Smallest Interval

$$\mathcal{O}^{\mathrm{S}}_{1-\alpha} = \{r^*\}$$

$$P(\mathcal{O}^{\mathrm{S}}_{1-\alpha}|N,p) \overset{?}{\geq} 1-\alpha$$

No

yes

done

$$P(r^*+1|N,p) \overset{?}{>} P(r^*-1|N,p)$$

$$\mathcal{O}^{\mathrm{S}}_{1-\alpha} = \{r^*, r^*+1\} \qquad \mathcal{O}^{\mathrm{S}}_{1-\alpha} = \{r^*, r^*-1\}$$

$$P(\mathcal{O}^{\mathrm{S}}_{1-\alpha}|N,p) \overset{?}{\geq} 1-\alpha$$

collect elements according to probability rank until get $\geq 1-\alpha$

Binomial Distribution – example

r	P(r|N=5,p=0.6)	F(r|N=5,p=06) "1 − F"	Rank	F(r|N=5,p=06) According to rank
0	0.01024	0.01024 1	6	1
1	0.0768	0.08704 0.98	5	0.98976
2	0.2304	0.31744 0.90	3	0.8352
3	0.3456	0.66304 0.67	1	0.3456
4	0.2592	0.92224 0.33	2	0.6048
5	0.07776	1 0.078	4	0.91296

$1-\alpha = 0.9$

$\frac{\alpha}{2} = 0.05$

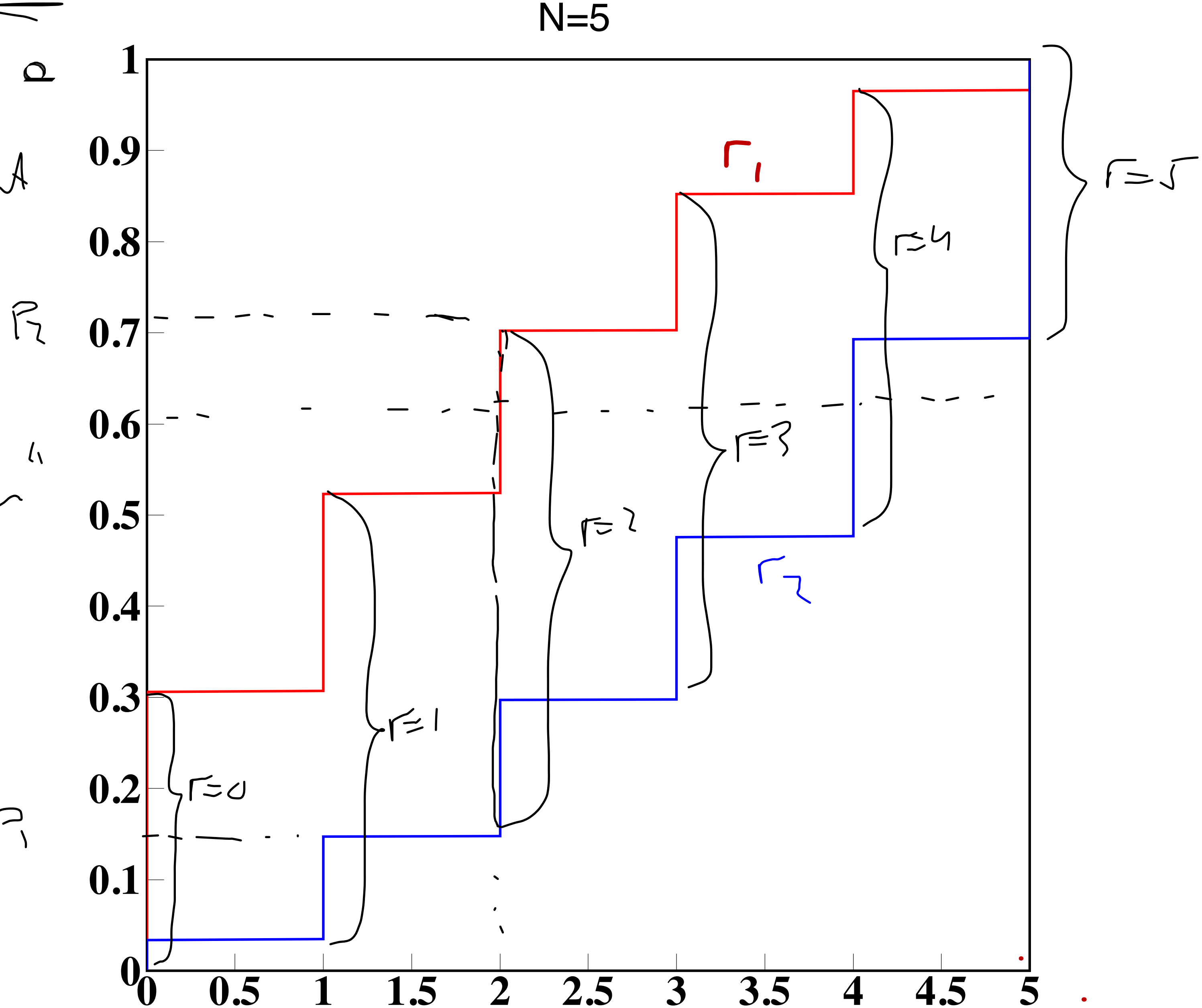
$r_1 = 1$

$r_2 = 5$

$O^c_{0.9} = \{1, 2, 3, 4, 5\}$

$O^s_{0.9} = \{2, 3, 4, 5\}$

68% Confidence Level Construction



make measurement

say $r=2$

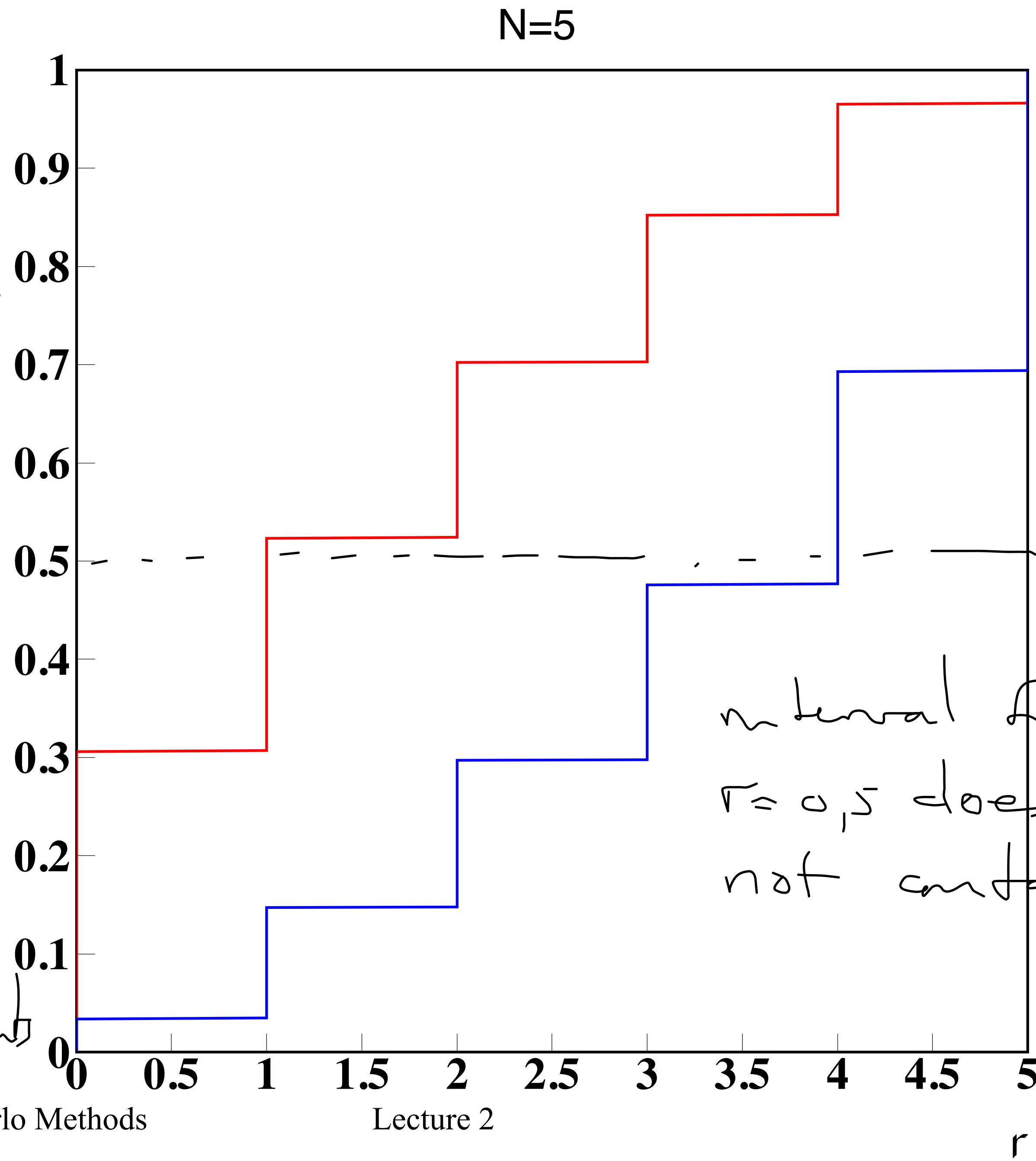
"Neyman construction"

$1-\alpha$ CL

$p_1 \leq p \leq p_2$

Not a probability for p

68% Confidence Level Construction



Suppose there is a "true value" p_0

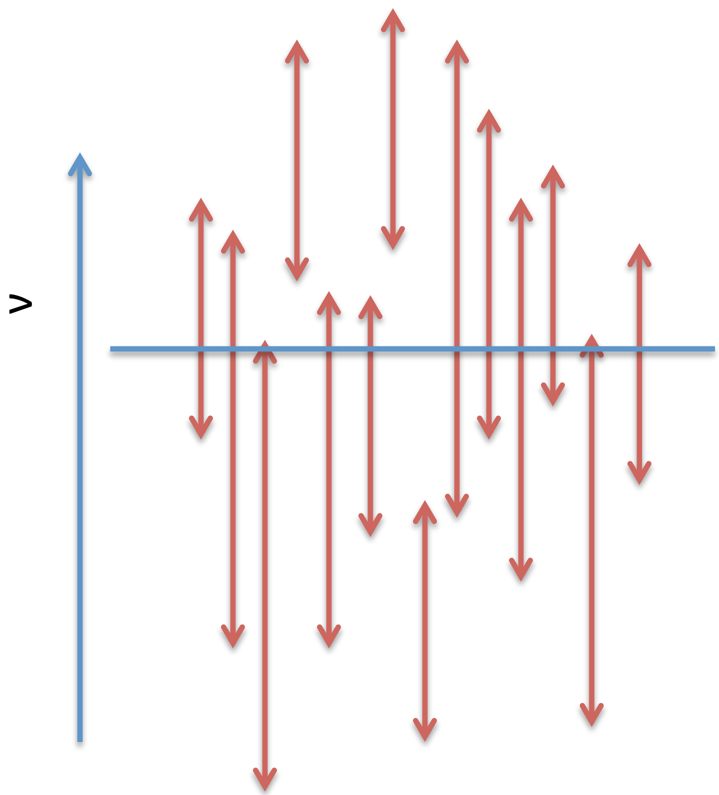
$p_0 = 0.5$

$[p_1, p_2]$ will contain p_0 in $\geq 1-\alpha$ of experiments

interval from $r = 0,5$ does not contain p_0

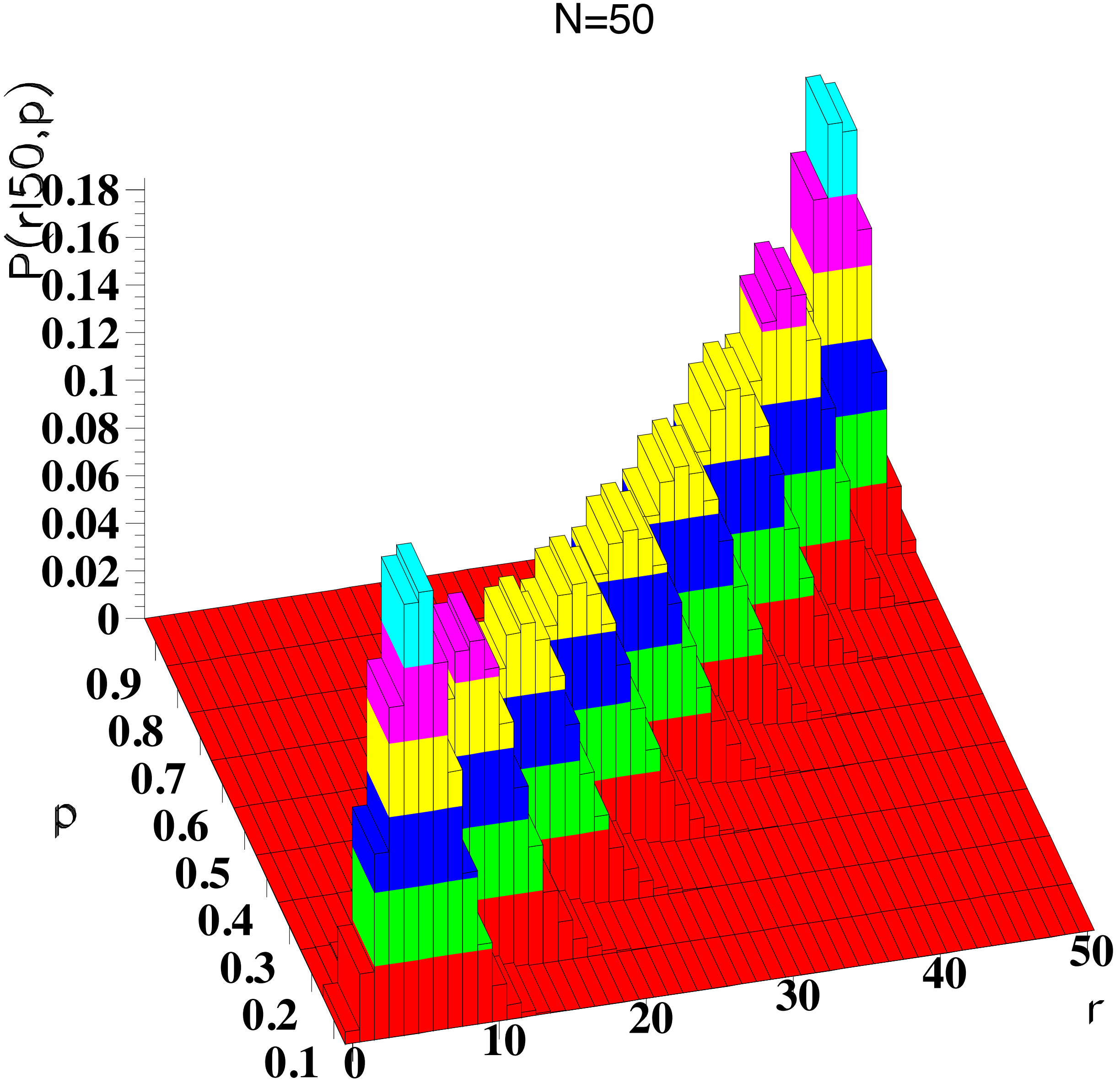
Use of Frequentist CL Intervals

Not intended to be used individually. Rather, collect a lot of intervals and use these to `find' the true value (not specified how).

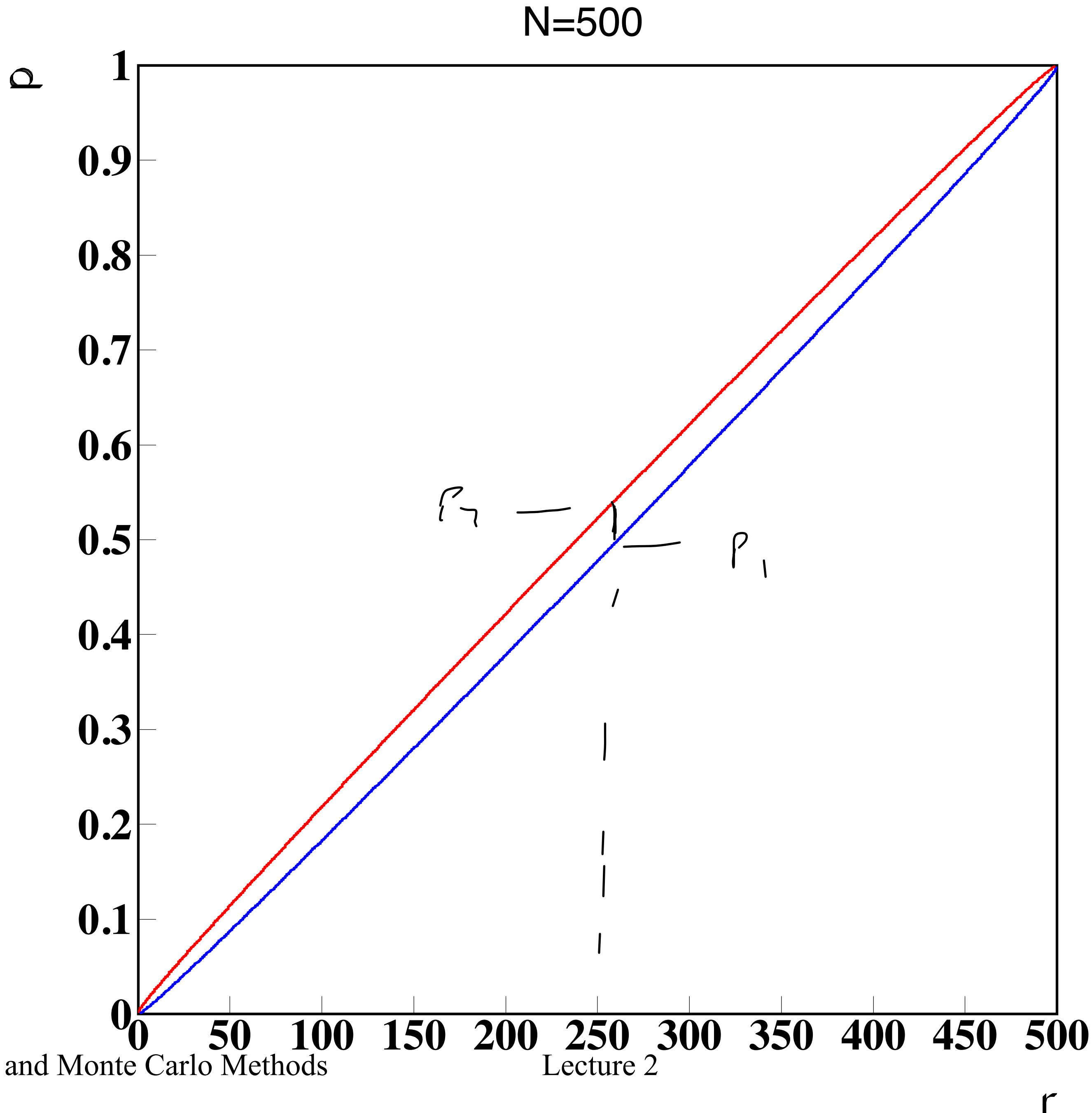


True value should be the one that is in the set of intervals the right fraction of the time (e.g., 68% of the time for the 68% CL intervals).

