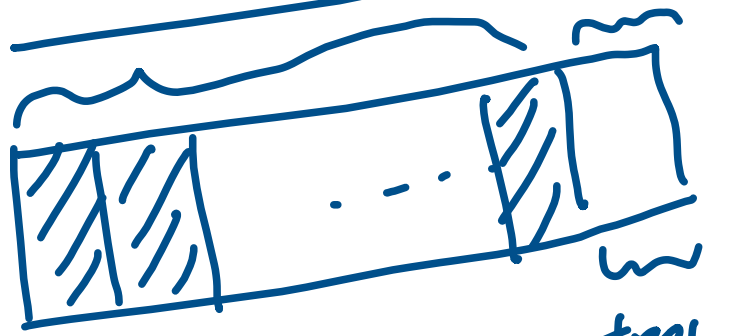
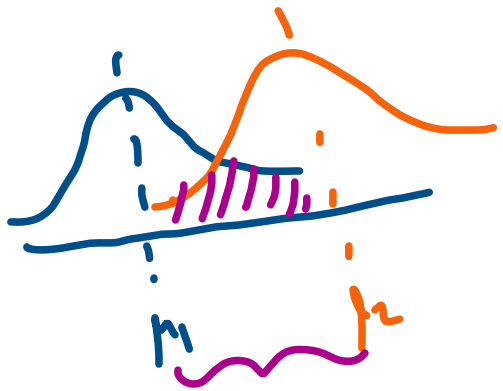
A_1	A_2
72	62
75	76
69	63
⋮	⋮

μ : ~ ~

Train-test splits.

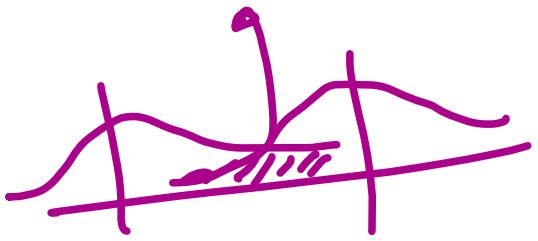


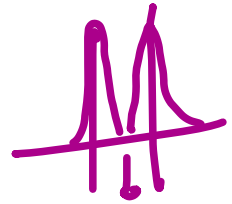
tol trip.



Sep. between means.

area / overlap.





{piece, peace} {desert, desiert}

$c_{-k} \quad \cdots \quad c_{-1} \quad w \quad c_1 \quad \cdots \quad c_k$

$P(w \mid c_{-k} \ldots c_{-1}, c_1, \ldots, c_k)$

$\propto P(c_{-k} \ldots c_{-1}, c_1, \ldots, c_k \mid w) \cdot P(w)$

$\prod_{i \in \{-k, \ldots, -1, 1, \ldots k\}} P(c_i \mid w)$

context words

collocation

$P(s_1, s_2 \ldots, s_n, d)$

$= P(s_1 | s_2 \ldots, d) \cdot P(s_2 | s_3 \ldots, d) \ldots P(s_n | d) \cdot P(d)$

$P(s_1 | d)$ $P(s_2 | d)$.

