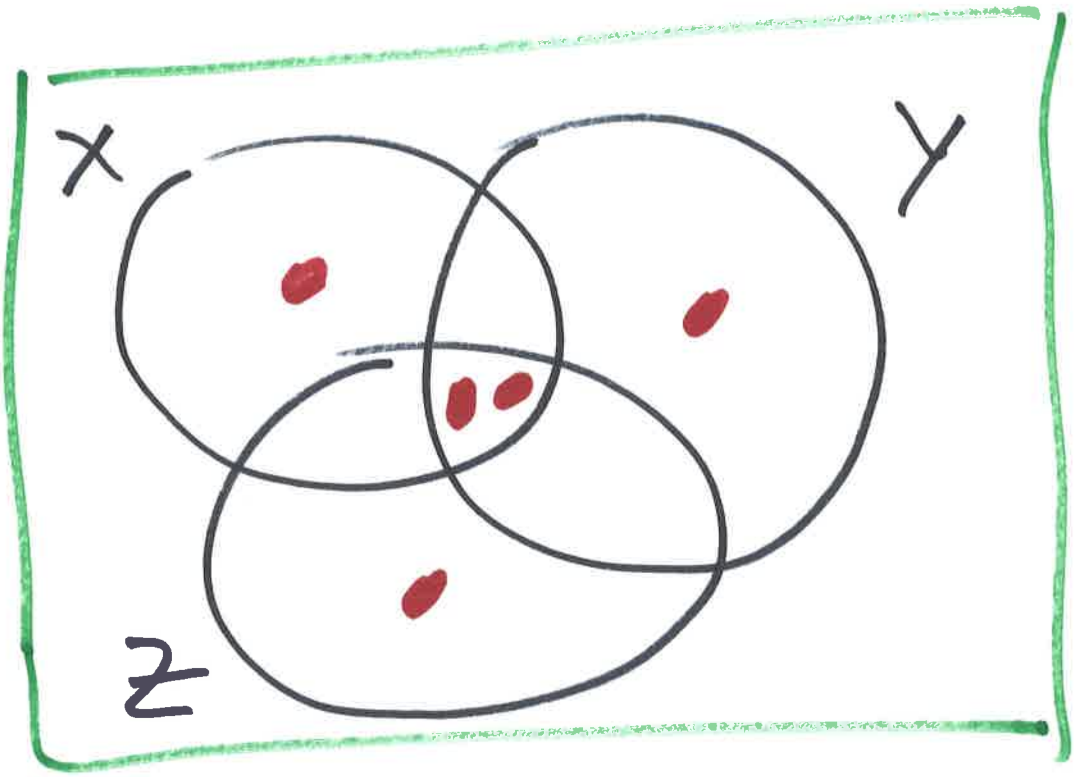
$$(CP \uparrow E \;\&\; C\bar{P} \uparrow E) \not\to (C \uparrow E)$$

WHICH IS AN INSTANCE OF:

$$(Y \uparrow X \;\&\; Z \uparrow X) \not\to (Y \vee Z \uparrow X)$$

$$\boxed{Y \uparrow X :\equiv p(X|Y) > p(X)}$$

A SIMPLE REFUTATION OF "$\uparrow$ DISJUNCTIVE SYLOGISM":



$$Y \uparrow X :\equiv p(X|Y) = \frac{2}{3} > \frac{3}{5} = p(X)$$

$$Z \uparrow X :\equiv p(X|Z) = \frac{2}{3} > \frac{3}{5} = p(X)$$

BUT

$$Y \vee Z \downarrow X :\equiv p(X|Y \vee Z) = \frac{1}{2} < \frac{3}{5} = p(X)$$

"ON PROBABLE CONDITIONALS"

Z. ŠIKIĆ   EUJAP 2016

ALMOST EVERY RULE VALID FOR $\rightarrow$ (EXPECT ONE) IS NOT VALID FOR $\uparrow$

VALID $\rightarrow$ DISJUNCTIVE SYLOGISM VS. NON VALID $\uparrow$ DISJ. SYLOG. IS AN EXAMPLE