Binomial Distribution – in more detail



68% Confidence Levels



Binomial Models – Bayesian Analysis

$P_0(\lambda|M)P_0(M)$

$$P(\vec{\lambda}, M|\vec{D}) = \frac{P(\vec{D}|\vec{\lambda}, M)P(\vec{\lambda}, M)}{P(\vec{D})}$$

Bayes Equation

$$P(M|\vec{D}) = \underbrace{\int P(\vec{\lambda}, M|\vec{D})d\vec{\lambda}}_{\text{marginalization}} = \frac{\left[\int P(\vec{D}|\vec{\lambda}, M)P_0(\vec{\lambda}|M)d\vec{\lambda}\right] P_0(M)}{P(\vec{D})}$$

evidence $\quad Z = \int P(\vec{D}|\vec{\lambda}, M)P_0(\vec{\lambda}|M)d\vec{\lambda}$

"marginal likelihood"

Posterior Odds $\quad \dfrac{P(M_1|\vec{D})}{P(M_2|\vec{D})} = \dfrac{Z_1}{Z_2}\dfrac{P_0(M_1)}{P_0(M_2)}$ Prior Odds

Bayes Factor

Binomial Distribution

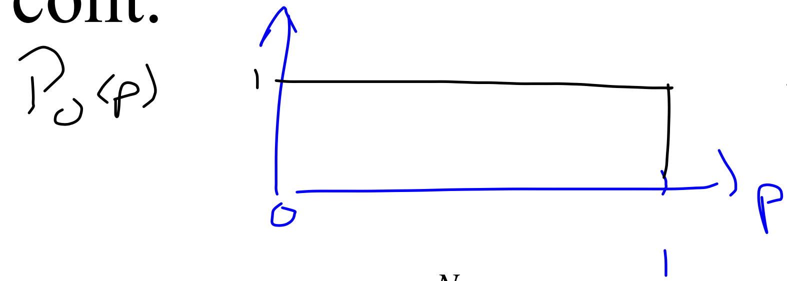
Suppose we perform N trials and have r successes. What is the probability distribution for p ?

E.g., you want to say something about the efficiency of your detector.
From Bayes' Theorem:

in general, you need to specify this

$$P(p \mid r,N) = \frac{P(r \mid p,N)P_0(p)}{\int_0^1 P(r \mid p,N)P_0(p)\,dp} = \frac{\frac{N!}{(N-r)!r!}p^r(1-p)^{N-r}P_0(p)}{\int_0^1 \frac{N!}{(N-r)!r!}p^r(1-p)^{N-r}P_0(p)\,dp}$$

Binomial – cont.



If we assume that $P_0(p)$ is a constant

$$P(p\,|\,r,N) = \frac{P(r\,|\,p,N)P_0(p)}{\int_0^1 P(r\,|\,p,N)P_0(p)dp} = \frac{p^r(1-p)^{N-r}}{\int_0^1 p^r(1-p)^{N-r}dp}$$

The integral is technically a 'β function', and for x,n integers we have

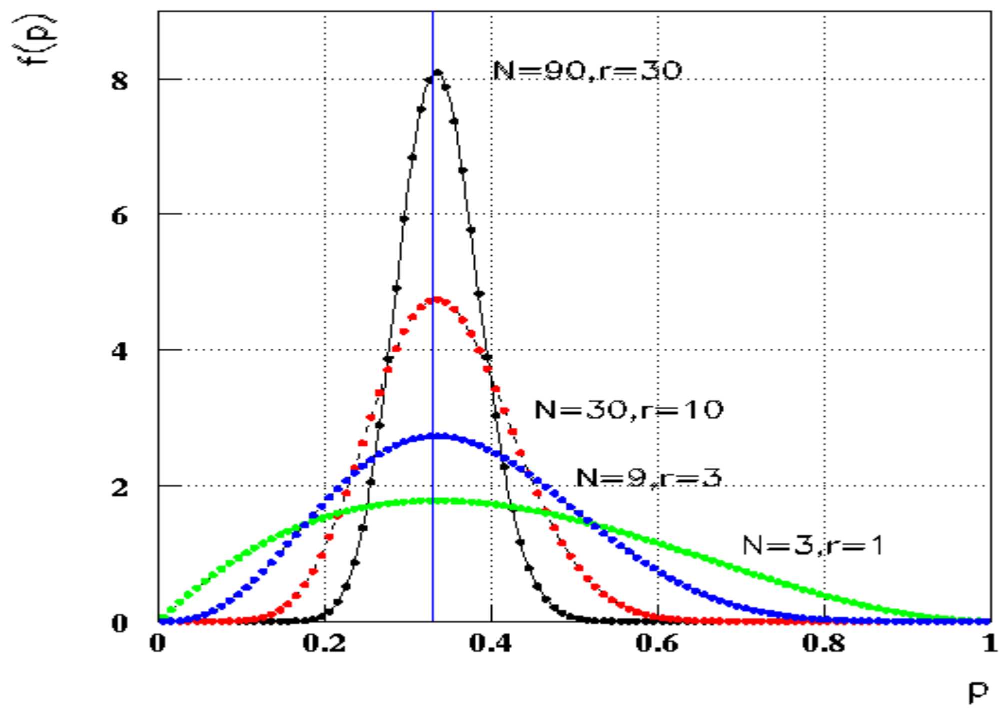
$$\int_0^1 p^x(1-p)^{n-x}dp = \frac{x!(n-x)!}{(n+1)!}$$

so

$$\boxed{P(p\,|\,r,N) = \frac{(N+1)!}{r!(N-r)!}p^r(1-p)^{N-r}}$$

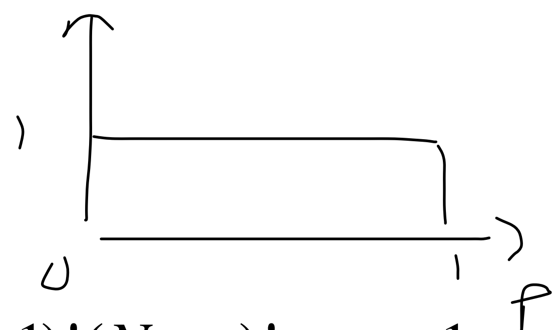
Master formula

Binomial - cont.



Note mode at p=r/N

all examples $p^* = 1/3$
but different N



The expectation value and variance are:

$$< p >= \int_0^1 \frac{(N+1)!}{r!(N-r)!} p^{r+1}(1-p)^{N-r}\,dp = \frac{(N+1)!}{r!(N-r)!}\frac{(r+1)!(N-r)!}{(N+2)!} = \frac{r+1}{N+2}$$

$$\sigma^2 = \frac{(r+1)(N-r+1)}{(N+3)(N+2)^2} = < p > (1- < p >)\frac{1}{N+3}$$

$\sigma = \frac{1}{\sqrt{12}}$
for flat dist

Detailed Example

Imagine we will perform a test of some equipment to determine how well it works (success rate). Two sets of data are taken:

N=100 trials, r=100 Successes

N=100 trials, r=95 Successes

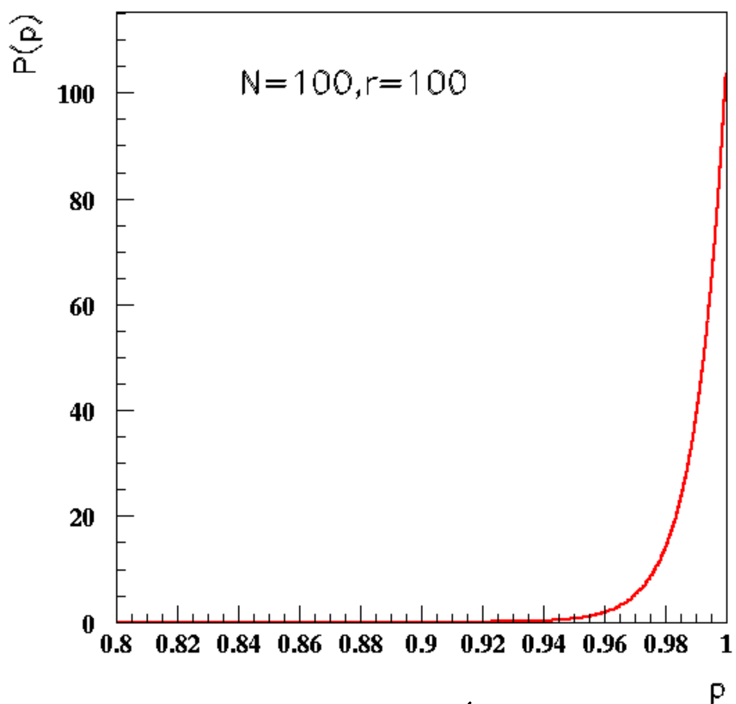
Questions:

- what can we say about the efficiency in the first case ?
- what about in the second case
- assuming the tests should be modeled with the same underlying model, what is our combined answer ?
- how does it compare to N=200, r=195 ? (you do this in the exercises)
- are the results compatible ?

Example

Data set 1: N=100, r=100

$$P(p \mid r,N) = \frac{(N+1)!}{r!(N-r)!} p^r (1-p)^{N-r} \qquad P(p|100,100) = 101 p^{100}$$



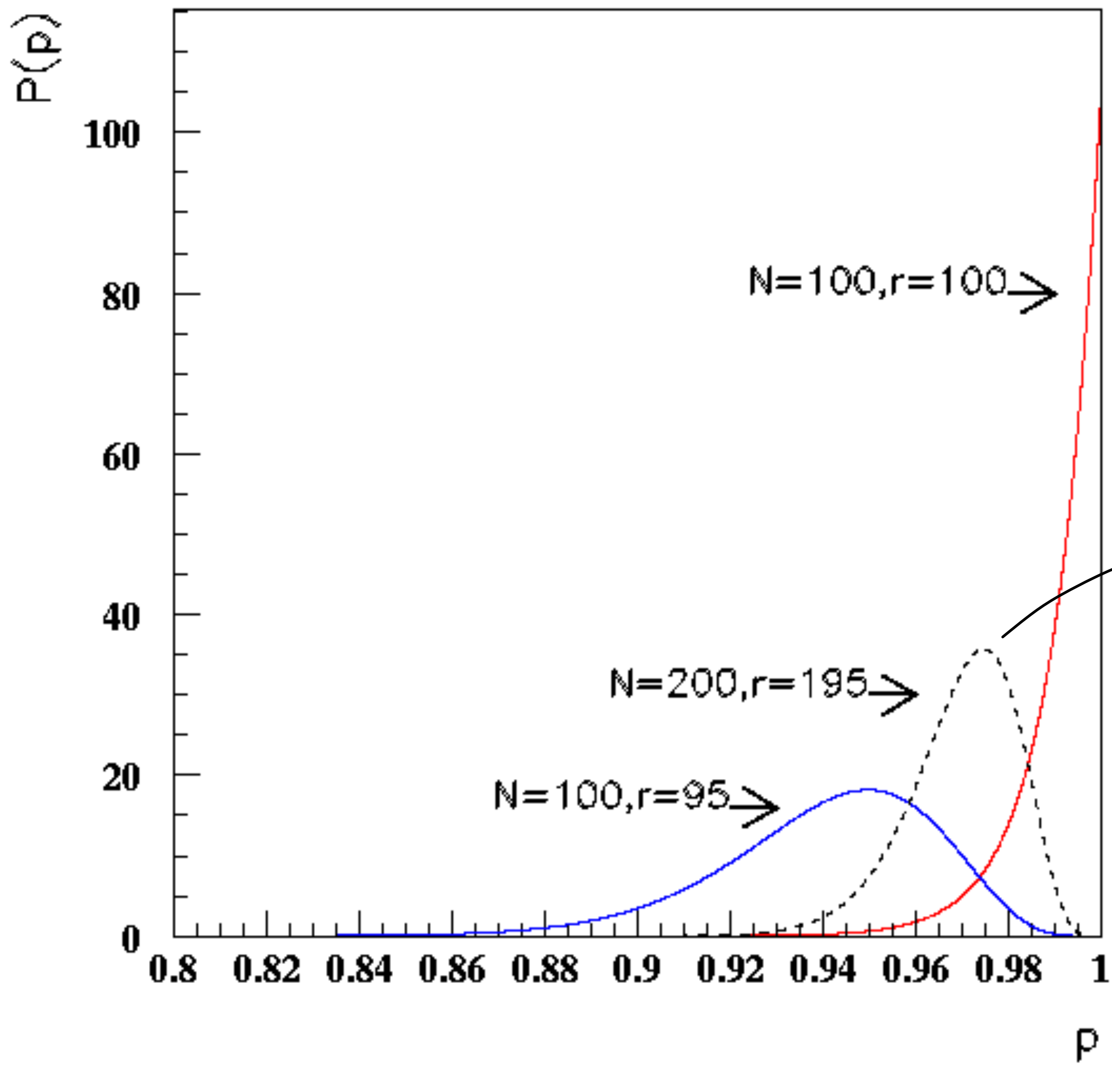
$$p^* = \frac{r}{N} = 1$$

$$\langle p \rangle \simeq \frac{r+1}{N+2} = \frac{101}{102} \simeq 0{,}99$$

$$\sigma^2 \simeq \frac{\langle p \rangle (1 - \langle p \rangle)}{N+3} \simeq \frac{0{,}99\,(0{,}01)}{103}$$

$$\sigma = \sqrt{\sigma^2} \simeq 0{,}01$$

Example-continued



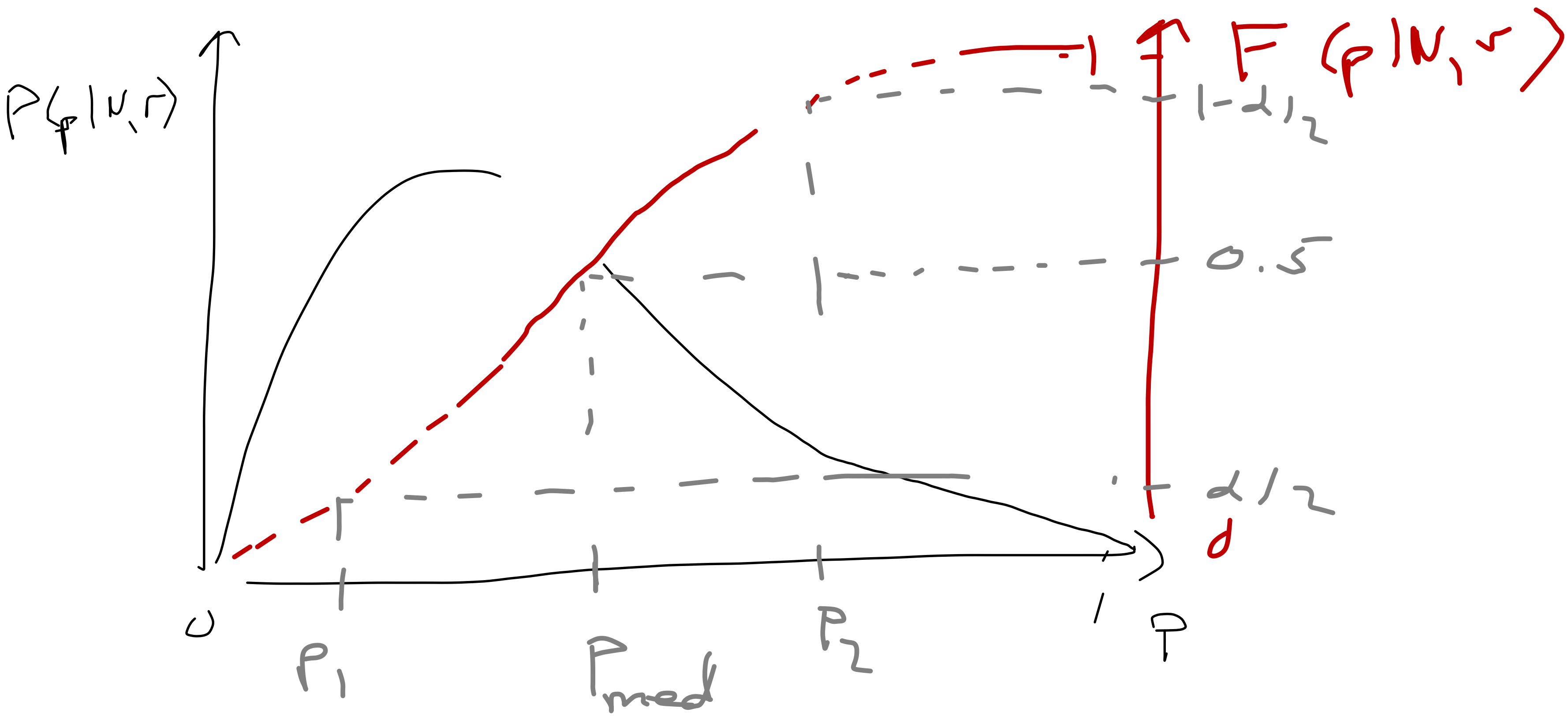
combined data

Take posterior pdf from E1 as prior for 2 → same result

Summarizing a Distribution

- Median
- Central interval $(1-\alpha)$

$$F(p_{\mathrm{median}}) = 0.5 \quad F(p_{\mathrm{lower}}) = \frac{\alpha}{2} \quad F(p_{\mathrm{upper}}) = 1 - \frac{\alpha}{2}$$



