

Ordering of Trials

Jim Linnemann

MSU

Banff Birs Statistics

July 14, 2010

Look Elsewhere

= “Trials Effects”

= “Multiple Comparisons” “Multiple Testing”

use this to talk to statisticians

Multiple discoveries possible?

False Discovery Rate

Look Elsewhere Problems

Worst: *unknown* number of trials

- blind analysis
- tuning sample, then freeze cuts

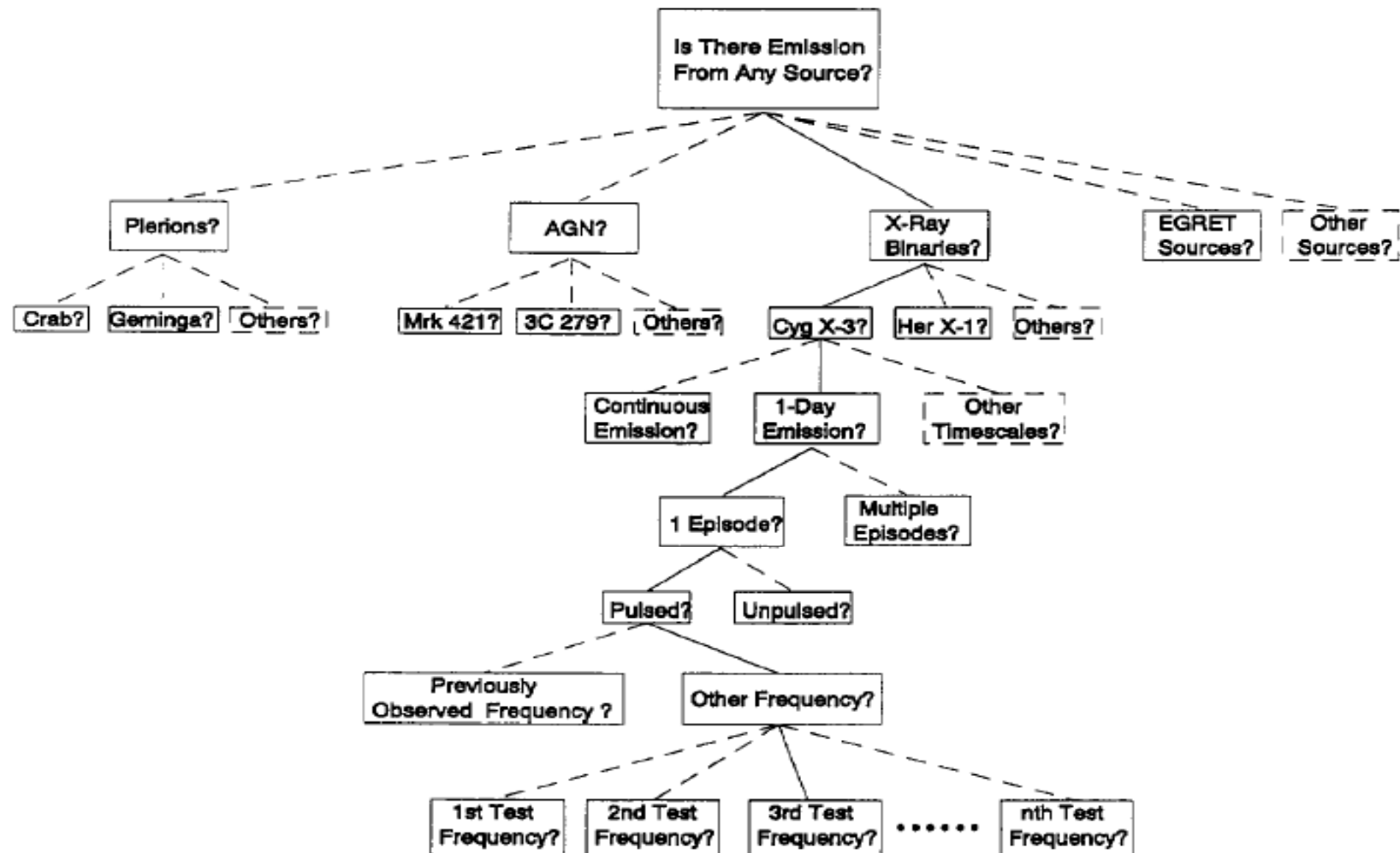
Next worst:

loss of power due to large N_{trials}
eg, lots of places to look on sky

How to “spend” trials

Importance ordering: write out a **protocol**

S.D. Biller / Astroparticle Physics 4 (1996) 285–291



How to order?