$P(s_1, \ldots s_n | d) = \prod_{i=1}^{n} P(s_i | d)$

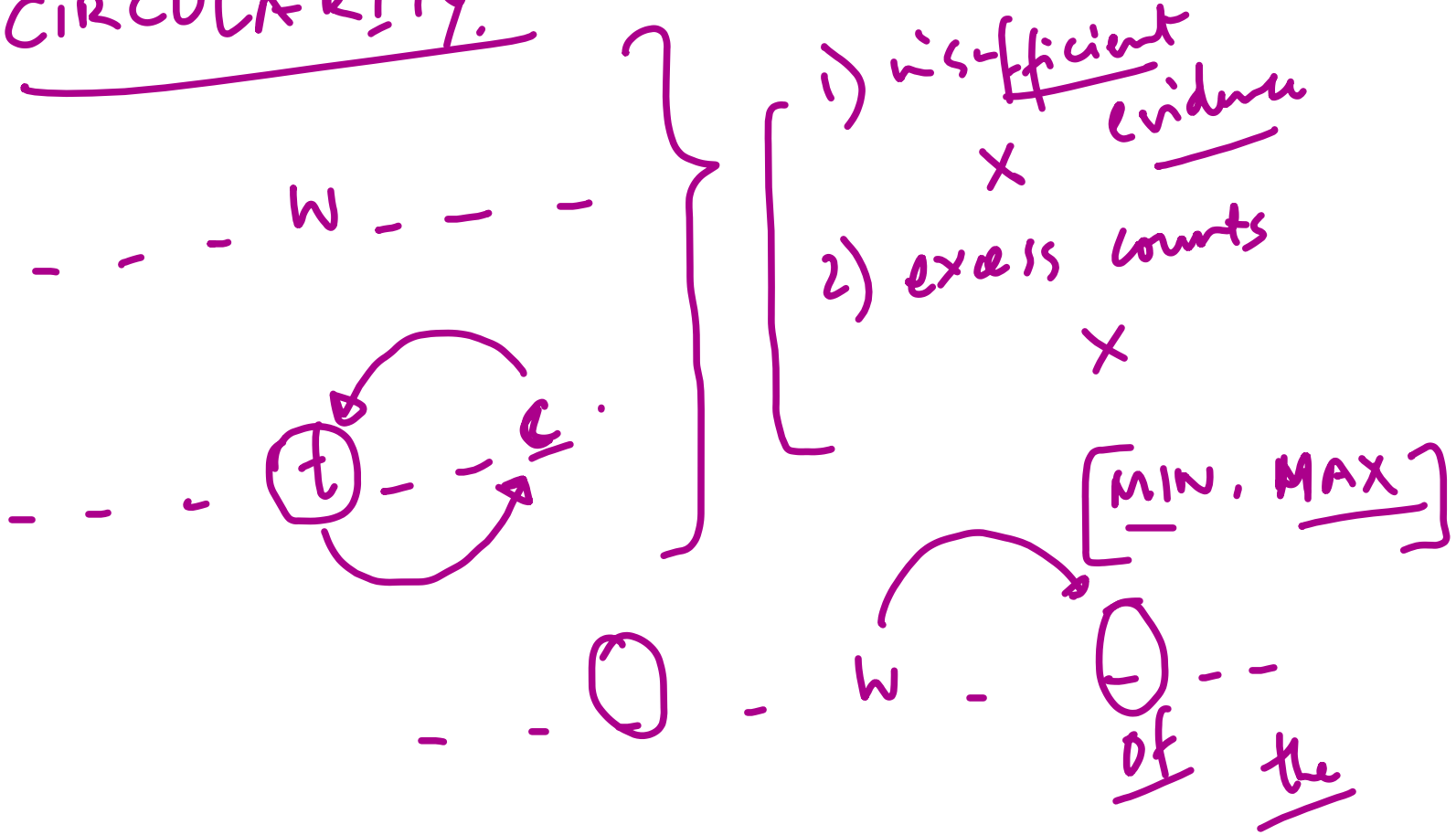
NAIVE BAYES ASSUMPTION

[features are independent] X [CONDITIONALLY INDEPENDENT GIVEN CLASS]