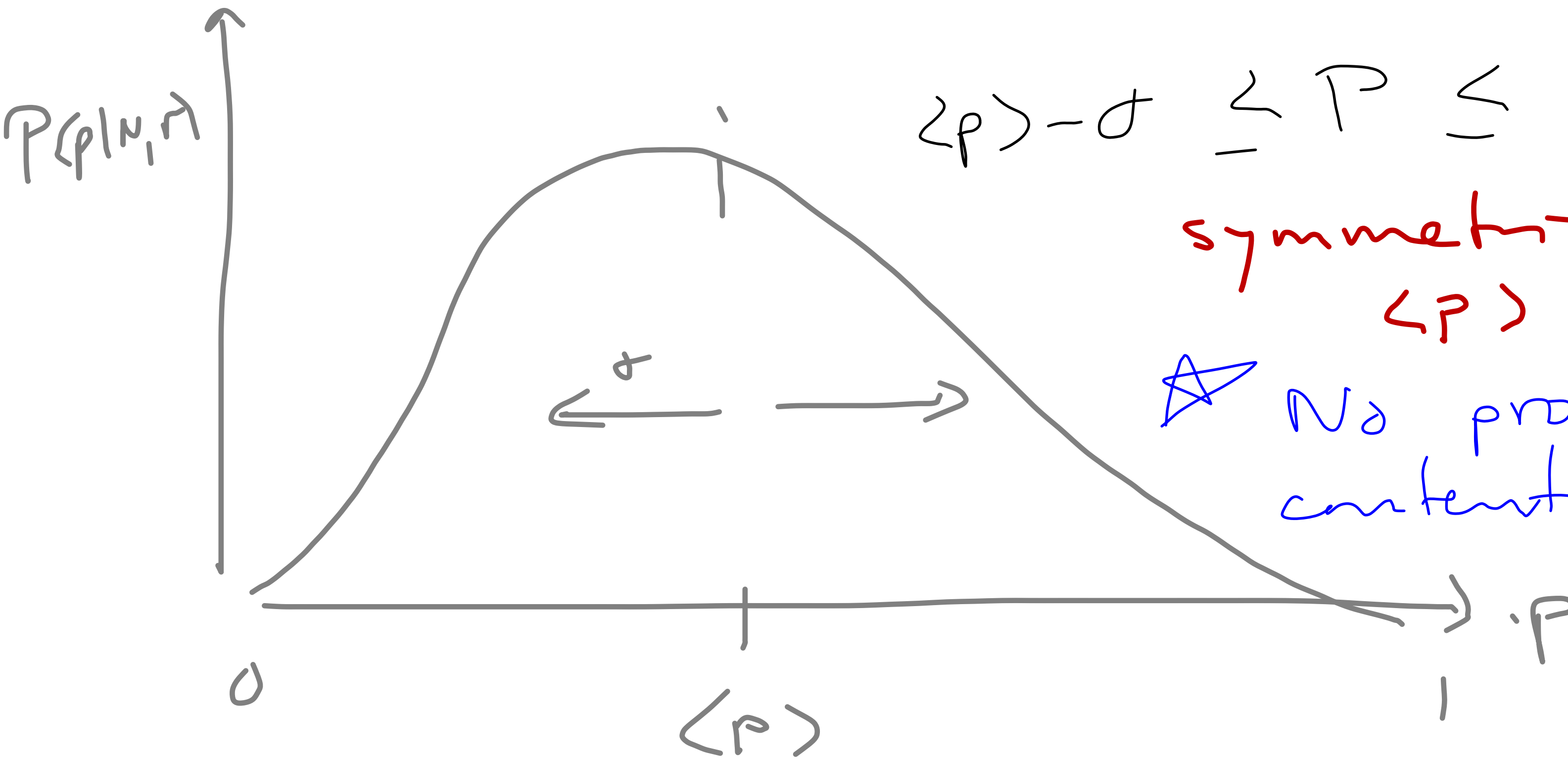
Summarizing a Distribution

- Mean

$$E[p] = \int_0^1 p P(p|N,r)\,dp$$

- rms

$$\sigma = \sqrt{E[p^2] - E[p]^2}$$



$\langle p \rangle - \sigma \leq P \leq \langle p \rangle + \sigma$

symmetric about $\langle p \rangle$

☆ No probability content defined

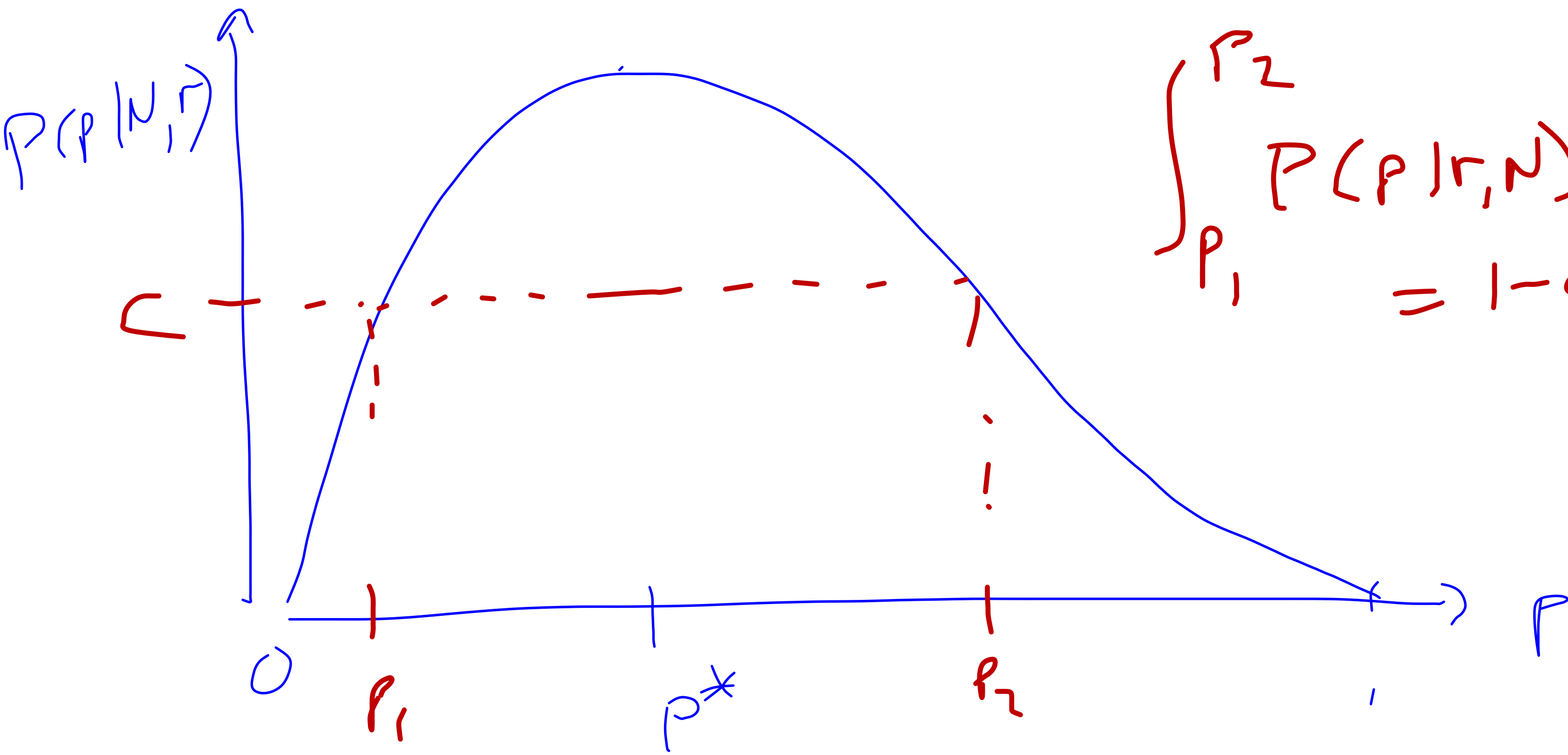
Summarizing a Distribution

- Most-probable value (mode)

$$p_{\text{mode}} = \max_p \left[P(p|r,N) \right]$$

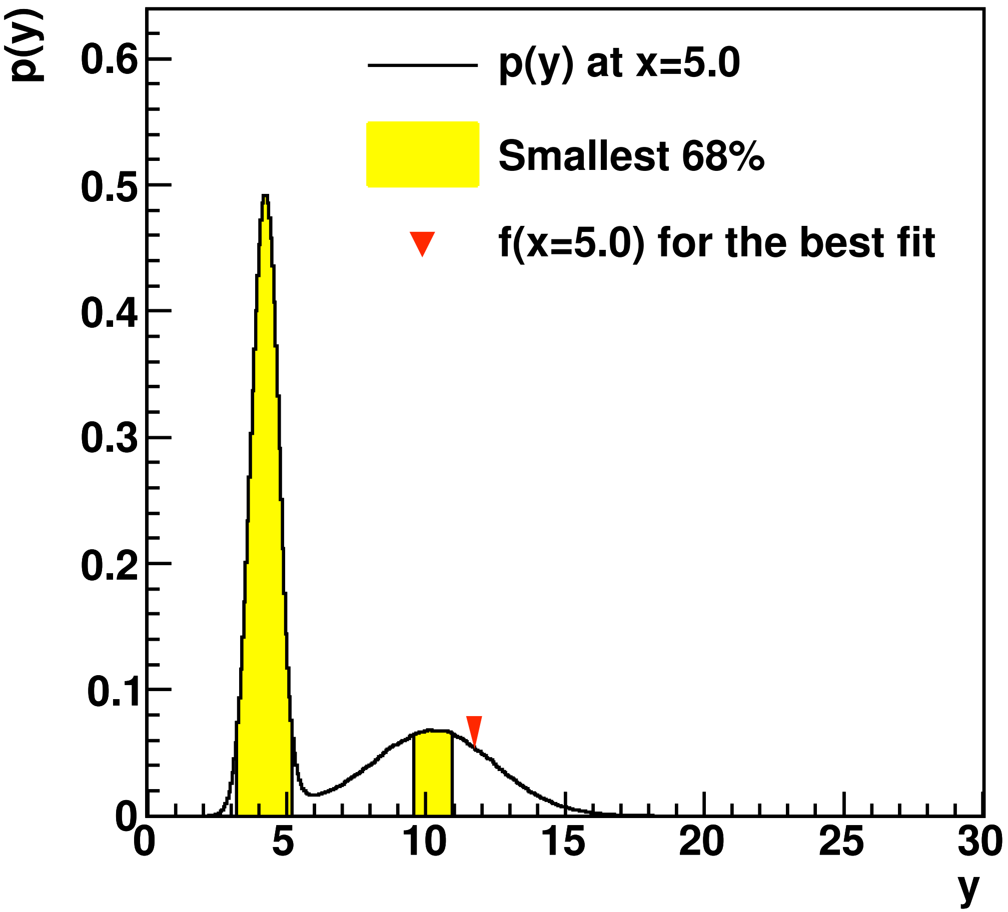
- Shortest interval(s)

$$1 - \alpha = \int_{P(p|r,N)>C} P(p|r,N)\,dp$$



$$\int_{p_1}^{p_2} P(p|r,N)\,dp = 1-\alpha$$

Example – multimodal distribution



Poisson Distribution

A Poisson distribution applies when we do not know the number of trials (it is a large number), but we know that there is a fixed probability of 'success' per trial, and the trials occur independently of each other.

Alternatively – a continuous time process with a constant rate will produce a Poisson distributed number of events in a fixed time interval.

High energy physics example: beams collide at a high frequency (10 MHz, say), and the chance of a 'good event' is very small. The resulting number of events in a fixed time will follow a Poisson distribution. A single trial is one crossing of the beams.

Nuclear physics example: a large sample of radioactive atoms will produce a Poisson distributed number of events in a fixed time interval (assuming a $\tau >> T$)

Poisson Distribution

The Poisson distribution can be derived from the Binomial distribution in the limit when N →∞ and p →0, but Np fixed and finite. Then

$\nu = Np$

$$\underbrace{P(r|N,p)}_{\text{Binomial pd}} \rightarrow \underbrace{P(n|\nu)}_{\text{Poisson}}$$

The expected number of events is calculated from a rate, or from a luminosity and cross section or some other way

total time luminosity

$$\nu = R \cdot T \;\; \text{or} \;\; \nu = \mathcal{L} \cdot \sigma \;\; \text{or...}$$

rate cross section 'probability'

Poisson Distribution - derivation

$\nu = Np$

$$P(n|N,p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

$$P(n|N,\frac{\nu}{N}) = \frac{N!}{n!(N-n)!} \frac{\nu^n}{N^n} \left(1 - \frac{\nu}{N}\right)^{N-n}$$

$\underline{N \to \infty}$ $\quad n \lesssim 10\sigma$

$$\frac{N!}{(N-n)!} = N \cdot (N-1) \cdot \ldots \cdot (N-n+1) \approx N^n \quad + O(N^{n-1})$$

$$\left(1 - \frac{\nu}{N}\right)^{N-n} \to \left(1 - \frac{\nu}{N}\right)^{N} \to e^{-\nu}$$

$$\boxed{P(n|\nu) = \frac{e^{-\nu}\nu^n}{n!}}$$ Poisson Distribution

$n = 0, 1, 2, \ldots \infty$

Poisson Distribution-alternate derivation

Process with a constant rate, R. What is the probability that no event has occurred up to time t ?

Divide t into many very small intervals Δt, t=nΔt. Then

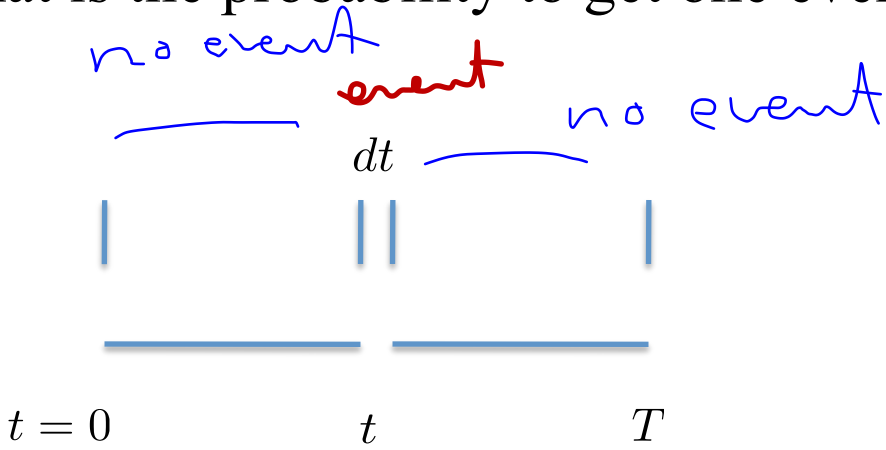
probability of no event in Δt

$1 - \underbrace{\text{prob of event}}_{R\Delta t}$

$$
\begin{aligned}
P(\text{no event}) &= (1 - R\Delta t)^n \\
&= (1 - Rt/n)^n \\
&= e^{-Rt} \ \text{ for } \ n \to \infty
\end{aligned}
$$

Poisson Distribution-alternate derivation

Now what is the probability to get one event in interval T ?



