P_1 versus P_0 plots

Luc Demortier (Rockefeller)

Louis Lyons (Imperial College & Oxford)

Statistical Issues in Searches SLAC, June 2012

These plots are used for demonstrating:

* Exclusion / Discovery / No-Decision regions

* Contours for fixed pdf separation under H_0 and H_1

* CL_S exclusion

* Punzi's sensitivity criterion

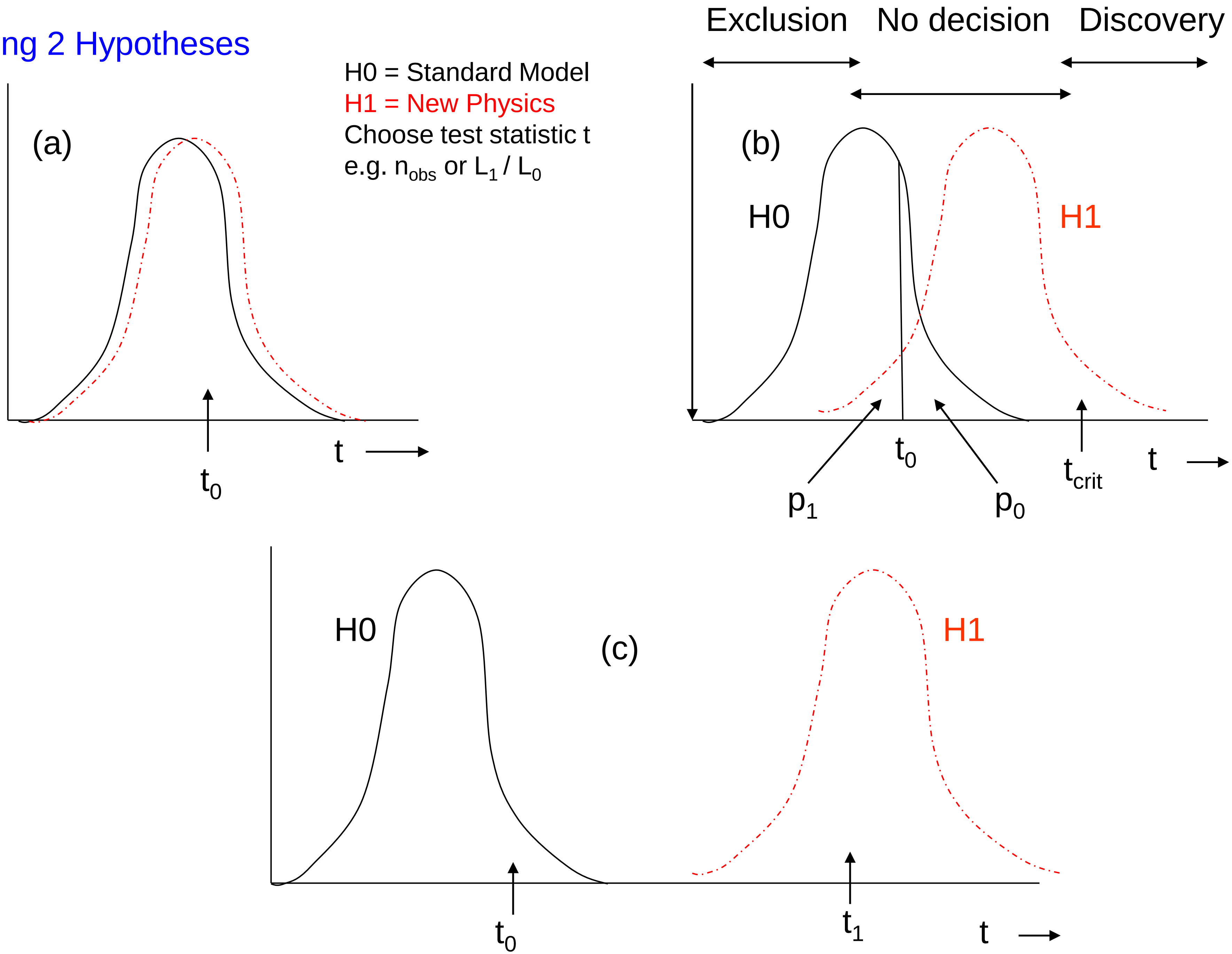
* Type-II error probabilities

* Likelihood ratios versus p-values

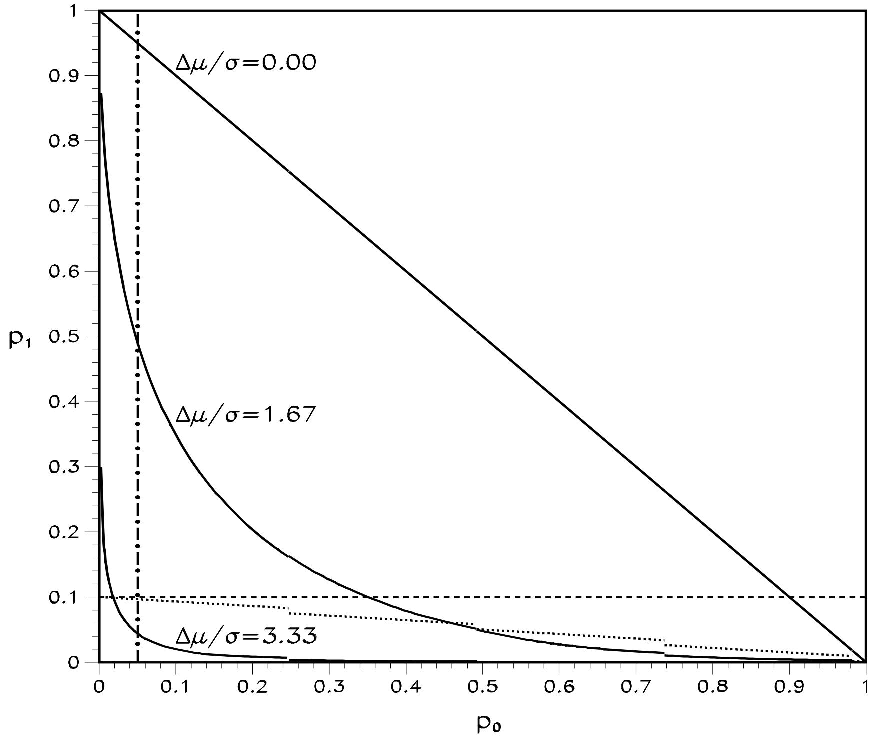
* A simplified version of Lindley's paradox