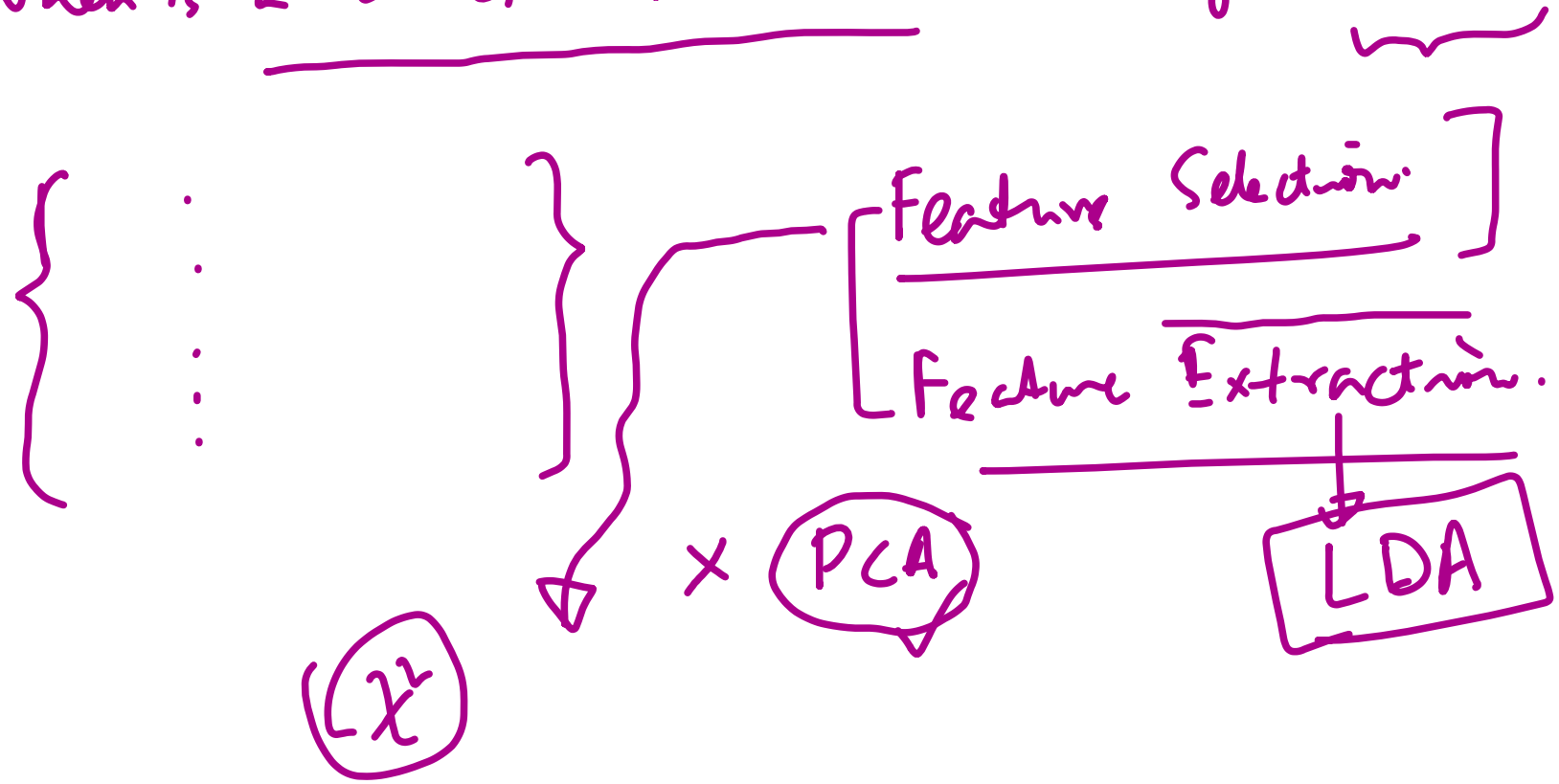
fly → cold, fever, thunderstorm

place