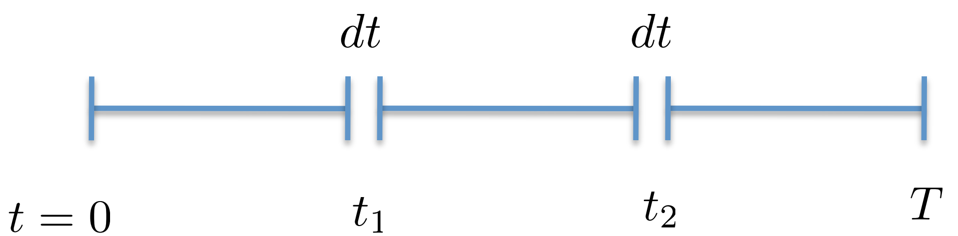
$$\int_0^T e^{-Rt} e^{-R(T-t)} (R dt) = RTe^{-RT}$$

Or, using $\nu = RT$ $\qquad$ $P(1|\nu) = \nu e^{-\nu}$

$n = 1$

Poisson Distribution-alternate derivation

Now what is the probability to get two events in interval T ?



$$P(2|R,T) = e^{-RT} \int_0^T R\,dt_1 \int_{t_1}^T R\,dt_2$$

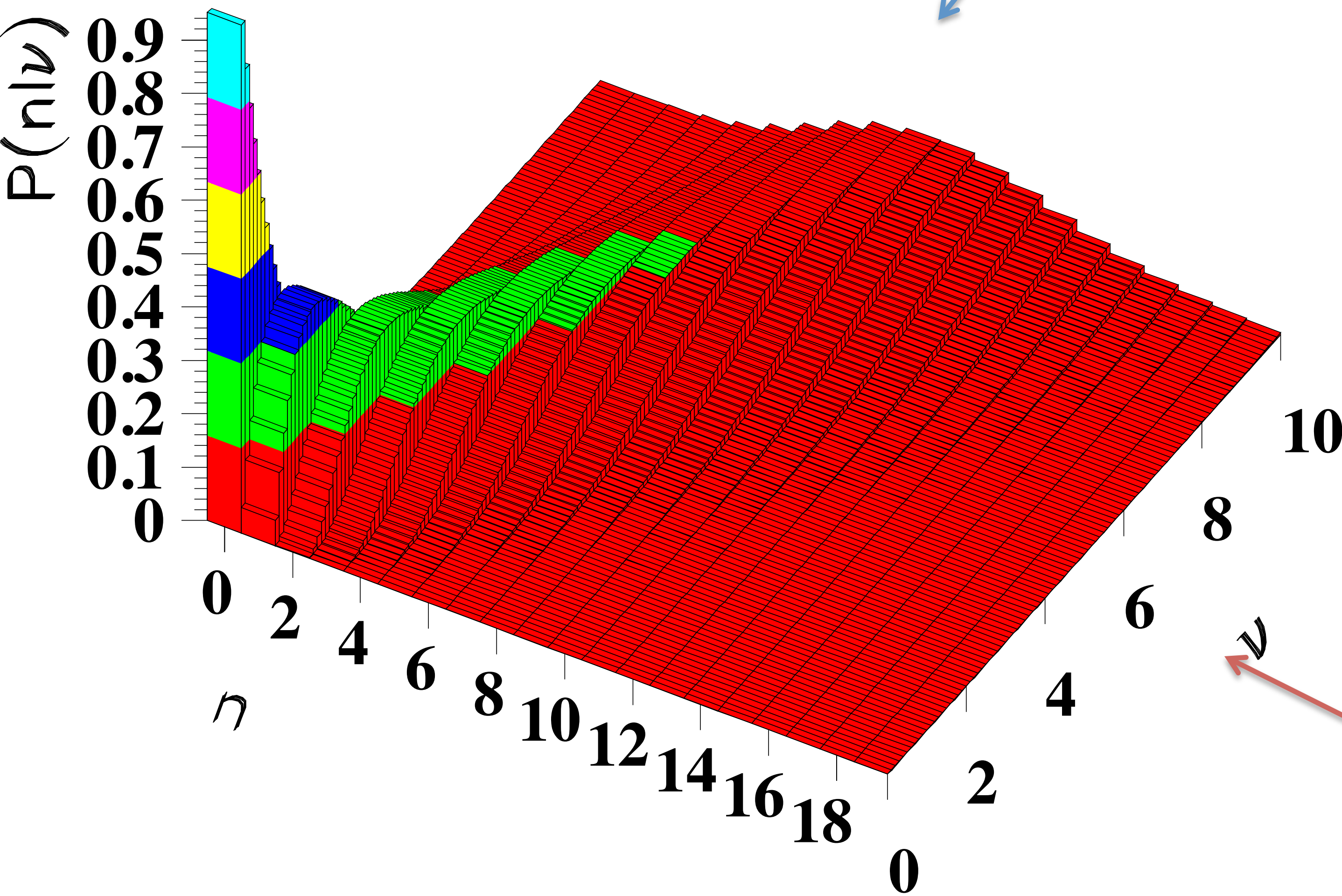
$$P(2|R,T) = R^2 e^{-RT} \int_0^T (T - t_1)dt_1 = \frac{R^2 T^2 e^{-RT}}{2}$$

$\nu = RT$

or $\quad P(2|\nu) = \dfrac{\nu^2 e^{-\nu}}{2} \quad$ and $\quad P(n|\nu) = \dfrac{\nu^n e^{-\nu}}{n!}$

Poisson Example

Likelihood used in likelihood analysis



Probability of the data used in confidence level setting

Poisson Distribution-cont.

$$P(n \mid \nu) = \frac{\nu^n e^{-\nu}}{n!}$$

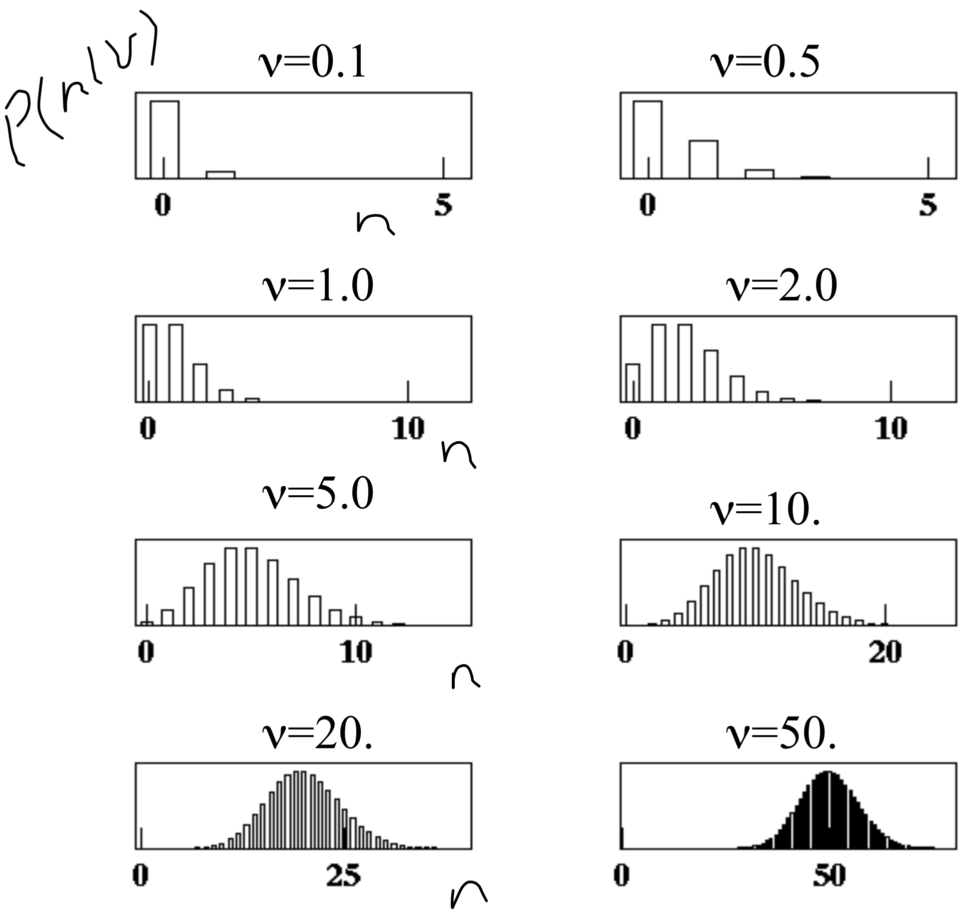
$n^* = \lfloor \nu \rfloor$

$n^* = \lceil \nu \rceil - 1$

mode } double if ν integer

$E[n] = \nu$

$\mathrm{V}[n] = \nu$ variance of n



Notes:

- As ν increases, the distribution becomes more symmetric
- Approximately Gaussian for large ν
- Poisson formula is much easier to use that the Binomial formula.

Poisson Distribution-cont.

Proof of Normalization, mean, variance:

$$Normalization: \sum_{n=0}^{\infty} \frac{\nu^n e^{-\nu}}{n!} = e^{-\nu} \sum_{n=0}^{\infty} \frac{\nu^n}{n!} = e^{-\nu} e^{\nu} = 1$$

P(n|ν)

$$E[n] = \sum_{n=0}^{\infty} n \frac{\nu^n e^{-\nu}}{n!} = e^{-\nu} \sum_{n=1}^{\infty} \nu \frac{\nu^{n-1}}{(n-1)!} = \nu e^{-\nu} e^{\nu} = \nu$$

$$V[n] = E[n^2] - E[n]^2$$

$$E[n^2] = \sum_{n=0}^{\infty} n^2 \frac{\nu^n e^{-\nu}}{n!} = e^{-\nu} \sum_{n=1}^{\infty} \nu n \frac{\nu^{n-1}}{(n-1)!} \qquad write\ n = (n-1+1)$$

$$= \nu e^{-\nu} \left(\sum_{n=1}^{\infty} (n-1) \frac{\nu^{n-1}}{(n-1)!} + \sum_{n=1}^{\infty} \frac{\nu^{n-1}}{(n-1)!} \right) = \nu^2 + \nu$$

$$V[n] = \nu^2 + \nu - \nu^2 = \nu$$

Example

Example for ν=10/3 $= 3.\bar{3}$

$1-\alpha \geq 0.9$

$n^* = \lfloor \nu \rfloor = 3$

n	$P(n\|\nu)$	$F(n\|\nu)$	R	$F_R(n\|\nu)$
0	0.0357	0.0357	7	0.9468
1	0.1189	0.1546	5	0.8431
2	0.1982	0.3528	2	0.4184
3	0.2202	0.5730	1	0.2202
4	0.1835	0.7565	3	0.6019
5	0.1223	0.8788	4	0.7242
6	0.0680	0.9468	6	0.9111
7	0.0324	0.9792	8	0.9792
8	0.0135	0.9927	9	0.9927
9	0.0050	0.9976	10	0.9976
10	0.0017	0.9993	11	0.9993
11	0.0005	0.9998	12	0.9998
12	0.0001	1.0000	13	1.0000

$\leq \alpha/2$ prob

$O^C_{0.9} = \{1, \ldots, 7\}$

$O^S_{0.9} = \{1, \ldots, 6\}$

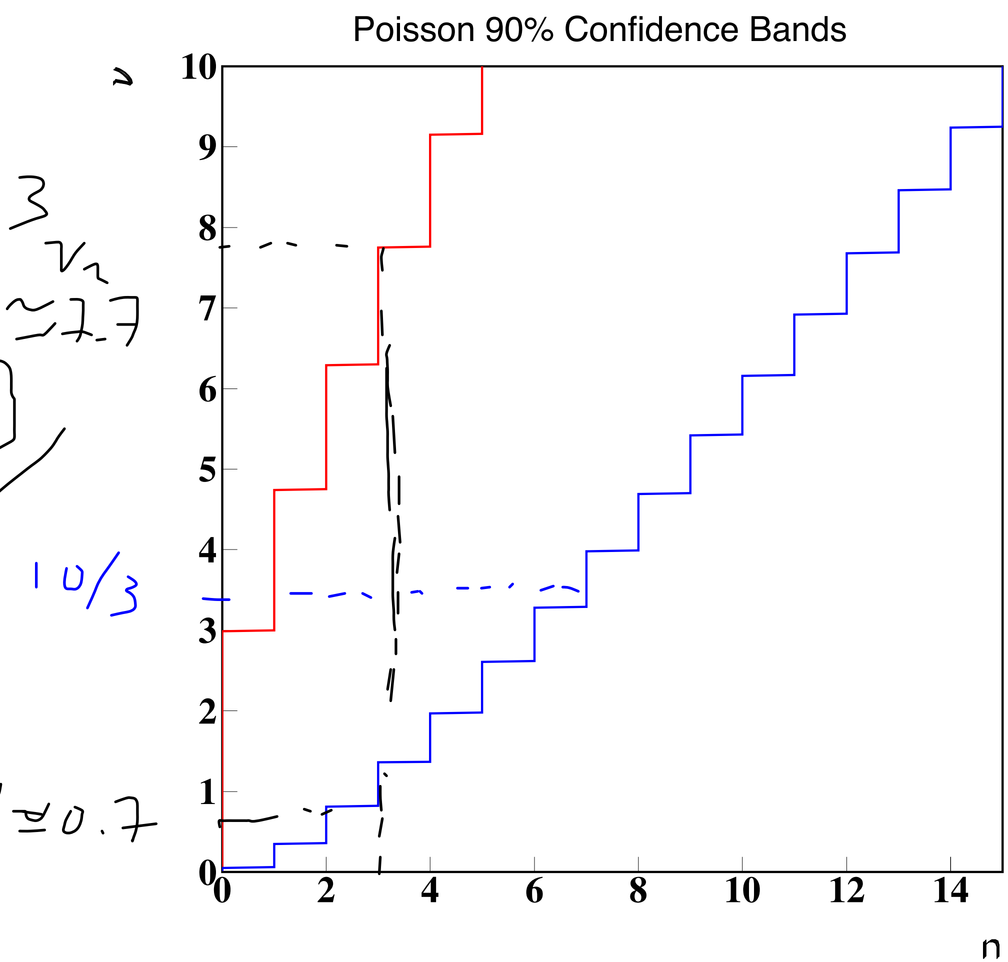
Confidence Level Calculation

We observe n events, and ask which values of ν are accepted with confidence level 1-α. For 1- α=0.9, central intervals:

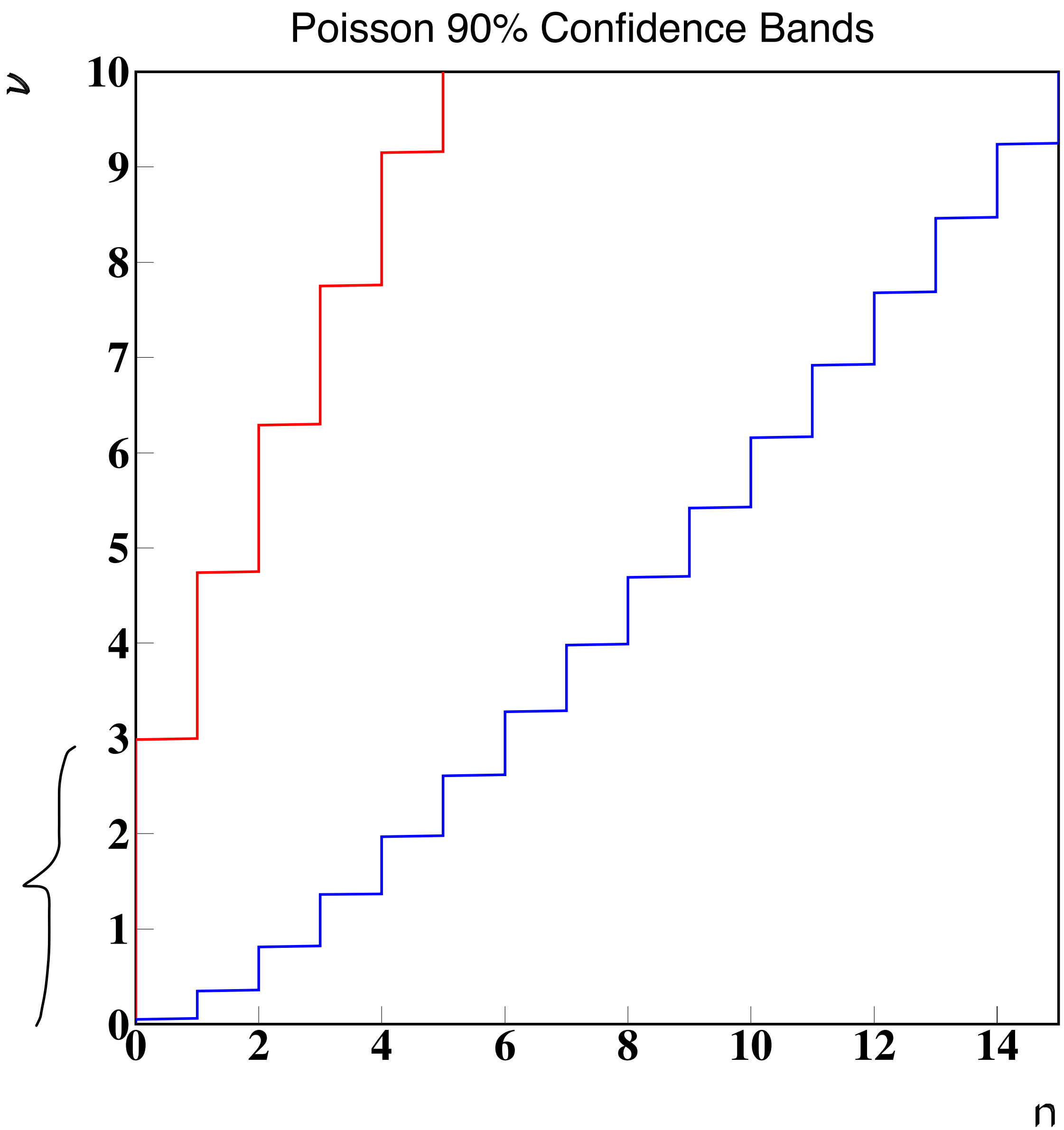
Suppose measure $n=3$

$\nu = \{0.7, 7.7\}$

90% CL



Confidence Level Calculation



n=0

>5% prob to get n=0

Upper limit on parameter @ 95 CL

$\nu < 3$ for n=0

Bayesian Data Analysis-Poisson Distribution

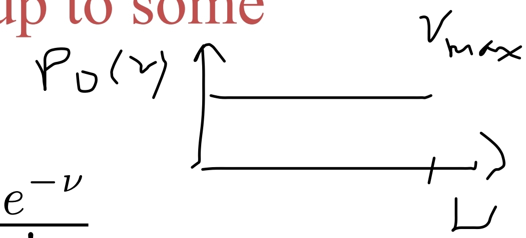
Typical examples – counting experiments such as source activity, failure rates, cross sections,…

$$P(\nu|n) = \frac{P(n|\nu)P_0(\nu)}{\int_0^\infty P(n|\nu)P_0(\nu)d\nu} = \frac{\frac{\nu^n e^{-\nu}}{n!}P_0(\nu)}{\int_0^\infty \frac{\nu^n e^{-\nu}}{n!}P_0(\nu)d\nu}$$

This is our master formula. Result in general will depend on choice of prior.

Poisson - cont.

If we assume a flat prior starting at 0 and extending up to some maximum of ν much larger than *n*.



$$P(\nu|n) = \frac{\frac{\nu^n e^{-\nu}}{n!}P_0(\nu)}{\int_0^\infty \frac{\nu^n e^{-\nu}}{n!}P_0(\nu)d\nu} = \frac{\frac{\nu^n e^{-\nu}}{n!}}{\int_0^{\nu_{max}} \frac{\nu^n e^{-\nu}}{n!}d\nu}$$

$\nu_{max} \gg n$

$$\int_0^{\nu_{max}} \frac{\nu^n e^{-\nu}}{n!}d\nu \approx \frac{1}{n!}\int_0^\infty \nu^n e^{-\nu}d\nu = \frac{1}{n!}n! = 1$$

n

$$P(\nu|n) = \frac{e^{-\nu}\nu^n}{n!} \qquad \nu^* = n$$

now probability density for ν !

need a ν_{max} for a "proper" prior

Poisson - cont.

