I EXAMPLE - TEST OF INDEPENDENCE

NULL HYPOTHESIS

$$P(\text{eye colour and hair colour}) = P(\text{eye colour})\,P(\text{hair colour})$$

$$P(A \cap B) = P(A)P(B) \leftarrow \text{independence}$$

$$\Rightarrow P(\text{brown eyes}) = \frac{31}{227}$$

$$P(\text{black hair}) = \frac{42}{227}$$

If the null hypothesis was correct,

$$P(\text{brown eyes + black hair}) = \frac{31}{227} \times \frac{42}{227}$$

and the expected frequency of brown eyes and black hair is

$$\boxed{n = 227}$$

$$\frac{31}{227} \times \frac{42}{227} \times 227 = 5.74$$

$$E_k = \frac{\text{column total}}{n} \times \frac{\text{row total}}{\cancel{n}} \cancel{\times n} \qquad X^2 \sim \chi^2_4$$

$$X^2 = \frac{(10-5.74)^2}{5.74} + \frac{(24-17.95)^2}{17.95} + \frac{(8-18.32)^2}{18.32} + \dots = 34.9$$

II EXAMPLE

the row totals are fixed, under the null hypothesis.

$$\frac{E_K}{\text{row total}} = \frac{\text{column total}}{n}$$

or

$$E_k = \frac{\text{column total} \times \text{row total}}{n}$$

$$E_1 = \frac{50 \times 40}{90} = 22.22$$

$$E_2 = \frac{50 \times 30}{90} = 16.67$$

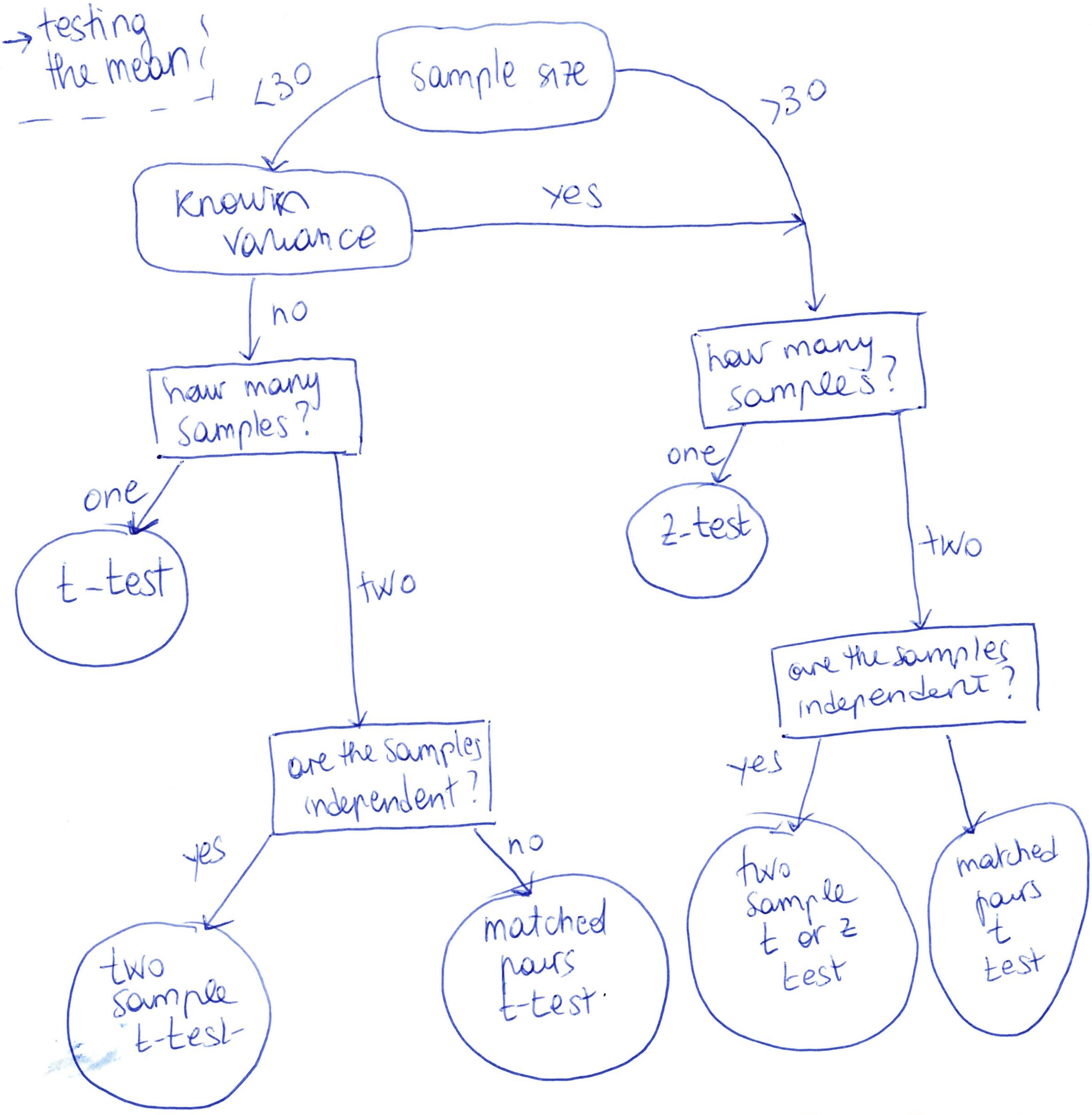
...

$$X^2 = \frac{(16 - 22.22)^2}{22.22} + \frac{(20 - 16.67)^2}{16.67} + \ldots = 7.11$$

$$X^2 \sim \chi^2_{(2-1)(3-1)} \equiv \chi^2_2$$

continuous and normally distributed data

→ testing the mean



→ testing the variance

how many samples?

→ one → chi square test

→ two → F test

→ testing the correlation → z test

discrete data / proportions

to test proportions for two samples → z test

goodness of fit / independence / similarity test

→ chi squared test