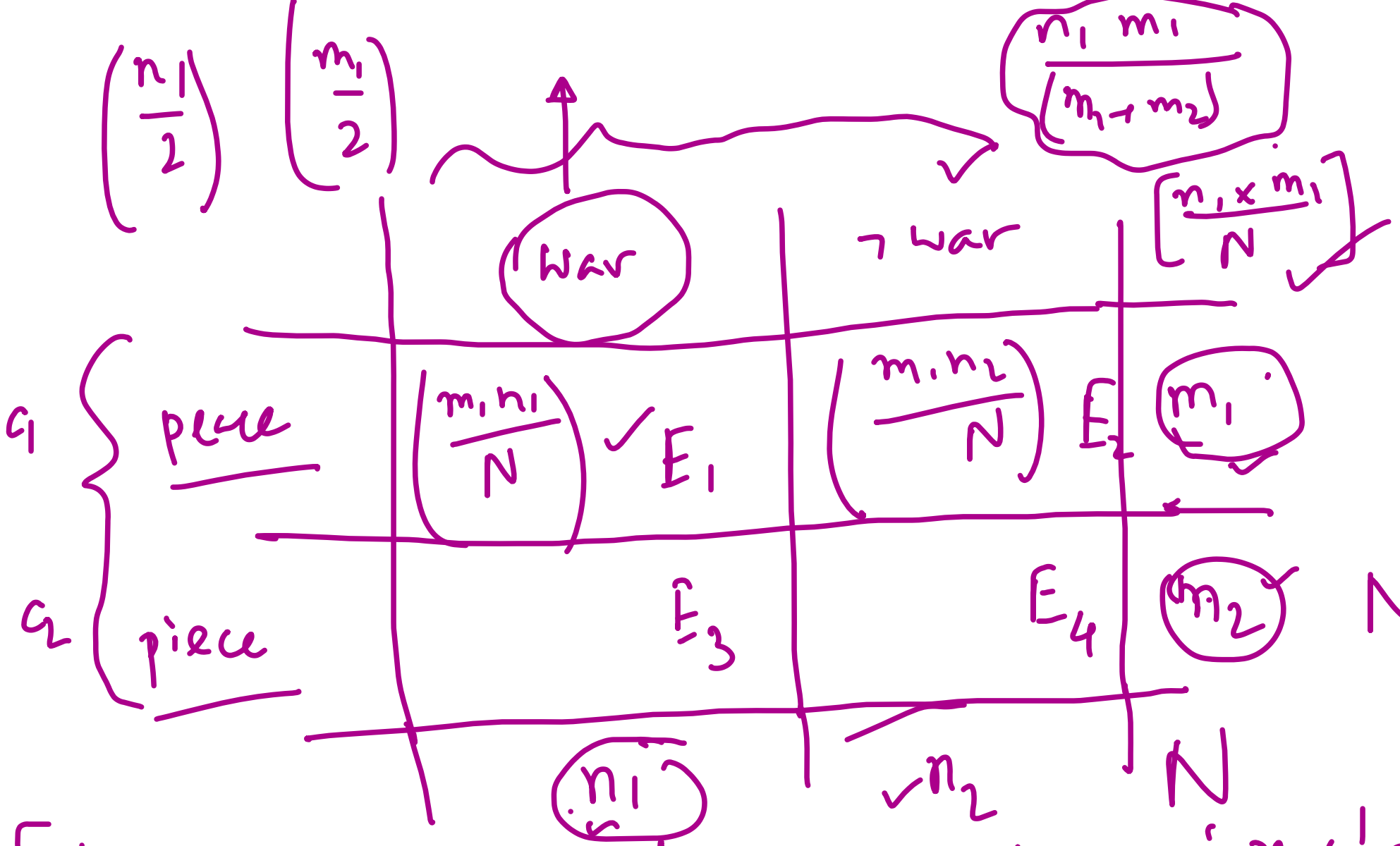
CIRCULARITY.



When is a context word discriminating?



Chi-squared test.