The expectation value:

$$< \nu >= \int_0^\infty P(\nu|n)\nu d\nu = \int_0^\infty \frac{\nu^n e^{-\nu}}{n!}\nu d\nu = \frac{(n+1)!}{n!} = n+1$$

The variance:

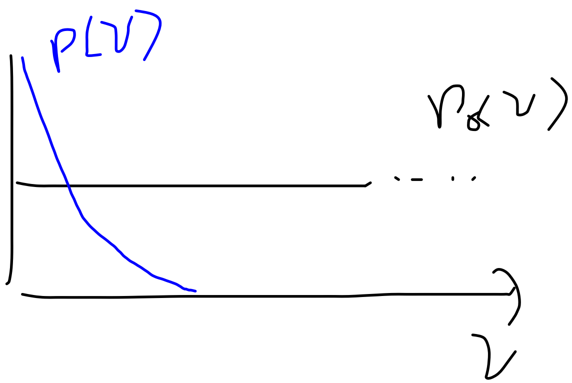
1

$$\begin{aligned}
\sigma^2 &= \int_0^\infty P(\nu|n)(\nu - <\nu>)^2 d\nu \\
&= \int_0^\infty \frac{\nu^n e^{-\nu}}{n!}\nu^2 d\nu - <\nu>^2 \int_0^\infty \frac{\nu^n e^{-\nu}}{n!} d\nu \\
&= \frac{(n+2)!}{n!} - (n+1)^2 = n+1
\end{aligned}$$

Poisson - cont.

Note: $n=0$ $\quad <\nu>=1$ $\quad ???$

From prior, expect
$$< \nu > \;=\; \int_0^{\nu_{max}} P_0(\nu)\nu d\nu = \int_0^{\nu_{max}} \frac{\nu}{\nu_{max}} d\nu$$
$$= \left[\frac{\nu^2}{2(\nu_{max})}\right]_0^{\nu_{max}}$$
$$= \frac{\nu_{max}}{2}$$



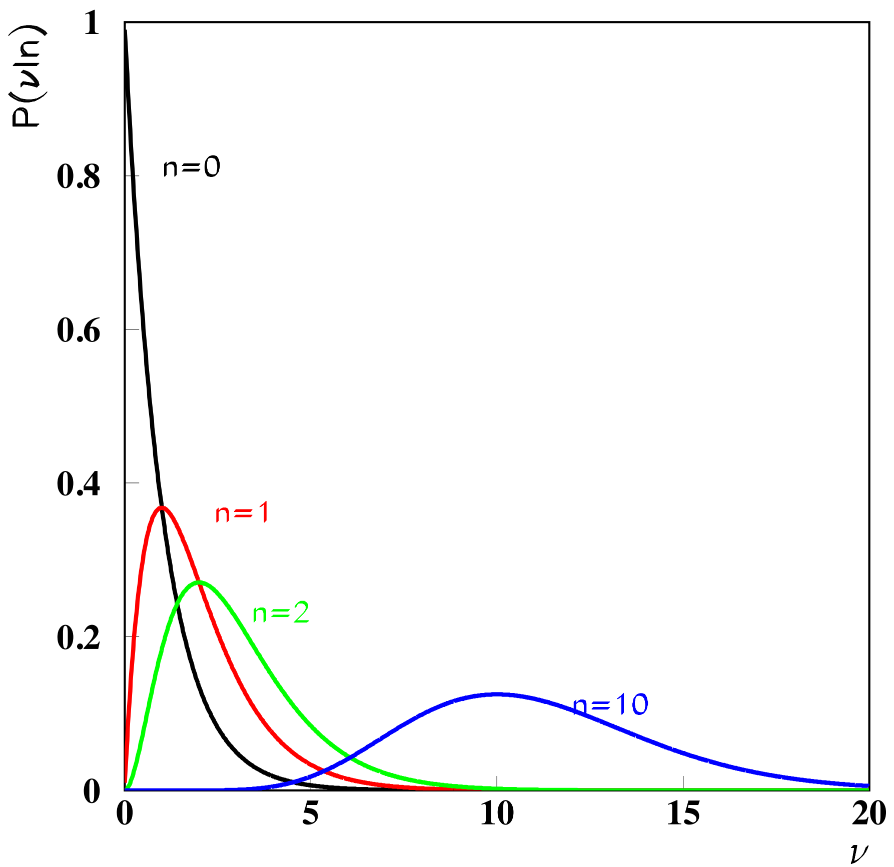
What happened ?

n=0 is a measurement !

$$P(\nu|0) = e^{-\nu}$$

Poisson – cont.

Some examples



Comments:

If you decide to quote the mode as your nominal result, you would use $\nu^{*}=n$. For large enough n, the 68% probability region is then approximately

$$n-\sqrt{n}\rightarrow n+\sqrt{n}$$

Poisson - cont.

The cumulative distribution function:

$$
\begin{aligned}
F(\nu|n) &= \int_0^{\nu} \frac{\nu'^n e^{-\nu'}}{n!} d\nu' \\
&= \frac{1}{n!}\left[-\nu'^n e^{-\nu'}|_0^{\nu} + n\int_0^{\nu} \nu'^{n-1} e^{-\nu'} d\nu'\right] \\
&= 1 - e^{-\nu}\sum_{i=0}^{n} \frac{\nu^i}{i!}
\end{aligned}
$$

Poisson – Examples

Assume measure zero counts.

With flat prior assumption

$$P(\nu|n=0) = e^{-\nu}$$
$$F(\nu|n=0) = 1-e^{-\nu}$$

For a 95% credibility upper limit

$$0.95 = 1-e^{-\nu}$$
$$\nu \approx 3$$

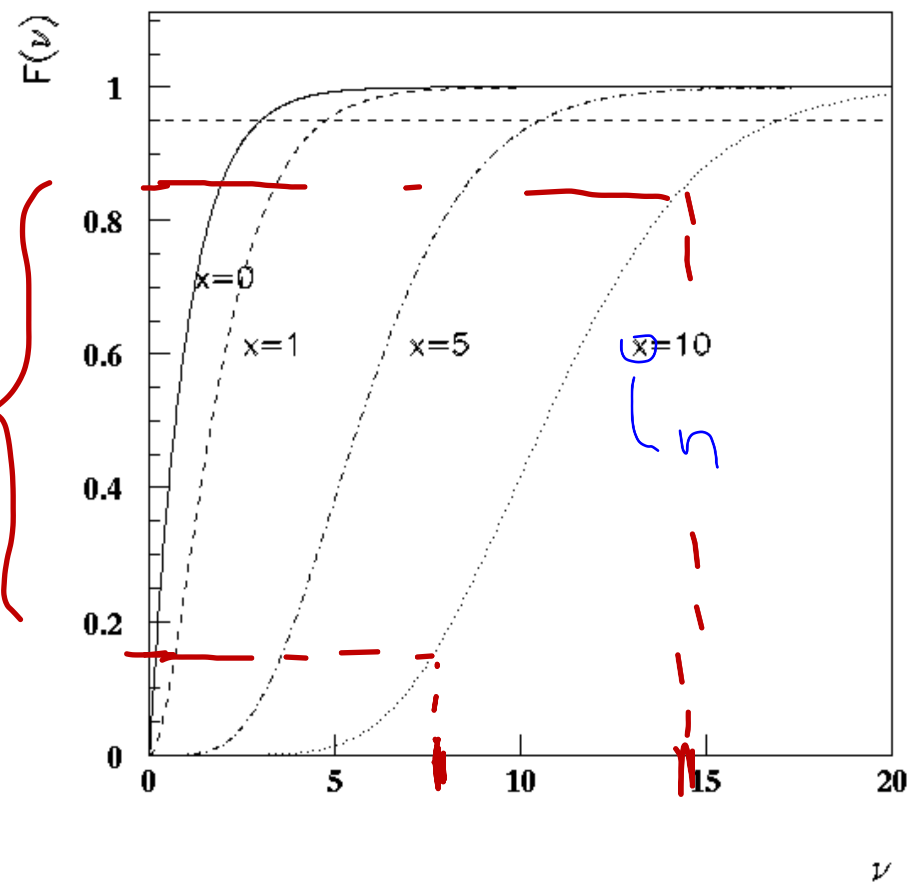
numerically same result as 95% CL

$\frac{\alpha}{2} = 0.16$

e.g. 68% central credibility interval

for $n=10$

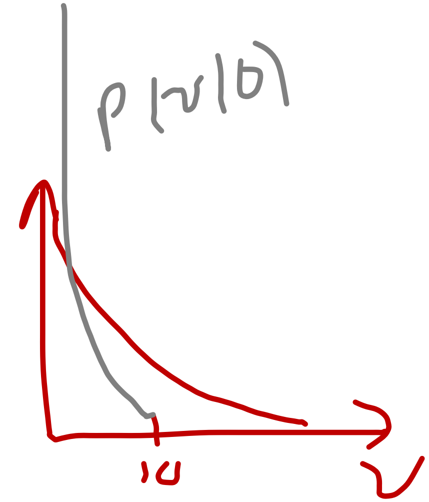
$7.5 \lesssim \nu \lesssim 14$



Poisson – cont.

What if we cannot (or do not want to) take a flat prior

$P_0(\nu)$



Suppose we can model the prior belief as $P_0(\nu) = \dfrac{1}{10} e^{-\nu/10}$

Now Bayes tells us $P(\nu \mid n = 0) = \dfrac{P(0 \mid \nu) P_0(\nu)}{\int_0^\infty P(0 \mid \nu) P_0(\nu) d\nu} = \dfrac{e^{-\nu} \frac{1}{10} e^{-\nu/10}}{\int_0^\infty \frac{1}{10} e^{-11\nu/10} d\nu} = \dfrac{11}{10} e^{-11\nu/10}$

$< \nu >= \int_0^\infty \dfrac{11}{10} e^{-11\nu/10} \nu d\nu = 0.91$

$P(\nu \le 2.7) = 95\%$, i.e., $\nu \le 2.7$ with 95% probability

Poisson Distribution-cont.

We often have to deal with a superposition of two Poisson processes – the signal and the background, which are indistinguishable in the experiment. Usually we know the background expectations and want to know the likelihood of a signal in addition.

Example, the signal for large extra dimensions may be the observation of events where momentum balance is (apparently) strongly violated. However this can be mimicked by neutrinos, energy leakage from the detector, etc.

Use the subscripts B for background, s for signal,

and assume n events are observed

signal background expectation $n = n_B + n_s$

$$P(n) = \sum_{n_s=0}^{n} P(n_s \mid \nu_s) P(n - n_s \mid \nu_B)$$

$$= e^{-(\nu_B + \nu_s)} \sum_{n_s=0}^{n} \frac{\nu_s^{\,n_s} \nu_B^{\,n-n_s}}{n_s!(n-n_s)!}$$

Binomial formula with $p = \left(\frac{\nu_s}{\nu_s + \nu_B} \right)$

$$= e^{-(\nu_B + \nu_s)} \frac{(\nu_s + \nu_B)^n}{n!} \underbrace{\sum_{n_s=0}^{n} \frac{n!}{n_s!(n-n_s)!} \left(\frac{\nu_s}{\nu_s + \nu_B} \right)^{n_s} \left(\frac{\nu_B}{\nu_s + \nu_B} \right)^{n-n_s}}_{=1 \text{ by normalization}}$$

$$= e^{-(\nu_B + \nu_s)} \frac{(\nu_s + \nu_B)^n}{n!}$$

=1 by normalization

Poisson distribution with expectation $\nu_B + \nu_s$

Example

Example for ν=10/3 $= 3.\bar{3}$

$1-\alpha \geq 0.9$

n	$P(n\|\nu)$	$F(n\|\nu)$	R	$F_R(n\|\nu)$
0	0.0357	0.0357	7	0.9468
1	0.1189	0.1546	5	0.8431
2	0.1982	0.3528	2	0.4184
3	0.2202	0.5730	1	0.2202
4	0.1835	0.7565	3	0.6019
5	0.1223	0.8788	4	0.7242
6	0.0680	0.9468	6	0.9111
7	0.0324	0.9792	8	0.9792
8	0.0135	0.9927	9	0.9927
9	0.0050	0.9976	10	0.9976
10	0.0017	0.9993	11	0.9993
11	0.0005	0.9998	12	0.9998
12	0.0001	1.0000	13	1.0000

$n^* = \lfloor \nu \rfloor = 3$

$\leq \alpha/2$ prob

$O^{c}_{0.9} = \{1, \dots, 7\}$

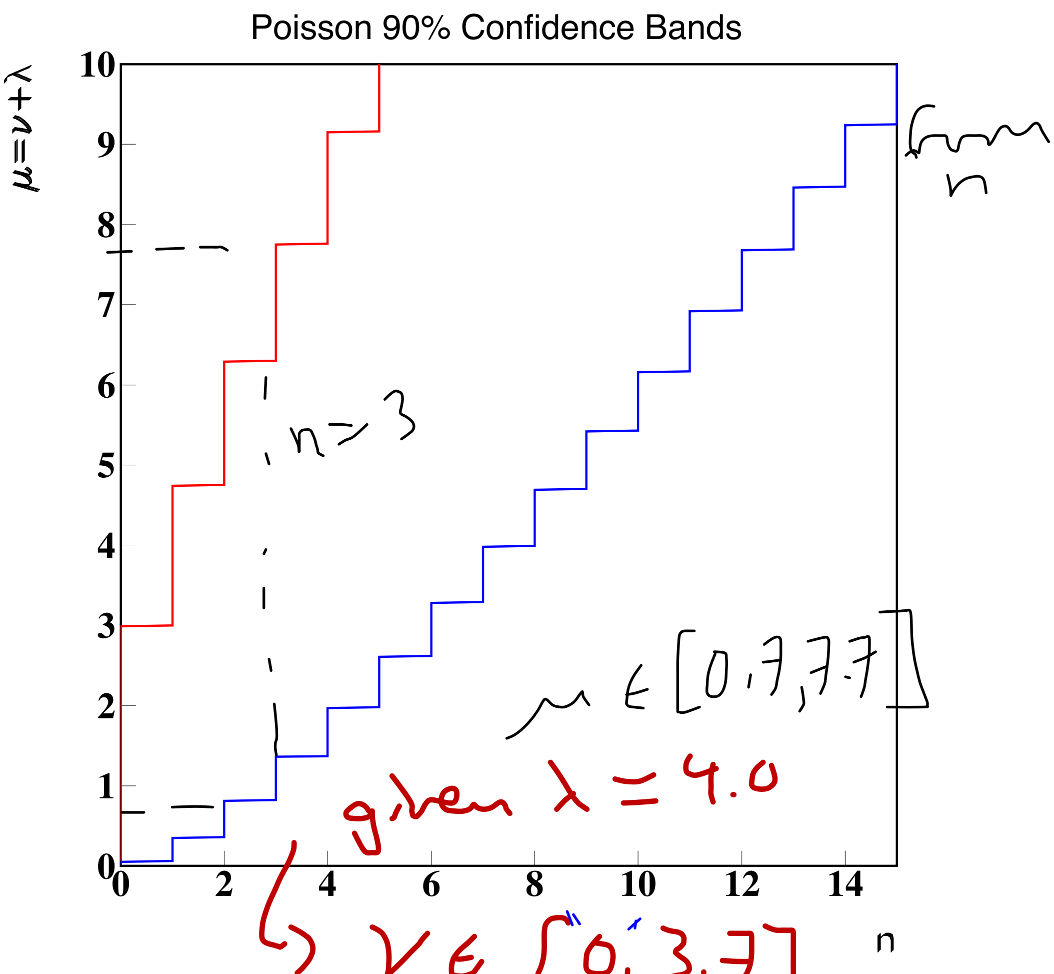
$O^{s}_{0.9} = \{1, \dots, 6\}$

Frequentist Statistics

λ for B
ν for signal

Poisson distribution in the presence of background, with mean λ. Then we have the same curves as for signal only, but replace ν with (ν+λ).

$\mu = \lambda + \nu$ total rate



- Traditional approach: find limit on μ, then subtract λ to get limit on ν

- limit for ν improves for a fixed n when we add background.

- can get negative limits ! For example, n=0, λ>3 gives ν<0.

Feldman-Cousins Confidence Levels

Imagine we have a Poisson process with known background expectation and unknown signal. If $\lambda \geq 3$ and $n = 0$ then the confidence interval for ν is empty (or includes unphysical values).

This has led to new definitions for the Confidence Intervals. The most popular (at least in particle physics) is the Feldman-Cousins construction, where a rank is assigned to possible outcomes based on

$$r = \frac{P(n|\mu = \lambda + \nu)}{P(n|\hat{\mu})}$$

Where $\hat{\mu}$ is the value of μ that maximizes $P(n|\mu)$ given the constraints.

