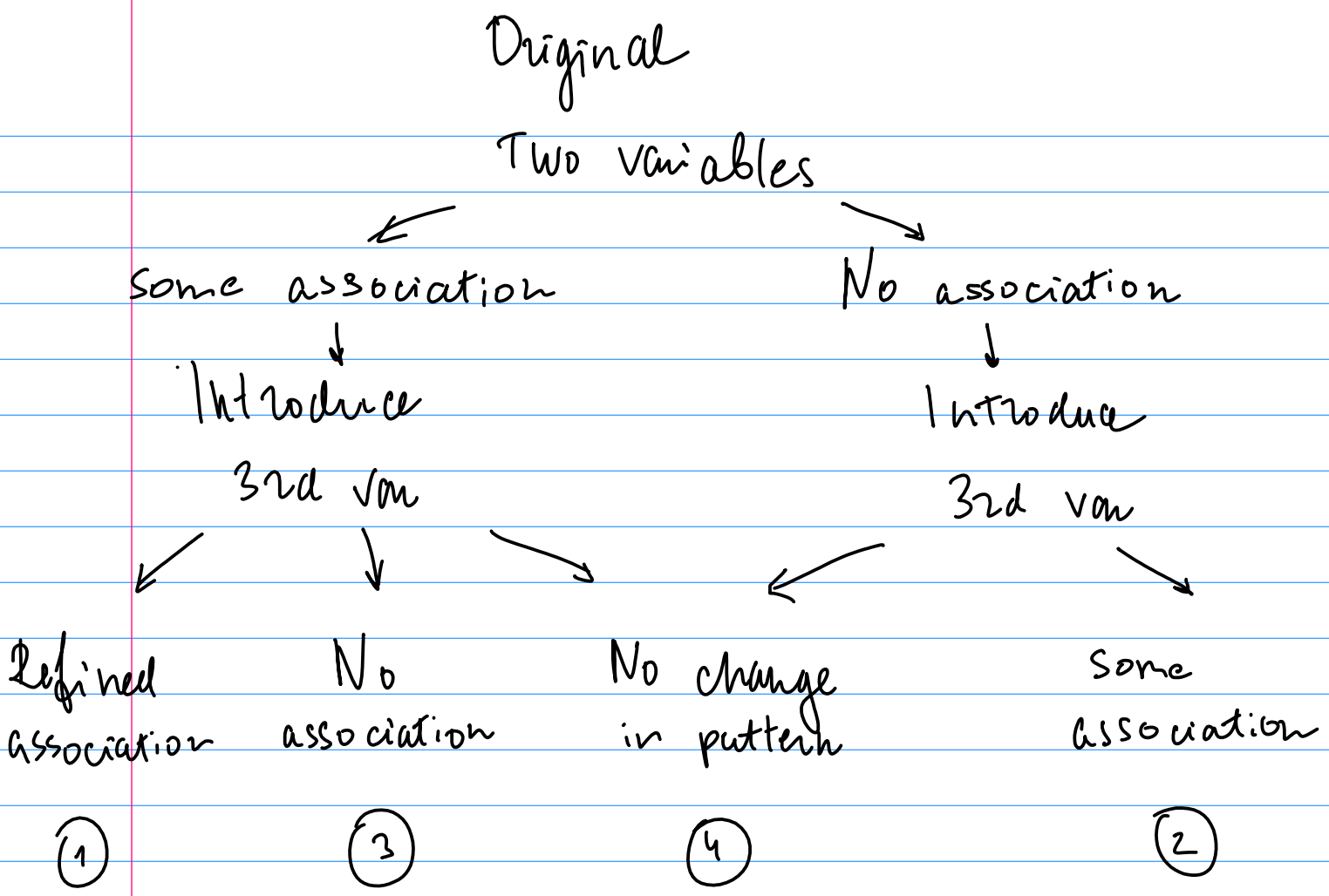
Contingency Tables

	M	F	
Light	5	10	15
Heavy	10	5	15
	15	15	

	M	F
L	33.3 %.	66.6 %
H	66,6 %.	33.3 %.


Original
Two variables
Some association
No association
Introduce 3rd var
Introduce 3rd var
Refined association
No association
No change in pattern
Some association
1
3
4
2

① Refined Association

	Man	Woman
H	31%	52%
L	69%	48%

∴

	M		F	
	M	U	M	U
H	35	40	25	60
L	65	60	75	40

② Spurious Association

% Owns exp. car \ Ed.	Degree	No degree
Yes	32	21
No	68	79

	Low-income ed.		High-income ed.	
	Degree	No degree	Degree	No degree
Yes	20 ↔	20	40 ↔	40
No	80 ↔	80	60 ↔	60

(3) Reveal suppressed association

%

Travel abroad desire	Under 45	Over 45
Yes	50	50
No	50	50

	Male		Female	
Desire	Under 45	Over 45	Under 45	Over 45
yes	60	40	35	65
No	40	60	65	35

(4)

Frequently fast-food	Small fam.	Large fam.
Yes	65	65
No	35	35

Frequently fast-food	Low income		High income	
	Small fam.	Large fam.	Small fam.	Large fam
Yes	65	65	65	65
No	35	35	35	35

Chi - Squared Test

H_0: no association
H_a: some association

$$\chi^2 = \sum_{i,j} \frac{(f_o - f_e)^2}{f_e} \overset{H_0}{\sim} \chi^2_{(r-1)\cdot(c-1)}$$

r — # row
c — # cols

$$f_e = \frac{n_r n_c}{n}$$

	M	F	
Light	5	10	15 $= n_r$
Heavy	10	5	15
	15 $= n_c$		$n = 30$

$$\chi^2 = 2 \cdot \left(5 - \frac{15 \cdot 15}{30}\right)^2 / 7.5 + 2 \cdot \left(10 - 7.5\right)^2 / 7.5 =$$

$$= 3.333$$

$$\chi^2_{crit,\ 0.95} = 3.841$$

Measure Strength of Association:

① Phi coefficient (2x2)

$$\phi = \sqrt{\frac{\chi^2}{n}} \quad \in [0, 1]$$

$$\phi = \sqrt{\frac{3,333}{30}} = 0,33$$

② Contingency coefficient (any size)

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} \quad \in [0, 1]$$

(*) max value of C depends on table size

③ Cramer's V (any size tables)

$$V = \sqrt{\frac{\phi^2}{\min\{r-1, c-1\}}} \quad \in [0, 1]$$

(4) Lambda coefficient

↳ ↑% prediction improved

Asymmetric (predicting dependend var based on ind. Var)

Symmetric (no assumption on direction of dependency)

(5) Statistics for ordeal data

T_b } non-parametric coef.

T_c } for ordeal data

γ } [-1 ; 1]