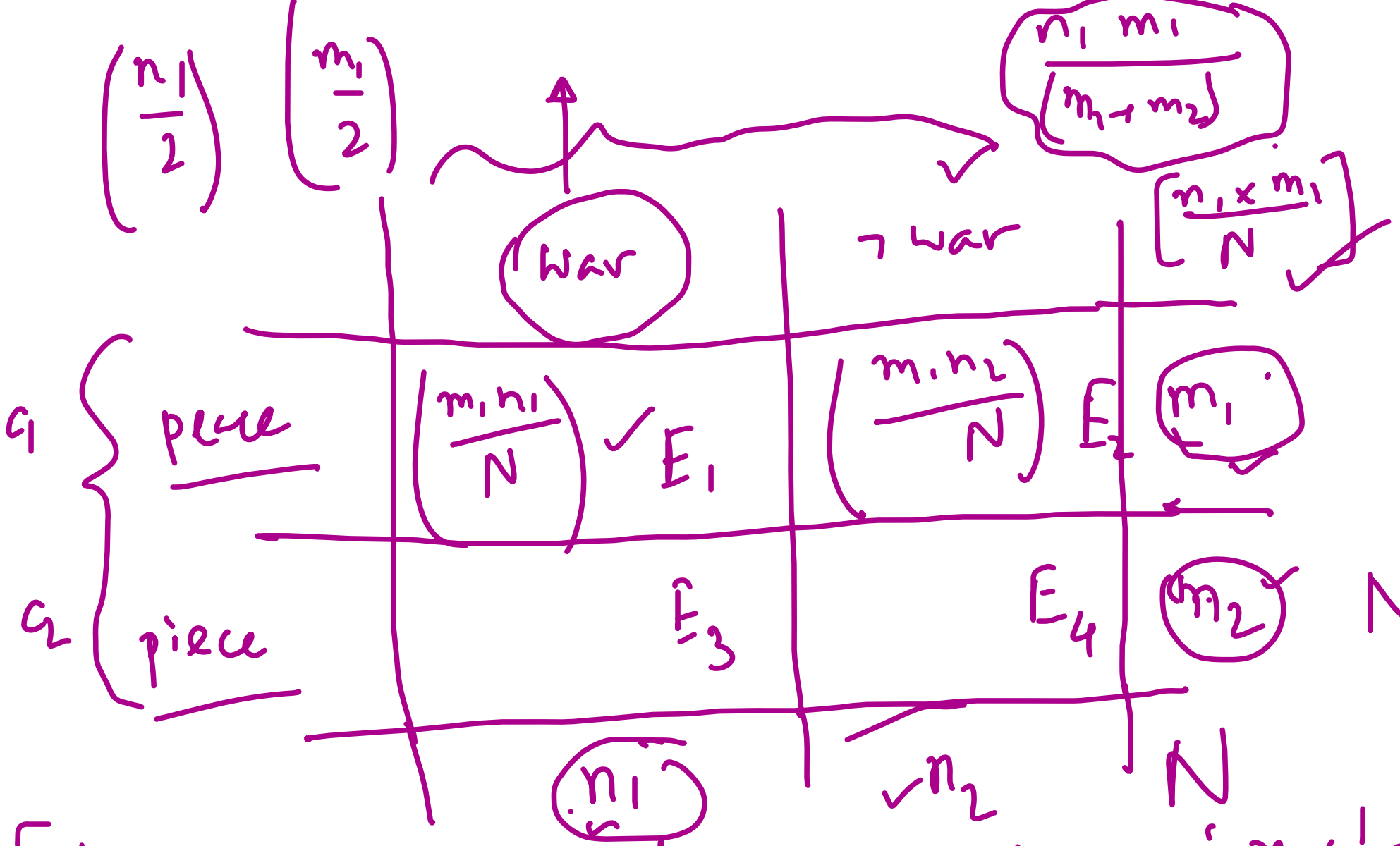


$N = m_1 + m_2$

$= n_1 + n_2$

H_0: War is not discriminating between 'piece' & 'piece'

H_1: " " discriminating " " " "