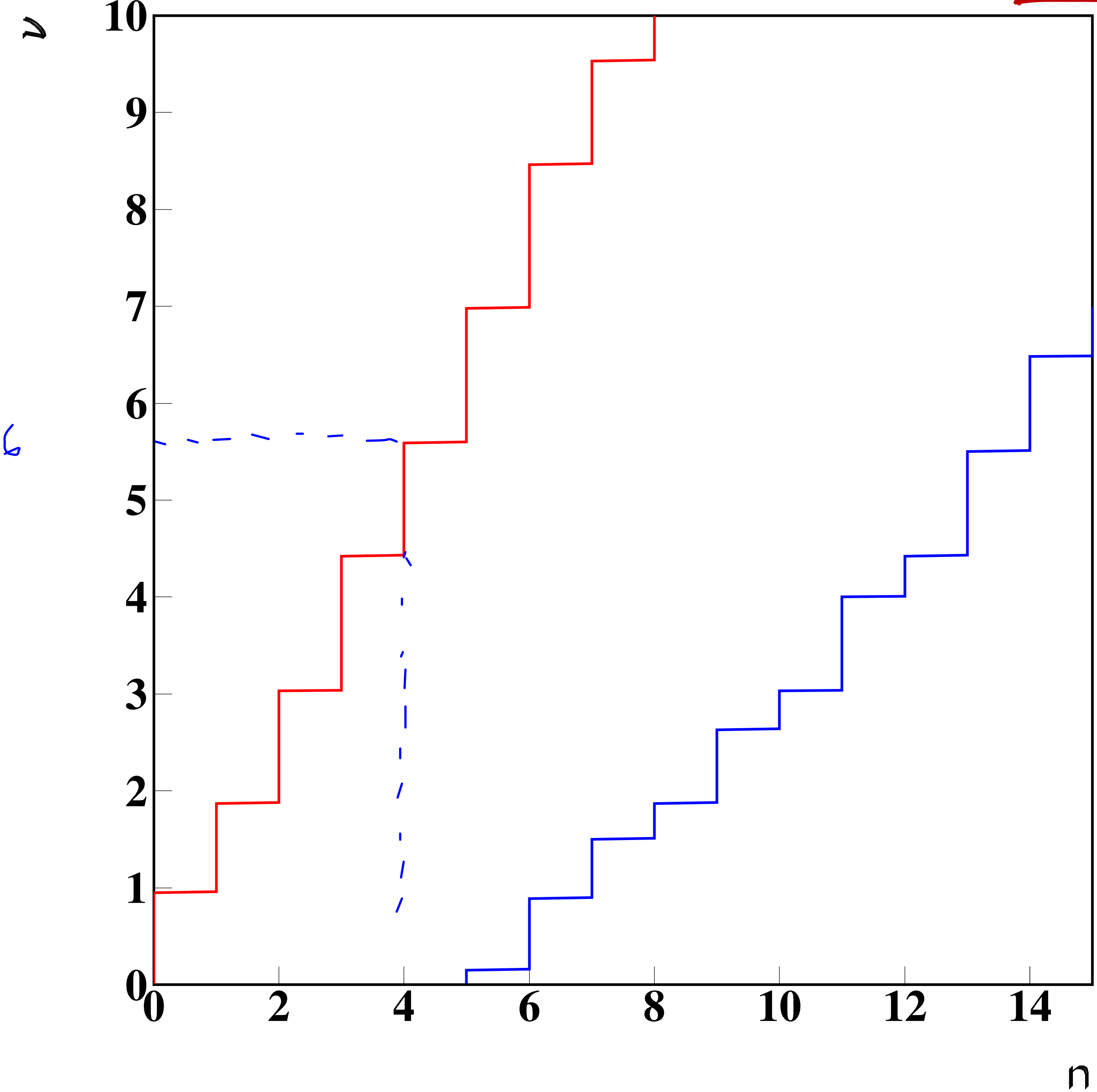
Concrete example: $\lambda = 3.0 \quad \nu = 0.\bar{3}$ $\mu = 10/3 = 3.\bar{3}$

n	$P(n\|\overset{\mu}{\cancel{\nu}})$	$\hat{\mu}$	$P(n\|\hat{\mu})$	r	Rank	$F_R(n\|\overset{\mu}{\cancel{\nu}})$
0	0.0357	3.0	0.050	0.717	5	0.7565
1	0.1189	3.0	0.149	0.796	4	0.7208
2	0.1982	3.0	0.224	0.885	3	0.6091
3	0.2202	3.0	0.224	0.983	1	0.2202
4	0.1835	4.0	0.195	0.941	2	0.4037
5	0.1223	5.0	0.175	0.699	6	0.8788
6	0.0680	6.0	0.161	0.422	7	0.9468
7	0.0324	7.0	0.149	0.217	8	0.9792
8	0.0135	8.0	0.140	0.096	9	0.9927
9	0.0050	9.0	0.132	0.038	10	0.9976
10	0.0017	10.0	0.125	0.014	11	0.9993
11	0.0005	11.0	0.119	0.004	12	0.9998

$\hat{\mu} = \max(n, \lambda)$

Procedure depends on λ

Poisson 90% CL Bands a la Feldman-Cousins for $\boxed{\lambda=3.0}$



5.6

$n = 4$

$\nu < 5.6$
@ 95%
CL

The Bayesian Way

$$\mu = \lambda + \nu \qquad P(n|\mu) = \frac{e^{-\mu}\mu^n}{n!}$$

Assuming (as before) that the background is perfectly known:

$$P(\nu|n,\lambda) = \frac{P(n|\nu,\lambda)P_0(\nu)}{\int P(n|\nu,\lambda)P_0(\nu)d\nu}$$

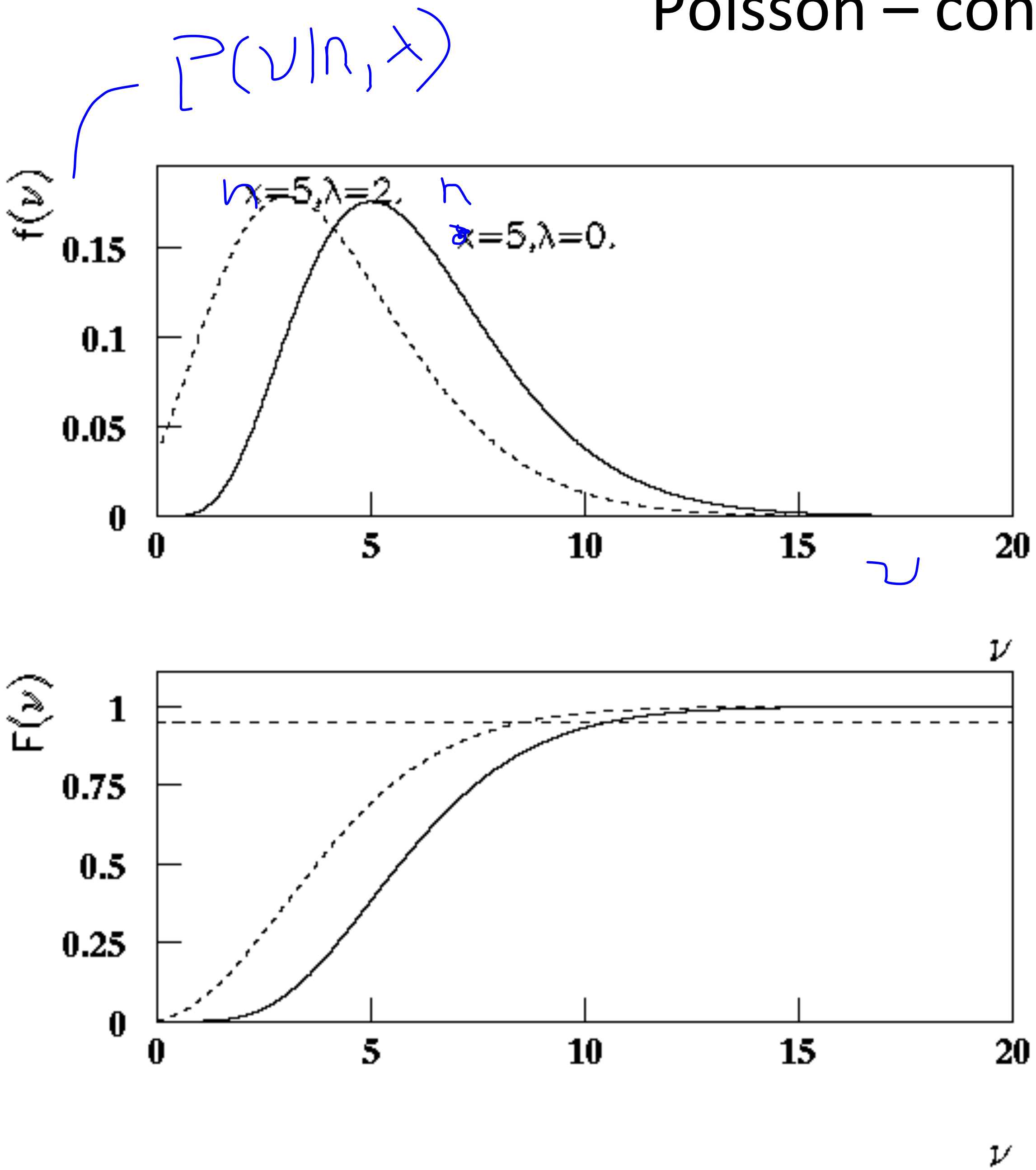
assuming a flat $P_0(\nu)$ and integrating by parts.

$$\boxed{P(\nu|n,\lambda) = \frac{e^{-\nu}(\lambda+\nu)^n}{n!\sum_{i=0}^{n}\frac{\lambda^i}{i!}}}$$

The cumulative pdf is

$$\boxed{F(\nu|n,\lambda) = 1 - \frac{e^{-\nu}\sum_{i=0}^{n}\frac{(\lambda+\nu)^i}{i!}}{\sum_{i=0}^{n}\frac{\lambda^i}{i!}}}$$

Poisson – cont.



Comment:

For n=0, $P(\nu|n, \lambda)=e^{-\nu}$. It does not matter how much background you have, you get the same probability distribution for the signal.

Comparing Feldman-Cousins with Bayesian Analysis with same background $\lambda = 3.0$ and a flat prior.

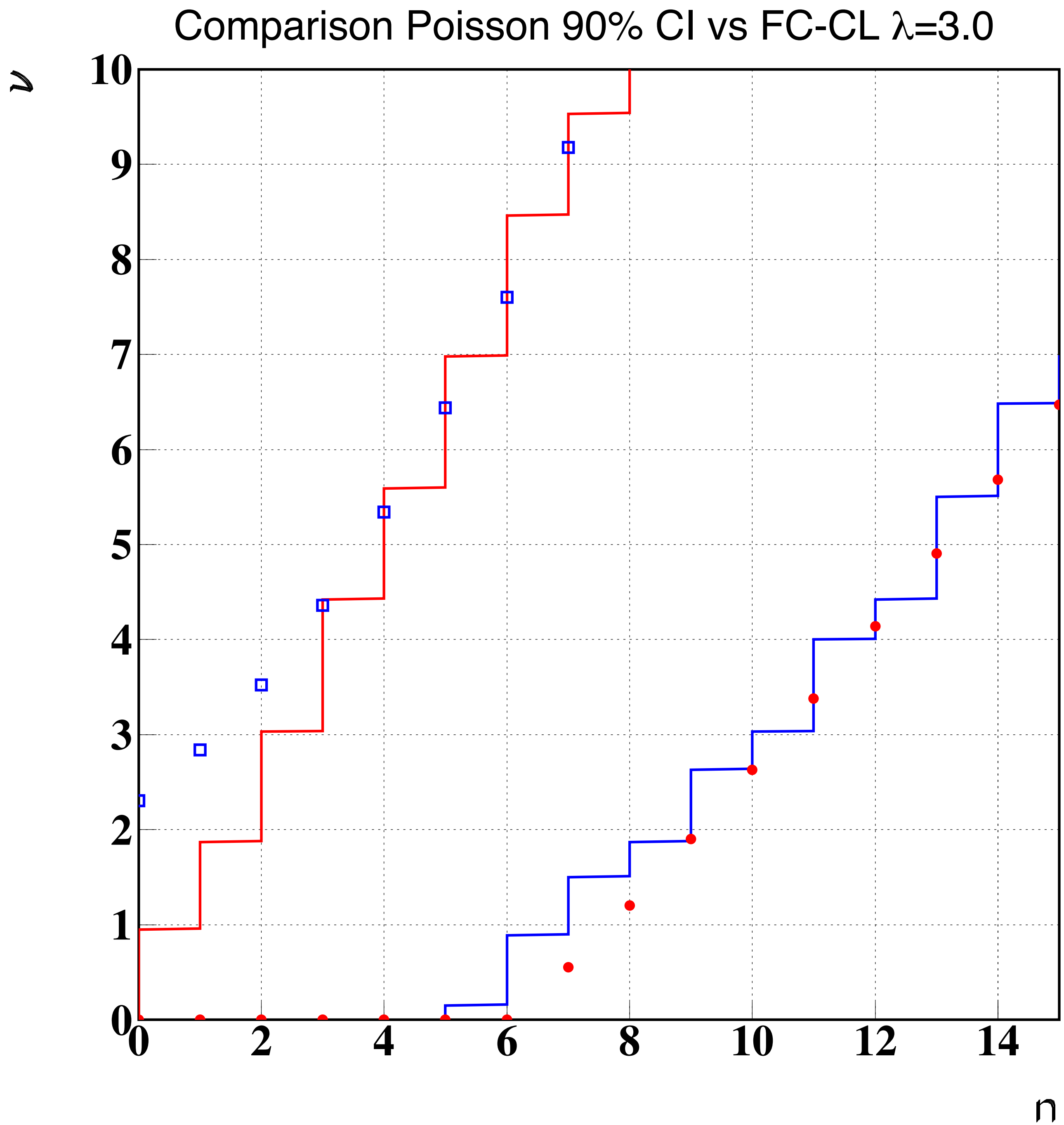
Recall: $P(\nu|n,\lambda) = \frac{e^{-\nu}(\lambda+\nu)^n}{n!\sum_{i=0}^{n}\frac{\lambda^i}{i!}}$

$$F(\nu|n,\lambda) = 1 - \frac{e^{-\nu}\sum_{i=0}^{n}\frac{(\lambda+\nu)^i}{i!}}{\sum_{i=0}^{n}\frac{\lambda^i}{i!}}$$

We will take the smallest interval with 90% credibility. I.e.,

$$\int_{P>C} P(\nu|n,\lambda)d\nu = 0.90$$

We find ν_{down} ν_{up} fulfilling this condition. Numerical integration.


Comparison Poisson 90% CI vs FC-CL λ=3.0
ν
0
1
2
3
4
5
6
7
8
9
10
0
2
4
6
8
10
12
14
n

Example

Probabilistic model:

$$P(n_B|\lambda) = \frac{e^{-\lambda}\lambda^{n_B}}{n_B!}$$

background

$$P(n_S|\nu) = \frac{e^{-\nu}\nu^{n_S}}{n_S!}$$

signal

$n = n_B + n_S$

$$P(n|\lambda,\nu) = \frac{e^{-\mu}\mu^{n}}{n!}$$

$$\mu = \lambda + \nu$$

combined is also Poisson process

Example

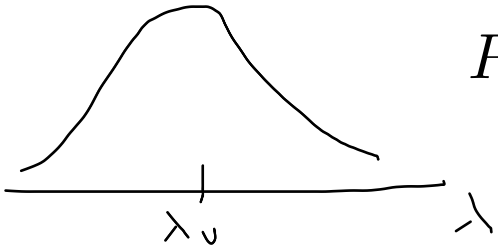
Compare two situations:

1) no knowledge on the background

2) Separate data help us constrain the background

Suppose we measure n=7 events, what can we say ?

With Background knowledge



$$P(\lambda) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} e^{-\frac{1}{2}\frac{(\lambda-\lambda_0)^2}{\sigma_\lambda^2}}$$

gleaned info from other expts, measurements

Can build this into the likelihood (e.g., frequentist analysis) or call it prior knowledge (either way for Bayes)

$$\mathcal{L}(\nu, \lambda) = P(n|\nu, \lambda)P(\lambda)$$

likelihood

Come back to likelihood analysis in ~2 lectures

With Background knowledge - Bayes

physics par

nuisance

Poisson

$P(\lambda,\nu) = P(\lambda)P(\nu)$

λ, ν independent

$$P(\nu,\lambda|n) = \frac{P(n|\nu,\lambda)P(\lambda)P(\nu)}{\int P(n|\nu,\lambda)P(\lambda)P(\nu)d\lambda d\nu}$$

$$P(n|\lambda,\nu) = \frac{e^{-(\lambda+\nu)}(\lambda+\nu)^n}{n!}$$

$$P(\lambda) = \underbrace{\frac{1}{\sqrt{2\pi}\sigma_\lambda}}_{A} e^{-\frac{1}{2}\frac{(\lambda-\lambda_0)^2}{\sigma_\lambda^2}}$$

$$P_0(\nu) = \text{constant}$$

require $\lambda \geq 0$, normalize $P(\lambda)$

We solve this numerically (here with the BAT package) https://www.mppmu.mpg.de/bat/