Testing 2 Hypotheses

H0 = Standard Model
H1 = New Physics
Choose test statistic t
e.g. n_{obs} or L_1 / L_0



P_1 versus P_0 plots



Contour = (p_0, p_1) values as t_{obs} varies, for Gaussian pdf's of **given** separation

Diagonal is limit of possible values

$p_0 = 0.05$ is for rejection of H_0 ("Discovery"). More realistically 3×10^{-7}

$p_1 = 0.1$ is for exclusion of H_1 ("Exclusion"). More realistically 0.05

Big rectangle is "No Decision"

Small rectangle is "Reject Both" (Discovery and exclusion of H_1?)

CL_S

$CL_S \equiv p_1 \,/\, (1-p_0)$

CL_S =0.1 exclusion line shown in Fig. 2

Protects against rejection of H_1 when little/no sensitivity (Fig. 1a)

Not liked by most statisticians (ratio of 2 p-values)

Punzi Sensitivity Criterion

Have just enough data so that H_0 and H_1 pdf's are well separated, and contour of Fig. 2 never enters "No decision" region

i.e. Always able to claim "Discovery" or "Exclusion"

N.B. Very conservative w.r.t. exclusion

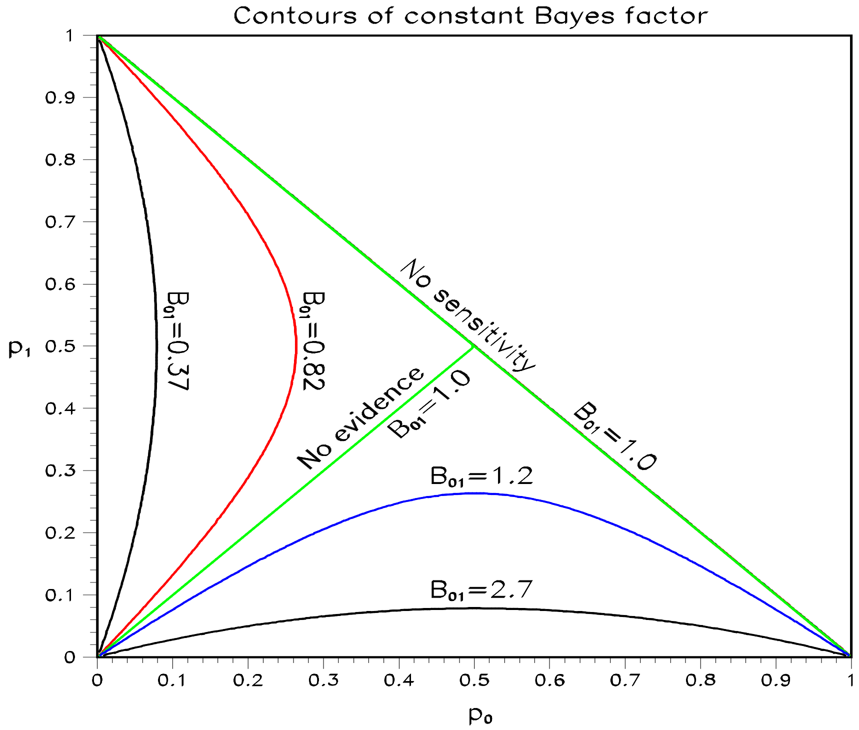
Type II Error Probabilities

Under H_1, expect uniform distribution of p_1 (modulo discreteness effects, treatment of nuisance parameters).

Hence, coordinate differences along y-axis correspond to probabilities under H_1.

In particular, the ordinate of the intersection of a given contour with the line $p_0=\alpha$ equals the Type-II error probability β for the corresponding pdf's.

L-ratio (Bayes factor) contours



For given pdf's, $B_{01} \equiv L_0 / L_1$ defined at each (p_0, p_1)

Contours shown for Gaussian pdf's

Diagonal through origin has $L_0 / L_1 = 1$

(Data halfway between pdf's)

Other diagonal also has $L_0 / L_1 = 1$

(H_0 and H_1 pdf's coincide – compare Fig. 1a)

Demonstrates that constant p_0 does not determine L_0 / L_1 (obvious reasons why not).

Minimum of L_0 / L_1 at constant p_0 visible (cf. Berger).

Useful for understanding why p_0 cut-off should depend on amount of data (related to Lindley's paradox, but no prior needed here) .

Poisson example:

H_0 = Poisson, $\mu_0 = 1$ H_1 = Poisson, $\mu_1 = 10$, $n_{obs} = 10$

$p_0 = 1.1\ 10^{-7}$ $L_0 / L_1 = 8\ 10^{-7}$ **Strongly favours H_1**

Now have 10 times as much data. Basic physics hypotheses stay same, but expected rates increase by factor of 10

H_0 = Poisson, $\mu_0 = 10$ H_1 = Poisson, $\mu_1 = 100$, $n_{obs} = 31$

$p_0 = 0.8\ 10^{-7}$ $L_0 / L_1 = 1.2\ 10^{8}$ **Strongly favours H_0**