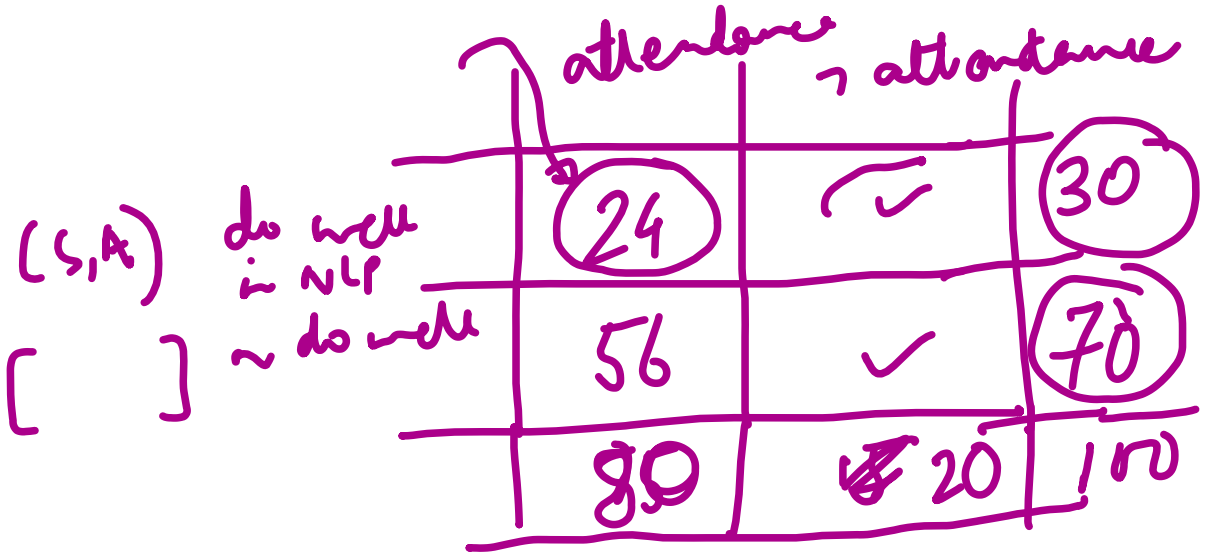


$N = m_1 + m_2$

$= n_1 + n_2$

H_0: War is not discriminating between 'piece' & 'piece'

H_1: " " discriminating " " " "