Your (collaboration's) choice:

Physics interest

we don't apply trials to top, Higgs, Z'

Prior probability

MC: expected sensitivity

Result of Ordering

First hypothesis:	1 trial (best sensitivity)
2 nd	2
..	
Nth	N (full Bonferroni penalty)

On average, $\frac{1}{2}$ the trials

Only “last” searches pay the full price

Trials Degrade Apparent Significance

Nominal significance must be de-rated by trials

$p_{\text{corrected}} \sim N p_{\text{nominal}}$ if you looked at N plots

$$\int_{\sigma}^{\infty} \text{Gau}(0,1) = N \int_{\sigma_{\text{nom}}}^{\infty} \text{Gau}(0,1)$$

To achieve significance σ after trials correction

$\sigma =$ 2 **3** 4 5 (corrected σ desired)

requires larger $\sigma_{\text{nom}}(N)$ before trials correction for N trials

N equivalent observed significance $\sigma_{\text{nom}} =$

-----	-----	-----	-----	-----
10	2.9	3.7	4.6	5.5
50	3.4	4.1	4.9	5.7
400	3.9	4.5	5.3	6.1
16K	4.7	5.3	5.9	6.7

bigger change for smaller corrected σ

Details: Bonferroni Correction Math:

Derive: $p_{\text{corrected}}$ is roughly $q = N * p_{\text{nominal}}$

Exact Binomial probability for ≥ 1 of N found above p_n :

$$p_c = 1 - (1 - p_n)^N = (1 - e^{-q}) + O\left(\frac{q^2}{N}\right) \approx q \left(1 - \frac{q}{2}\right)$$

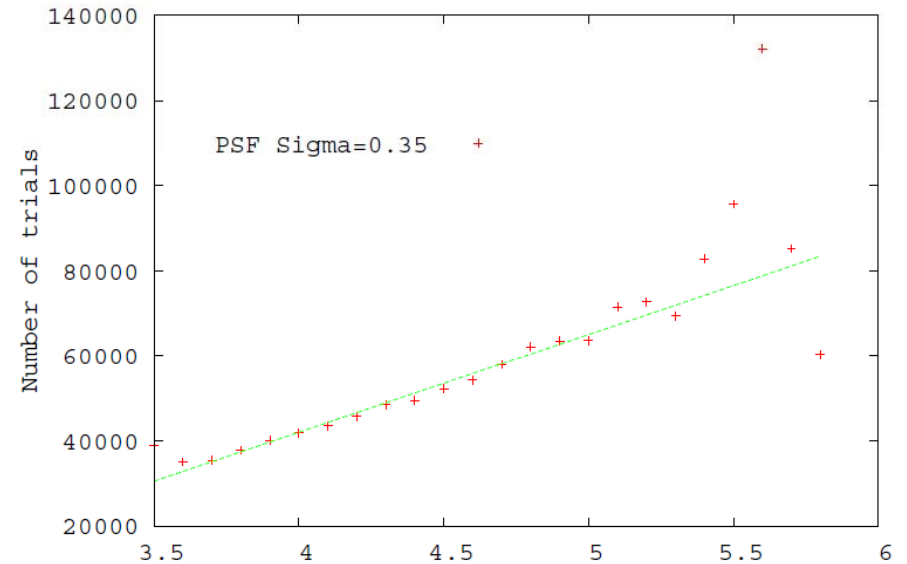
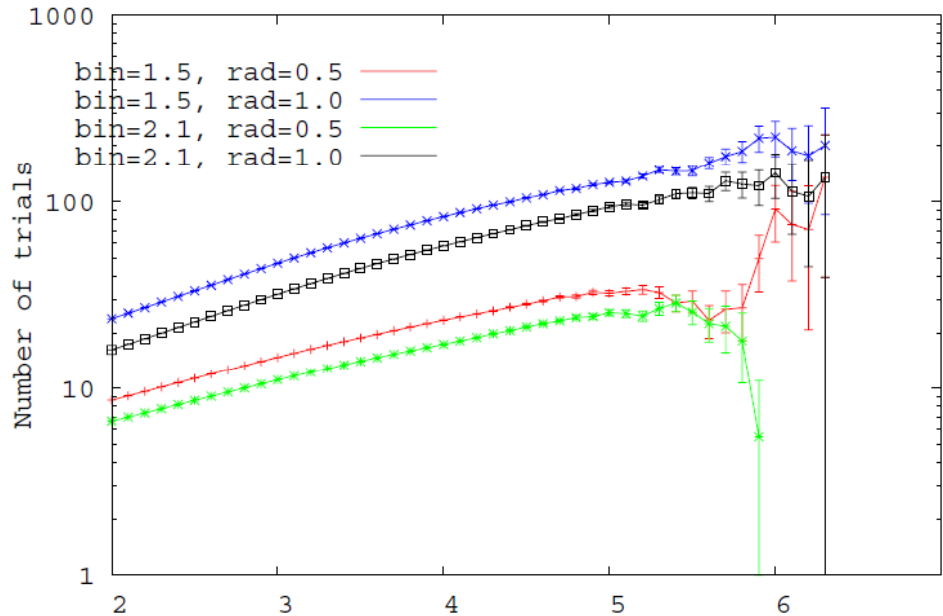
$p_c = q$ for $N=1$, and $p_c \sim q$ for $q \ll 1$

$Np=q$ sufficient for p_c of 2σ or more

$p_c=q=Np$ in terms of σ :

$$p_c = \int_{\sigma}^{\infty} \text{Gau}(0,1) = q = N \int_{\sigma_{\text{nom}}}^{\infty} \text{Gau}(0,1)$$

Effective # of trials vs Z



Chaun Chen, Milagro thesis 2008

similar behavior as mass search in 2D????