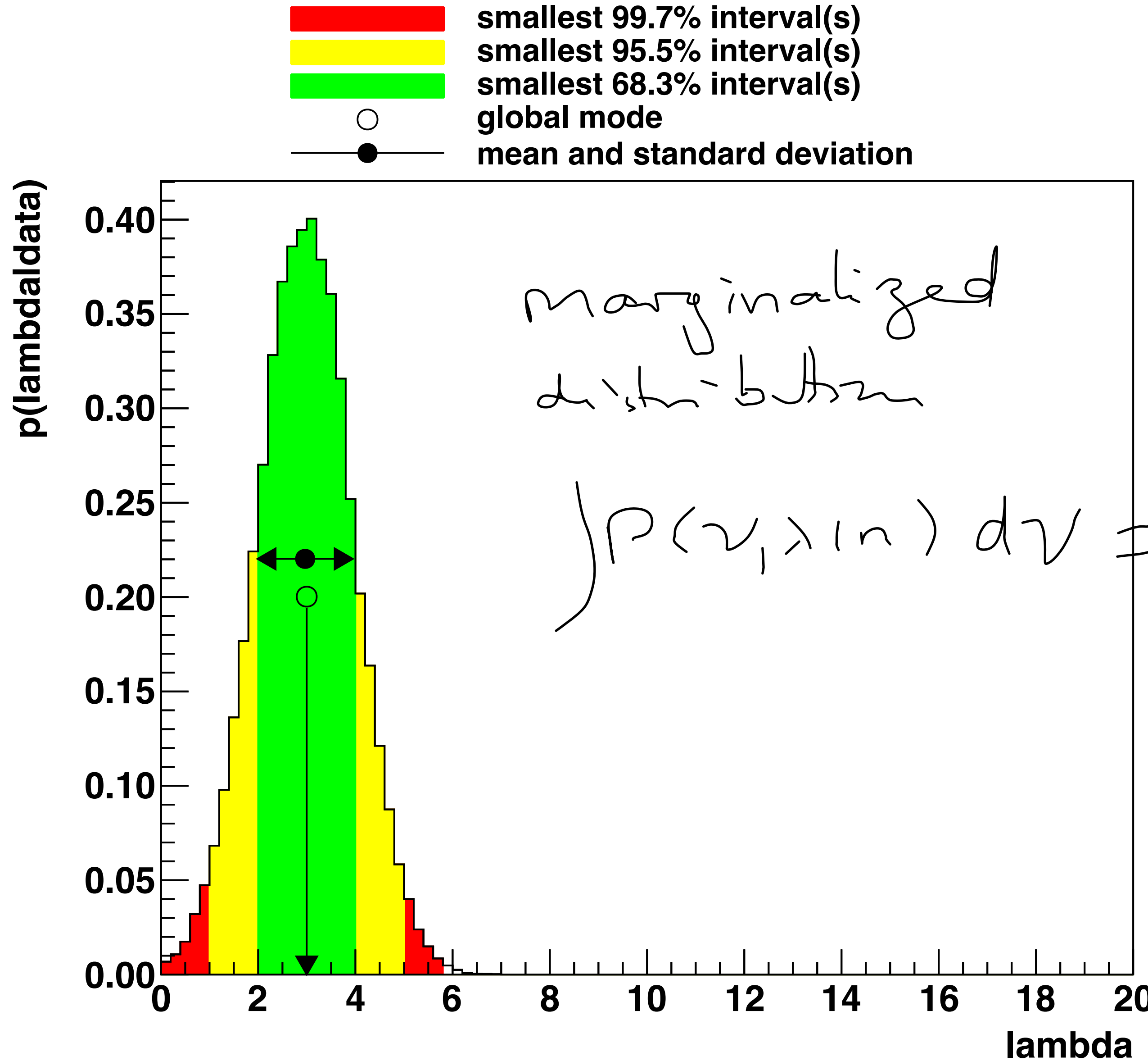
Bayesian Analysis Toolkit

To get a probability distribution for the physics parameter, we marginalize

$$P(\nu|n) = \int P(\nu,\lambda|n)d\lambda$$

Trivial in numerical program

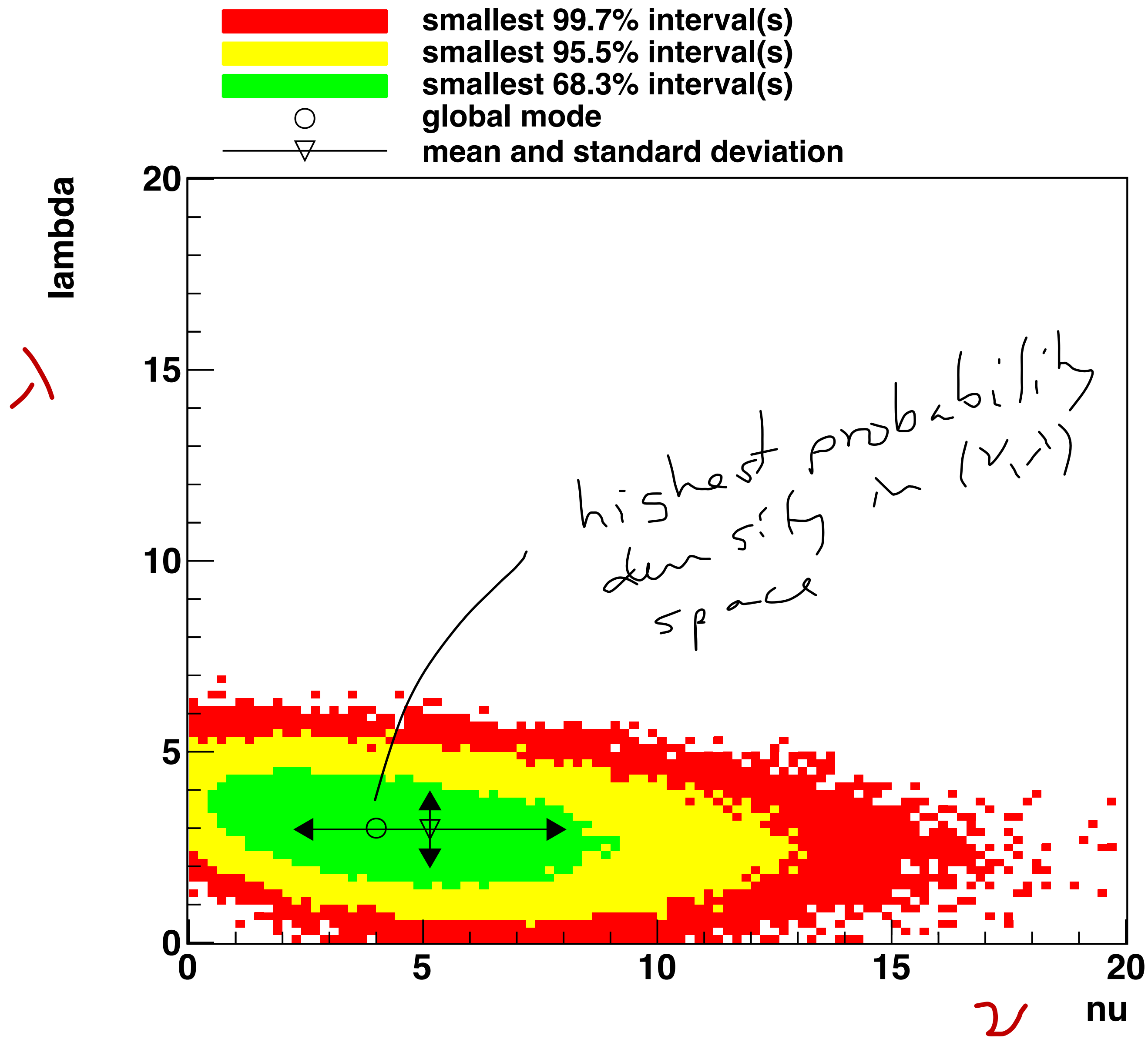
n=7 Constrained Background



marginalized distribution

$$\int P(\nu, \lambda | n)\, d\nu = P(\lambda | n)$$

n=7 Constrained Background

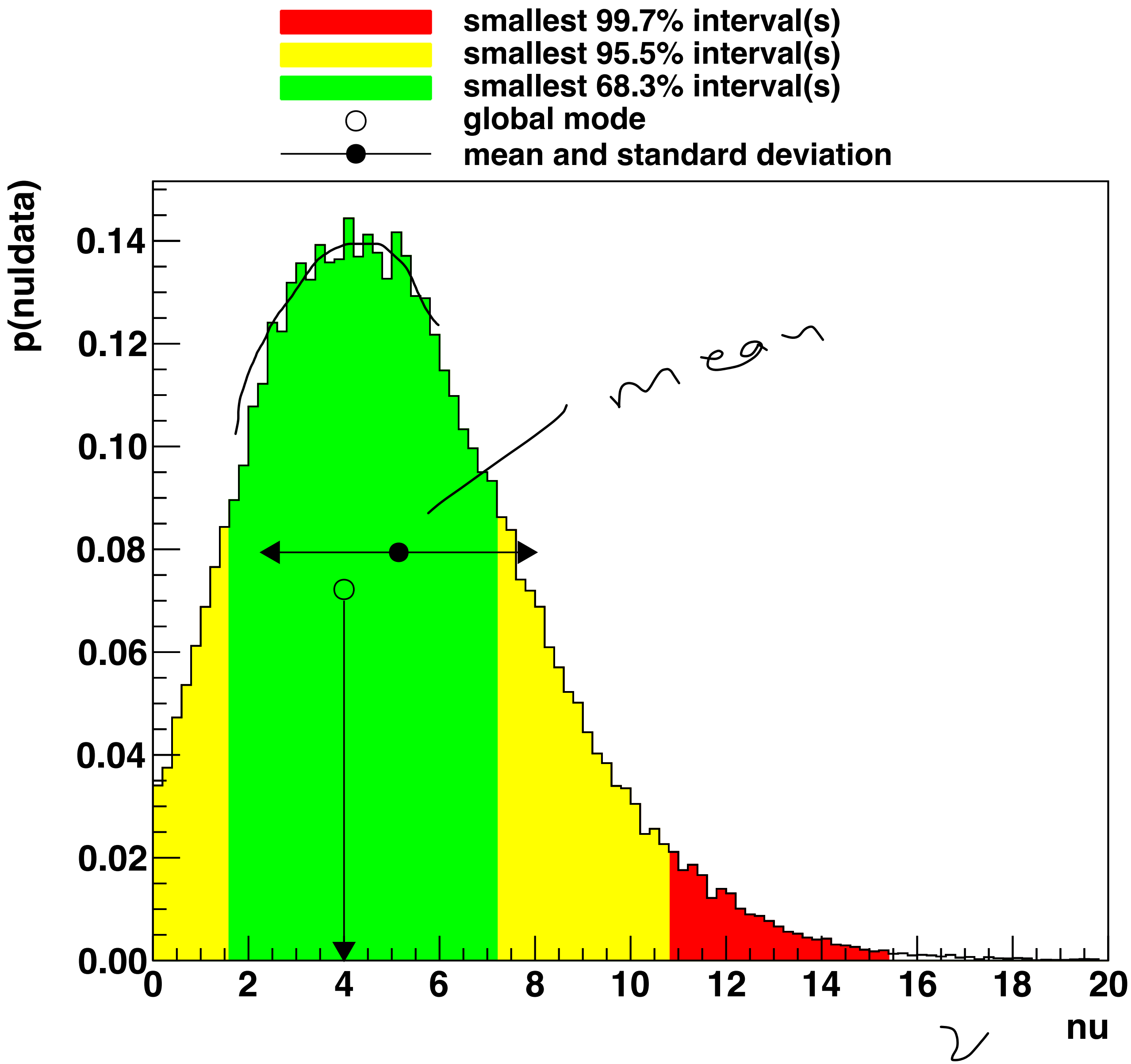


$\lambda_0 = 3$

$\sigma = 1$

$P(\nu, \lambda | n)$

n=7 Constrained Background



On/Off Example

As an example, we will consider measuring the decay rate for a radioactive isotope, in the presence of background. Prototype for off source/on source problem.

We take two measurements, one with the source absent, to measure the background rate, and once with the source present.

Data Set	Source in/out	Run Time	Events
1	Out	100	100
2	In	100	110

What can we say about the decay rate for our isotope ?

$$N = N_0 e^{-t/\tau} \quad \frac{dN}{dt} = -\frac{N}{\tau}$$

Radioactive Decay

Get an estimate of the background rate from the first data set. Assume we don't know very much. How do we represent this initial lack of knowledge ? Pick a simple form:

$$\underbrace{P_0(R_B)}_{\text{background rate}} = \text{constant}$$

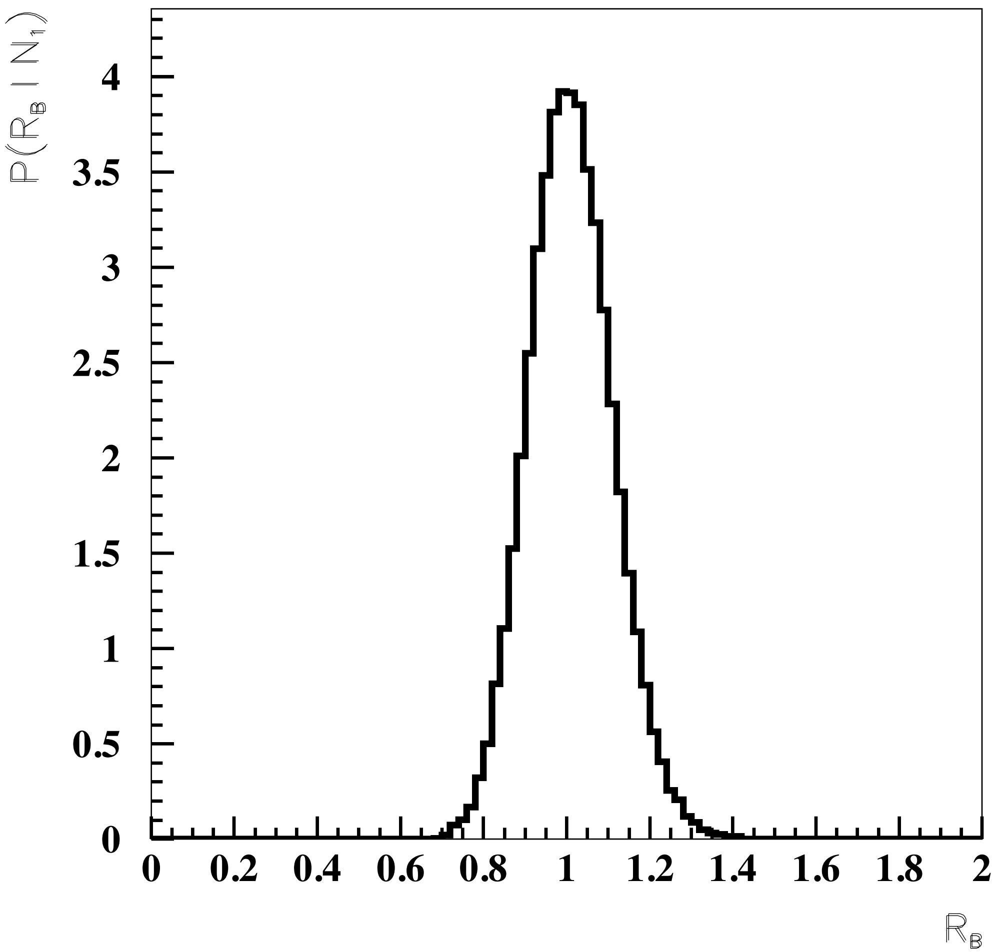
In this type of experiment, the number of counts in a time window follows a **Poisson Distribution**. For a flat prior, we found:

$$P(\lambda|N_1) = \frac{e^{-\lambda}\lambda^{N_1}}{N_1!} \qquad \lambda = R_B T$$

$$P(R_B)dR_B = P(\lambda)d\lambda = \frac{e^{-R_B T}(R_B T)^{N_1}}{N_1!} T dR_B$$

Radioactive Decay

Data Set	Source in/out	Run Time	Events
1	Out	100	100



$\boxed{\text{Poisson}}$

$\lambda^* = N$

$R^* = \frac{\lambda^*}{T} = \frac{N}{T} = 1$

$\sigma_\lambda^2 = N + 1$

$\sigma_\lambda = \sqrt{101} \simeq 10$

$\sigma_R = \frac{\sigma_\lambda}{T} \simeq 0.1$

Radioactive Decay

Now want to extract information on signal rate. We choose

$$P_0(R_B,R_S) = P_0(R_B)P_0(R_S) = \text{constant}, R_S > 0, R_B > 0$$

Analyze both data sets simultaneously

∴ const priors

$$P(R_B,R_S|N_1,N_2) \propto \underline{P(N_1,N_2|R_B,R_S)}$$

both signal and background can contribute

$$P(N_1,N_2|R_B,R_S) = P(N_1|R_B)P(N_2|R_B,R_S)$$

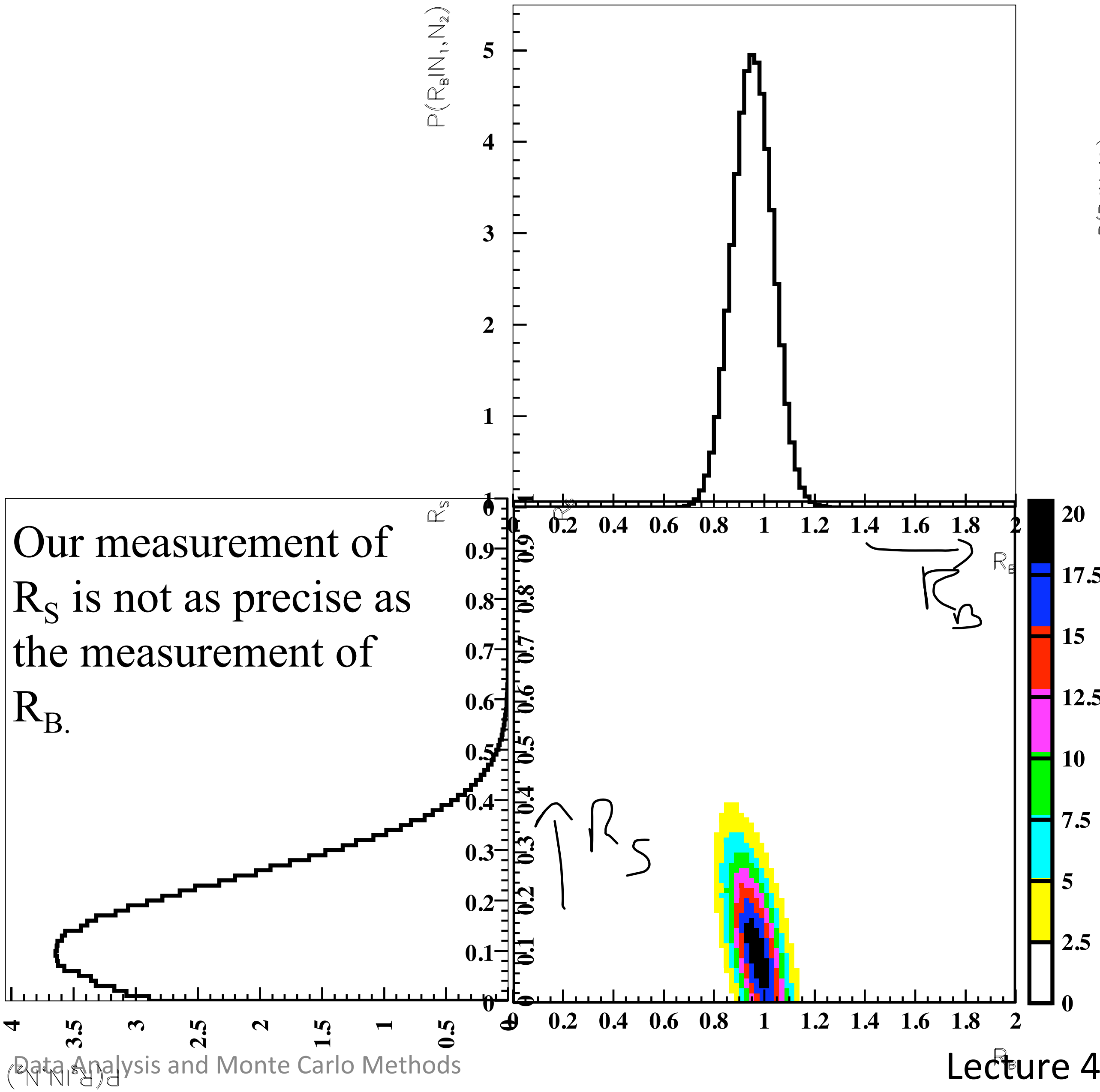
only background can contribute

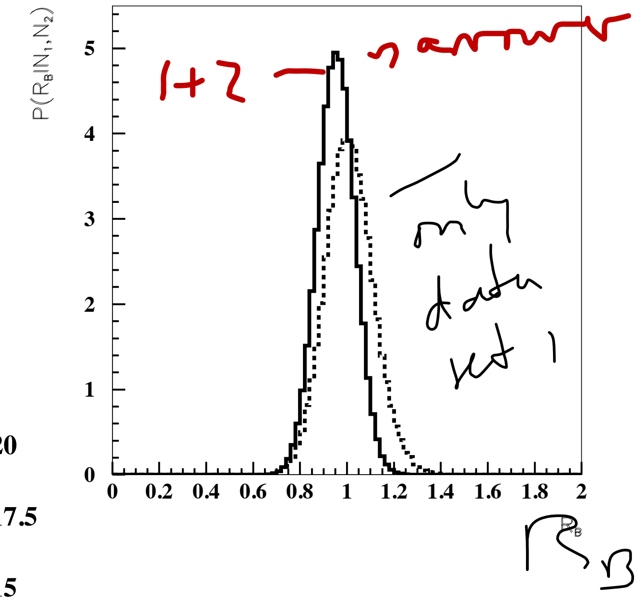
$$P(N_1,N_2|R_B,R_S) = \frac{(R_BT_1)^{N_1}e^{-R_BT_1}}{N_1!}\frac{((R_B+R_S)T_2)^{N_2}e^{-(R_B+R_S)T_2}}{N_2!}$$

Radioactive Decay

The measurement of R_B has improved.