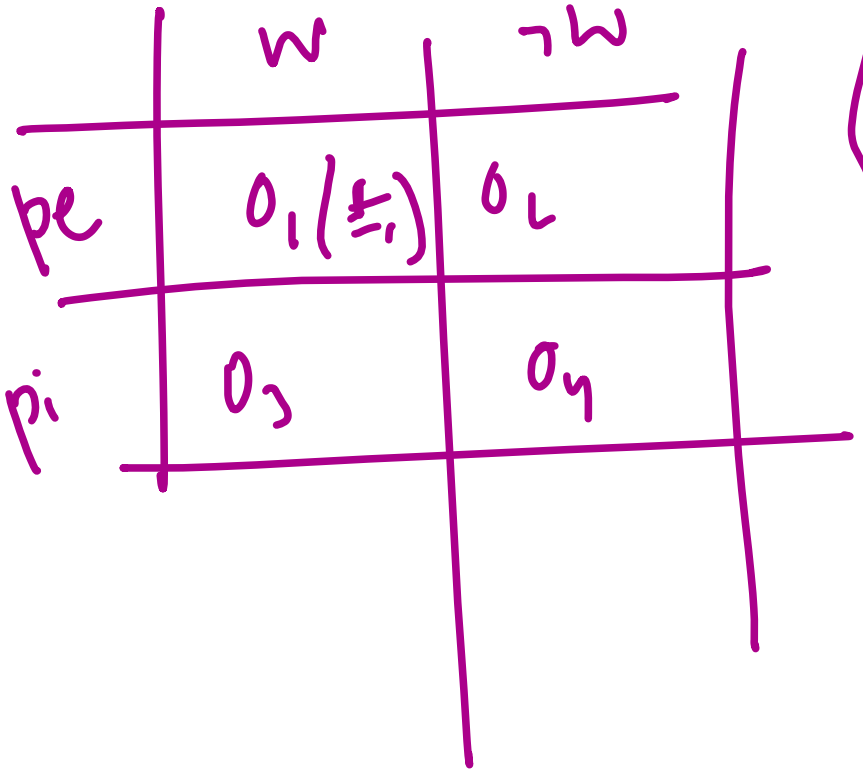
$$\sum \frac{(O_i - E_i)^2}{E_i}$$

d.f. = 1

degree of freedom.

$(m \times n)$

$(m-1)(n-1)$



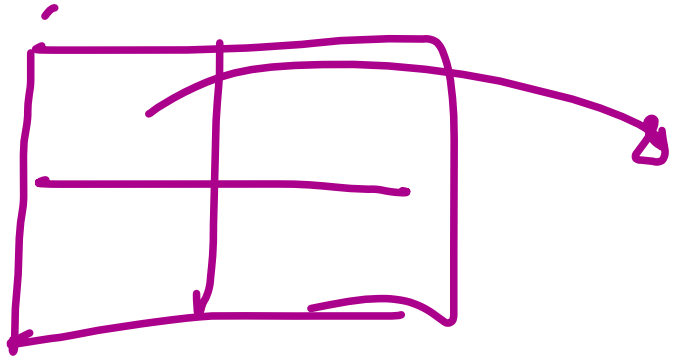
$$\sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

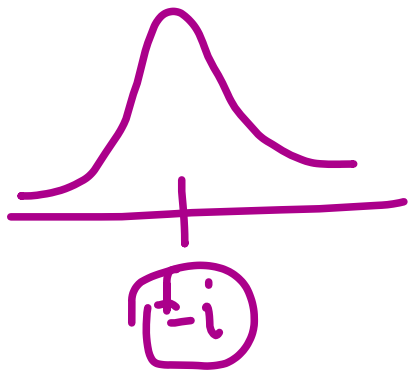
$= 4$ → Statistic

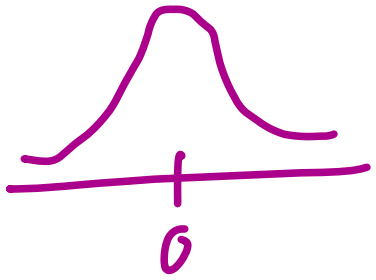
Except 3.4

hybrid

Collocation.





