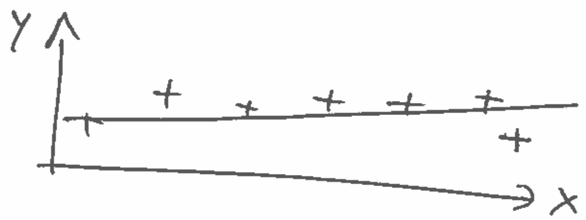
Decision making = Hypothesis testing

Barlow, chapter 8.1

yes - or - no situation

↳ rather than measuring parameters you ask a question about their behaviour, e.g. does y increase with x?



need:

- assertion that some hypothesis is true
- numerical test to be applied to data
- accept / reject hypothesis depending on outcome

examples

- ∟ interpretation of experiments
- ∟ goodness of fit
- ∟ two-sample problem
- ∟ analyses for several samples

Any test won't be infallible

⇒ need to choose a level of confidence at which to take the decision

Hypotheses

Example: - These data are drawn from a Poisson distribution of mean 3.4

- The new treatment has identical effects to the old.

↳ simple hypotheses: define probability distribution completely

- These data are drawn from a Poisson distribution of mean greater than 4.
- The new treatment is an improvement on the old.

↳ composite hypotheses: combine several prob. distr. functions

Alternative to hypothesis (compare to the first above)

- Poisson with another given mean, e.g. 5
- Poisson with any random mean other than 3.4
- Not Poisson

↳ needs to be spelt out

Generally, it is crucial to distinguish between one-tailed directional and two-tailed non-directional tests

Type I, II errors

hypothesis \ decision	accept	reject
true	☺	Type I error
false	Type II error	✓

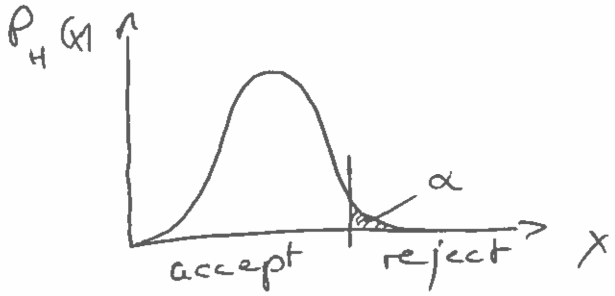
Law: not sentencing a guilty person: Type I error
sentencing a innocent person: Type II error

Significance

Type I errors are inevitable and the rate at which they occur is called significance. We know the probability distribution function of the hypothesis, $P_H(x)$, for the case that the test involves the determination of x.

e.g. Poisson with mean 3.4

Divide $P_H(x)$ in acceptance and rejection region



Take decision depending on where measured value of x falls.

The significance α is given by the integral

$$\alpha = \int_{Reject} P_H(x)\, dx$$

Typically, want to choose α to be small, eg. 1% or 5%.
In reality, we often need to work with the inequality

$$\alpha \geq \int_R P_H(x)\, dx$$

as we may have a range of possible $P_H(x)$ (composite) or a discrete distribution for which α cannot be reached exactly.

For composite hypotheses we would use the $P_H(x)$ that maximises α.
All other incarnations of the composite hypothesis would therefore result in a smaller integral and be equally accepted.

<u>Example</u>: Hyp.: Poisson with mean ≤ 5

Want significance 5%

$$\sum_{x=0}^{9} P_H(x; \lambda=5) = 96.8\%$$

$\Rightarrow$ accept $x = 9$ or smaller

for all $\lambda < 5$ $\sum_{x=0}^{9} P_H(x; \lambda) > 96.8\%$

$\Rightarrow$ $\lambda = 5$ maximises the integral

Power

Consider alternative hypothesis for which the prob. distr. function is known exactly: $P_A(x)$

With this we can calculate the rate of Type II errors which is the integral of $P_A(x)$ over the accept region

$$\beta = \int_{\text{Accept}} P_A(x)\,dx$$

With this, $1-\beta$ is called the power of the test.