Hypothesis Testing
(a very brief introduction)

Phys 310
Bucknell University

March 21, 2017

Parameter Estimation

So far in this course, we have used curve fitting and χ^2 minimization to estimate parameters.We worked on the assumption that the model was correct, and that what was to be determined was the parameter(s).

Hypothesis testing

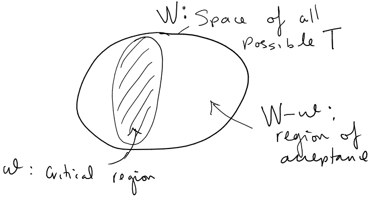
- Data can not prove a theory or hypothesis... because there may be other data that can contradict the theory
- However, data can be used to reject a theory if there is a contradiction to what may be expected.
- We can therefore use data to test hypotheses.

Basic Concepts

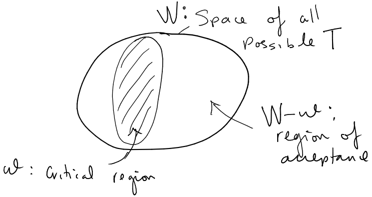
H_0: The Null Hypothesis (The model being tested)

H_1: The alternative hypothesis

t: test-statistic, a function of the data (mean, st. dev., χ^2 ...)



Types of Errors



$P(T \in \omega | H_0) = \alpha$
Type I Errors: T in critical region, given H_0 is true
$\alpha =$ probability of Type I error

$P(T \in W - \omega | H_1) = \beta$
Type II Errors: T in acceptable region, though H_0 is false
$\beta =$ probability of Type II error

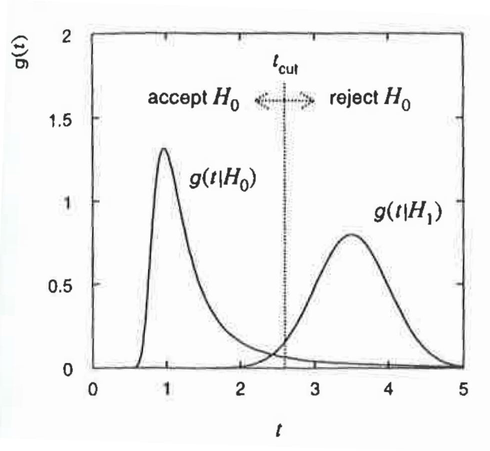


Figure: Probability densities for the test statistic t under the assumption of the hypothesis H_0 and H_1. H_0 is rejected if t is observed in the critical region, here shown as $t > t_{cut}$.(from Cowan: Statistical data analysis, 1998)

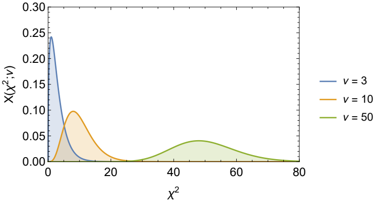
Goodness of Fit

H_0: data y_i (the sample distribution) is well modelled by a particular function $y(x_i)$ (The parent distribution).

T: Test statistic $= \chi^2$

We know the χ^2 pdf given ν the number of degrees of freedom:

$$X(\chi^2;\nu) = \frac{(\chi^2)^{\left(\frac{\nu}{2}-1\right)}\exp[-\chi^2/2]}{2^{\nu/2}\,\Gamma(\nu/2)}$$



Goodness of Fit

In Mathematica, the command

$$PDF[ChiSquareDistribution[\nu], x]$$

calculates $X(\chi^2; \nu)$
Likewise, to get

$$P(0 \leq \chi^2 \leq \chi^2_{min}; \nu) = \int_0^{\chi^2_{min}} X(\chi^2; \nu) d\chi^2$$

we enter

$$CDF[ChiSquareDistribution[\nu], xmin]$$

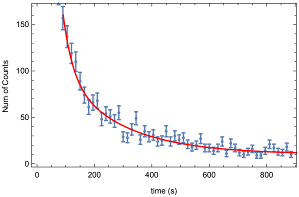
Exercise1

What is the most probable value of χ^2, given ν degrees of freedom for:

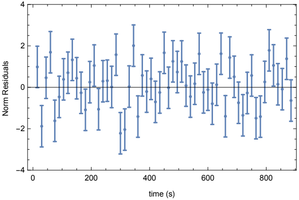
- $\nu = 5$
- $\nu = 10$
- $\nu = 30$

Exercise2

Consider a fit to a data shown below:



Model: y(x) =	Bg+N1∗exp[−x/T1]+N2∗exp[−x/T2]
χ^2	66.1
No. of data points	59



	Estimate	Standard Error	t- Statistic	P- Value
Bg	10.1358	2.09963	4.82742	0.0000117718
N1	128.302	23.4401	5.47361	1.1724×10^{-6}
N2	957.788	54.7796	17.4844	6.78919×10^{-24}
T1	209.657	35.1231	5.96921	1.91063×10^{-7}
T2	34.242	2.788	12.2819	2.90782×10^{-17}

Exercise2 (cont)

- Is the data consistent with the model at the 95 % confidence level. (i.e. Can we reject the hypothesis with Type I error $< 5\%$?)

- If not, for what value of χ^2 can we conclude that the data is not consistent with the model at the 95 % CL?