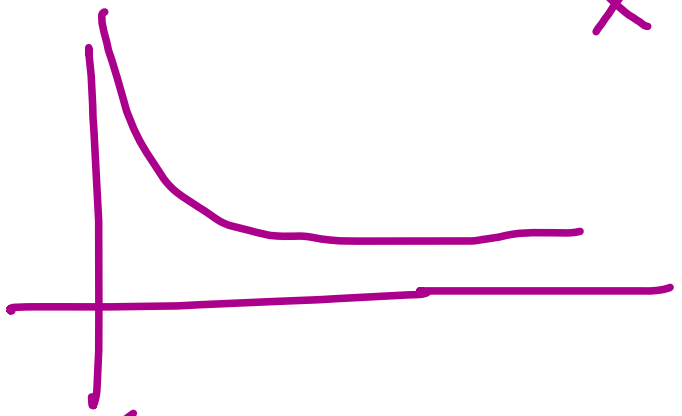


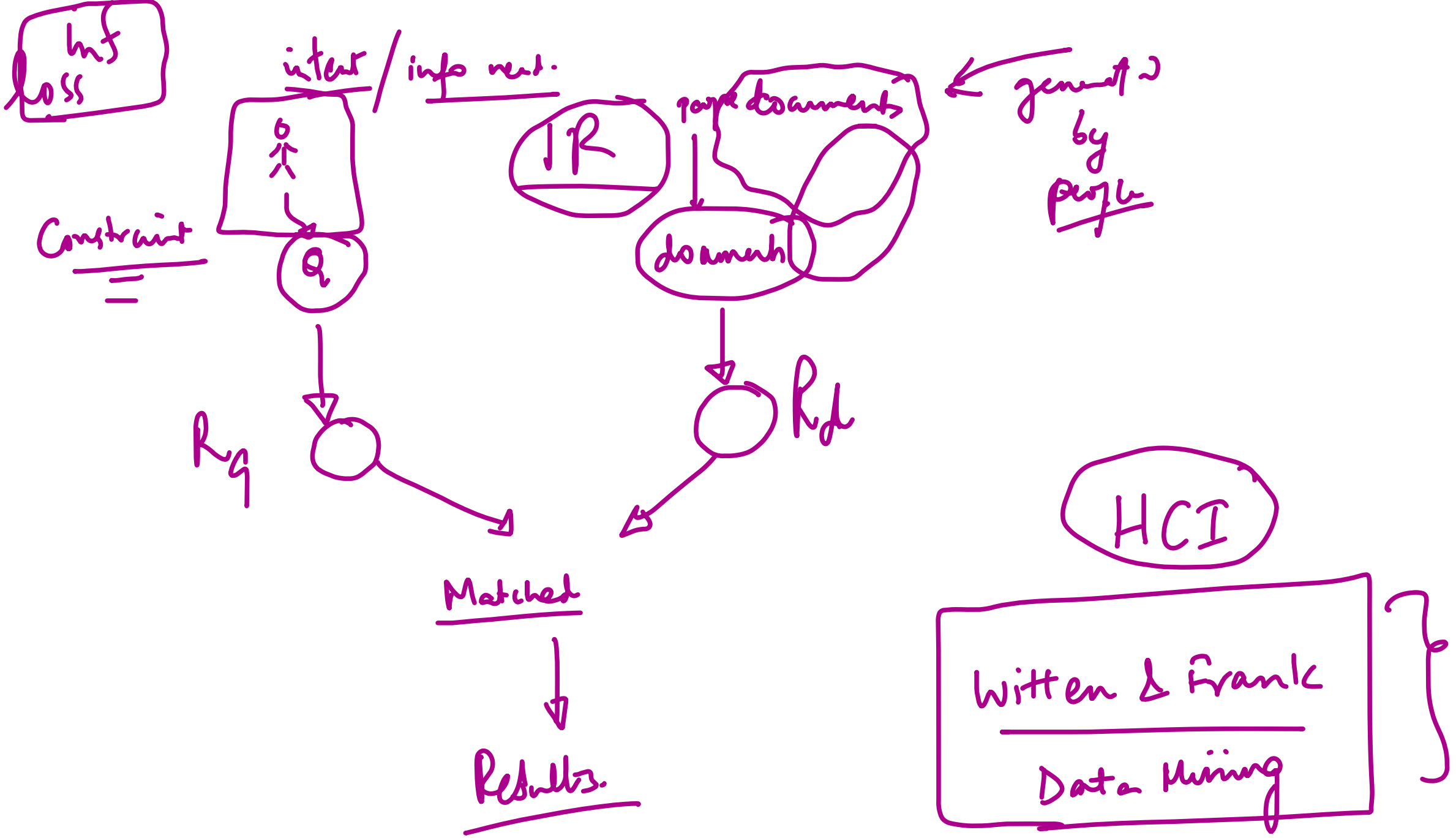
$X \sim N(0,1)$

χ^2 $\xrightarrow{1}$ X^2

$\chi^2 \rightarrow X_1^2 + X_2^2$

$\chi^2 \rightarrow X_1^2 + X_2^2 + X_3^2$




Inf
loss
intent / info req.
Constraint
Q
IR
page
documents
generated
by
people
document
R_q
R_d
Matched
Results.
HCI
Witten & Frank
Data Mining