Hypothesis testing

Suppose we see a stream of data $X_1 \ldots \ldots X_n$.

We have two computing models for the data.

<u>The "Null" model</u> :- Says that the data is coming from some probability distribution P_0.

<u>The "alternative" model</u> :- Says that the data is coming from another distribution P_1.

<u>Test</u> :- $T : \Omega \to \{0, 1\}$: a fn. from observations to "models".

Some terms :-

"Type 1 error" : 'Rejecting' the null when the 'alternative' is false

"Type 2 error" : Failing to 'reject' the null when the 'alternative' is true.

Significance level of a test :

The probability, when the null is true, of rejecting the null

Power of a test

The probability, when the alternative is true, of rejecting the null.

(Q):- Which tests are 'optimal'?

A possible solution:- Fix a significance level α. What is the highest **'power'** that one can achieve?

Neyman - Pearson lemma

Consider a null rejecting region of the form:

$$R_T(\eta) = \{ x \mid P_1(x) \geq \eta P_0(x) \}.$$

for some $\eta \geq 0$ real.

Let α be the significance level of $R_T(\eta)$. The power of any test with significance level at most α is at most that of $R_T(\eta)$.

Proof:- Let R be any other region satisfying $P_0(R) \leq \alpha$.

$$R_1 := R \cap R_T(\eta)$$

$$R_2 := R \setminus R_1.$$

Now

$$\text{Power}(R_T(\eta)) = P_1(R_1) + P_1(R_T(\eta) - R_1)$$

$$\text{Power}(R) = P_1(R_1) + P_1(R_2)$$

So,

$$\text{Power } R_T(\eta) - P_1(R_1) + \eta P_0(R_1)$$

$$= P_1(R_T(\eta)) + \eta P_0(R) \geq \eta P_0(R_T(\eta) - R_1) + \eta P_0(R_1)$$

$$= \eta P_0(R_T(\eta))$$

$$\mathrm{Power}(R) - P_1(R_1) + \eta P_0(R_1)$$

$$= P_1(R_2) + \eta P_0(R_1) \leqslant \eta P_0(R_2) + \eta P_0(R_1)$$

$$= \eta P_0(R)$$

$$\Rightarrow \mathrm{Power}(R_T(\eta)) - \mathrm{Power}(R)$$

$$\geqslant \eta\left(P_0 R_T(\eta) - P_0(R)\right). \quad \blacksquare$$