Our measurement of R_S is not as precise as the measurement of R_B.





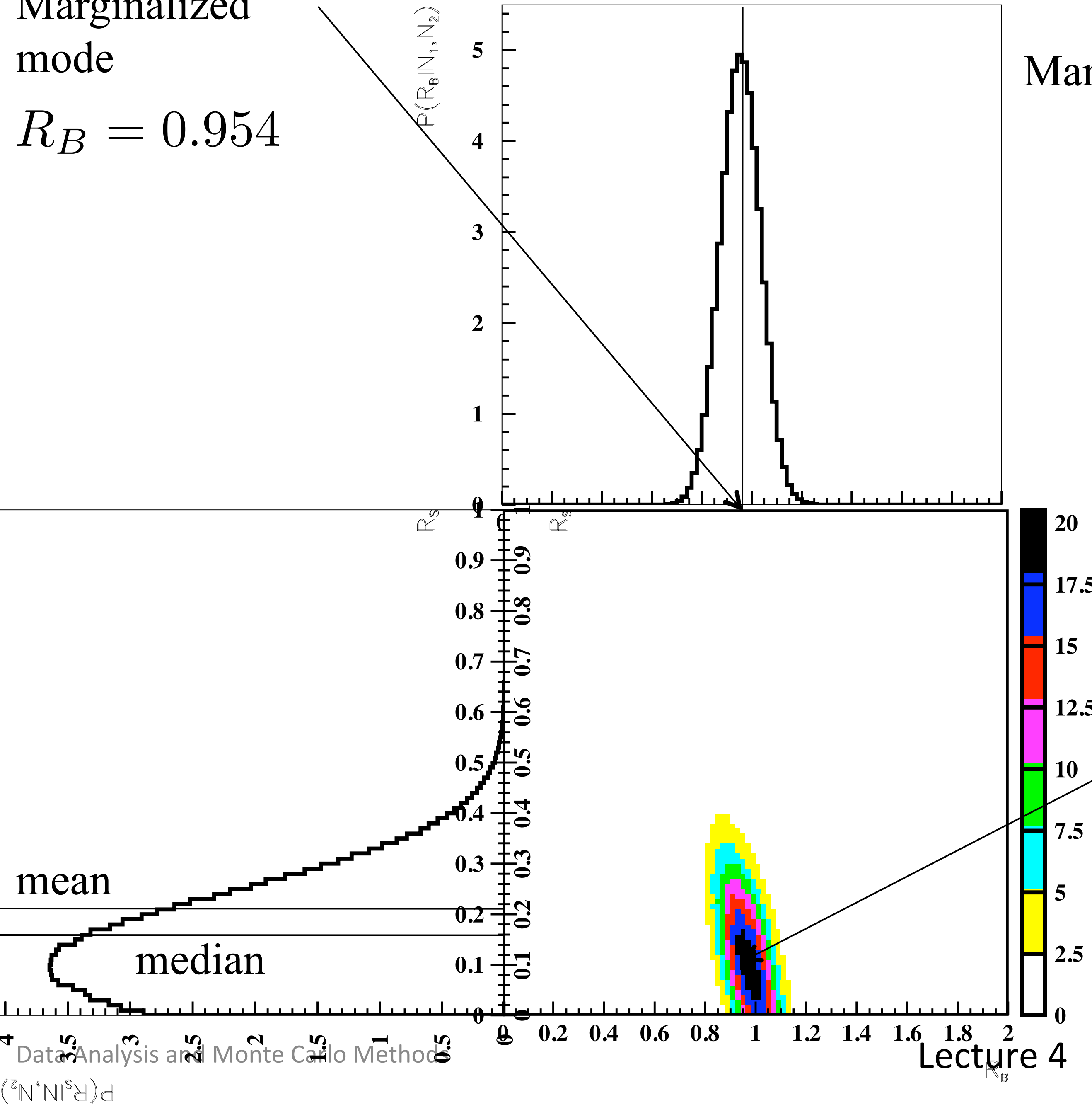
Data Set	Source in/out	Run Time	Events
1	Out	100	100
2	In	100	110

Radioactive Decay

Marginalized mode

$R_B = 0.954$

Marginalized distribution



mean

median

Global mode

$$R_B = 0.99999$$
$$R_S = 0.10011$$

Supernova 1987a

On the night of February 23, 1987 astronomers saw something they hadn't seen for 400 years... a **supernova explosion** close enough to be seen with the naked eye. A massive "blue giant" star, 50 times as large as our sun, had exploded in the Large Magellenic Cloud (a small suburb of our galaxy). The explosion actually occurred 170,000 years before... it took that long for the light to get here.

When a large star has burned up all of the nuclear fuel in it's center it becomes, in a few seconds, an almost empty shell and suddenly collapses. The rebounding matter and energy becomes a very dense, and very bright, source of light. Suddenly the object becomes hundreds of time brighter than its progenitor star.

The "before" and "after" pictures for SN1987a are shown below.

http://www-personal.umich.edu/~jcv/imb/imbp4.html

The number of neutrinos emitted is extremely large... about **10^57** escape in a few seconds. After 170,000 years this pulse of neutrinos is spread out over the surface of a sphere 170,000 light-years in radius.... big enough to encompass our **whole galaxy.** Spreading out the 10^57 neutrinos over the surface of this huge sphere gives 10^13 neutrinos per square meter. All of the neutrinos are contained in a thin shell on the surface. The shell is only a few light-seconds thick (about the distance from here to the moon).

Of the 10^16 neutrinos that went through the IMB tank only 8 interacted with enough energy to be detected.... all near the lower limit of our energy threshold. The normal rate of events from atmospheric neutrinos at these low energies was only about one per week, so seeing 8 in a few seconds meant something truly unique had happened.

Compare:

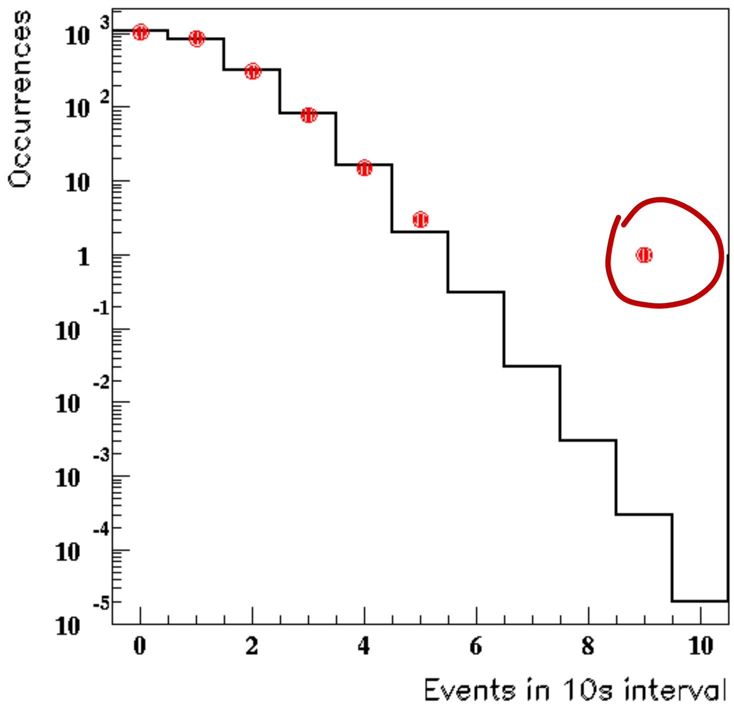
Did not see the light flash, no other detector recorded a signal. How significant is the results ?

Light observed (see pictures) – now what is the significance ?

IMB Observations

Example: Observation of Supernovae – IMB experiment

Number of events in 10 sec interval:	0	1	2	3	4	5	6	7	8	9
Frequency	1042	860	307	78	15	3	0	0	0	1
Poisson with mean 0.77	1064	823	318	82	16	2	0.3	0.03	0.003	0.0003



Significant?

Discovery?

IMB Observations

$$P(n \geq 9|0.77) = \sum_{n=9}^{\infty} \frac{e^{-0.77} 0.77^n}{n!} = 1.3 \cdot 10^{-7}$$

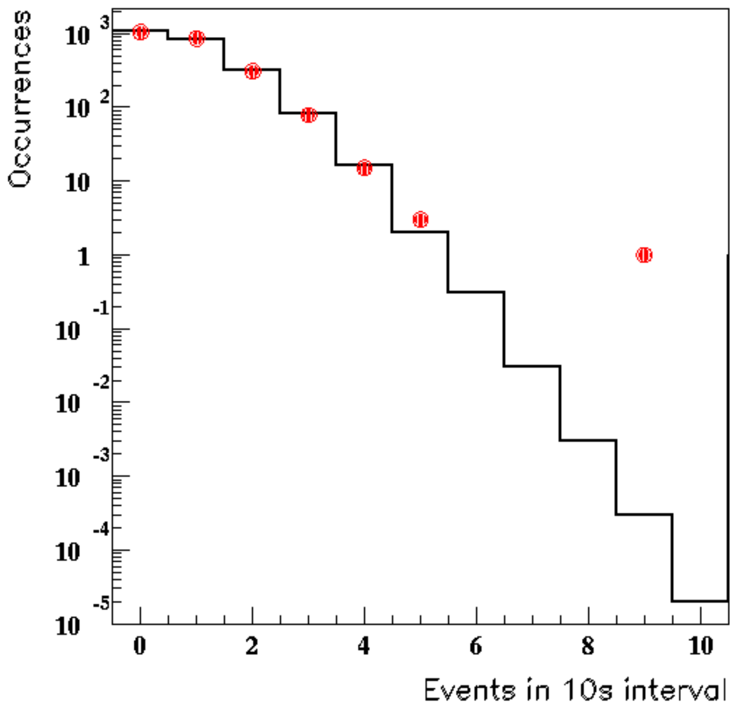
2306 10s intervals analyzed.

$$\begin{aligned}
P(r \geq 9|0.77, 2306 \text{ trials}) &= 1 - P(r < 9|0.77, 2306) \\
&= 1 - P(r < 9|0.77)^{2306} \\
&= 1 - (1 - 1.3 \cdot 10^{-7})^{2306} \\
&\approx 2306 \cdot 1.3 \cdot 10^{-7} \\
&= 3 \cdot 10^{-4}
\end{aligned}$$

Test Statistic and p-value

Example: Observation of Supernovae – IMB experiment

Number of events in 10 sec interval:	0	1	2	3	4	5	6	7	8	9
n_i	1042	860	307	78	15	3	0	0	0	1
ν_i	1064	823	318	82	16	2	0.3	0.03	0.003	0.0003



Test-statistic: a scalar quantity that summarizes the data. E.g.,

$$\eta = \log\left[\prod_{i=0}^{10} \frac{e^{-\nu_i}\nu_i^{n_i}}{n_i!}\right]$$

$$\nu_i = 2306 \cdot \frac{e^{-0.77}(0.77)^i}{i!}$$

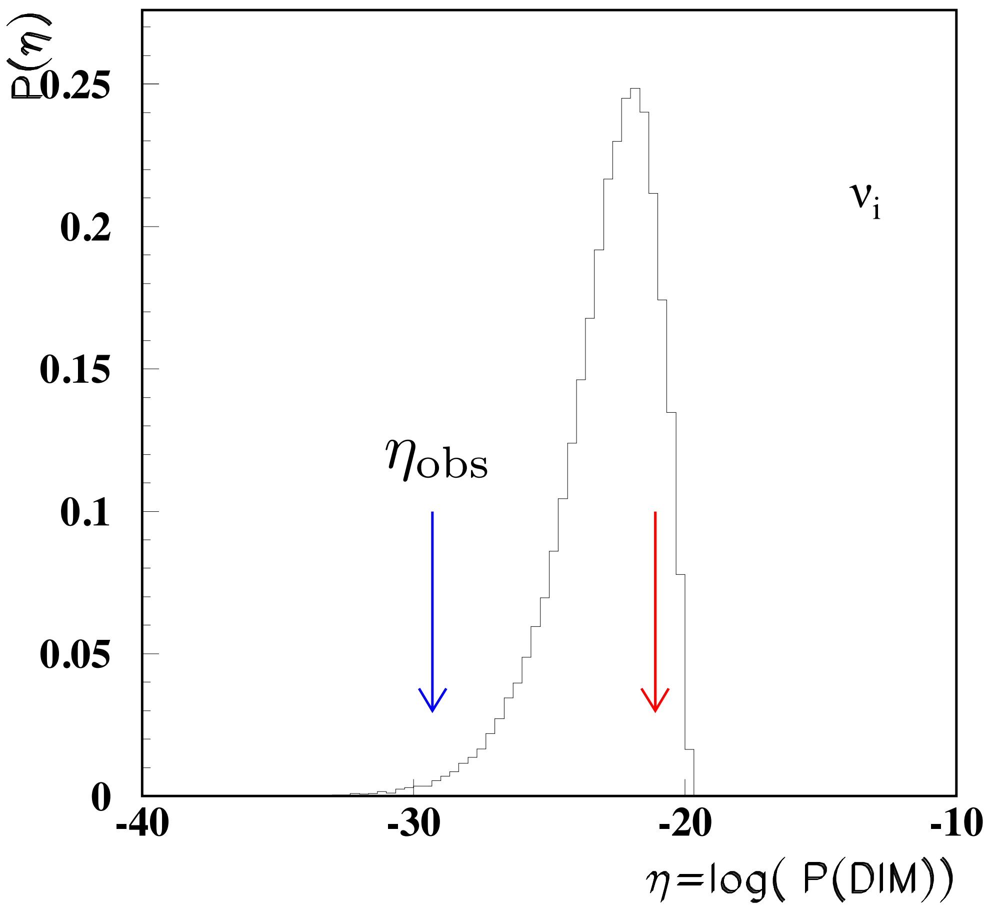
expectation for i events

n_i observations for i events

Distribution of test statistic

Null Hypothesis:
H_0 The bin contents are distributed according to Poisson processes with means ν_i

Assuming H_0, we can make the distribution of what we expect for η. Here it is:



The histogram is from a large number of simulations of the experiment.

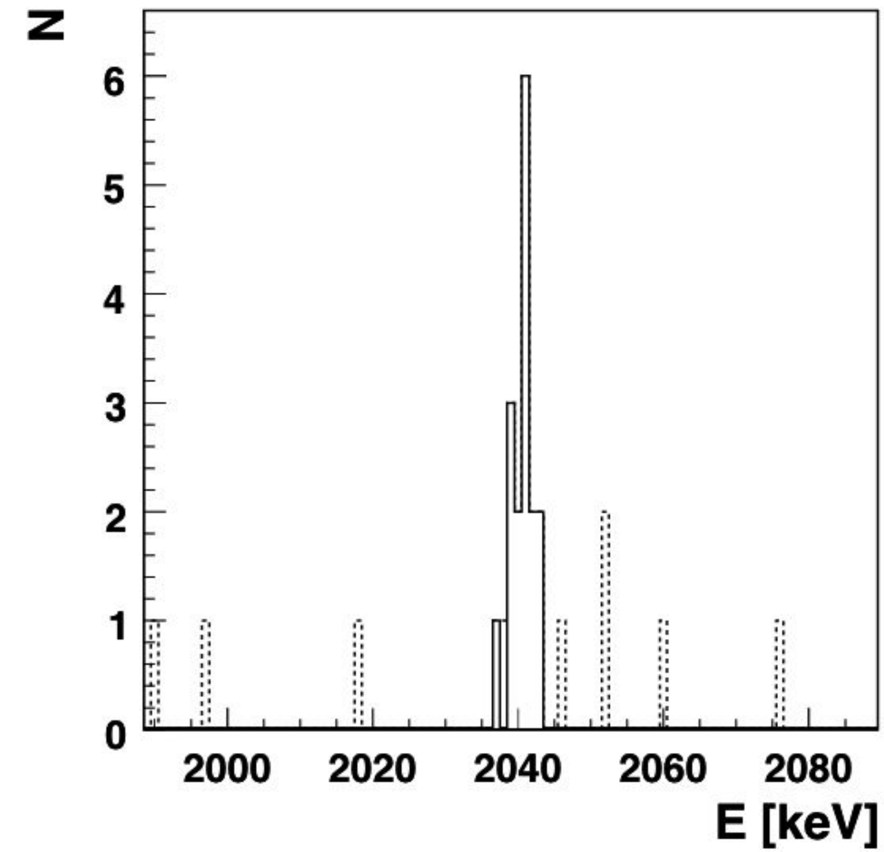
The blue arrow shows the observed value of η

The red arrow shows the value of η if we leave out the observation of the 9 events.

$$p = \int_{\eta_{\rm min}}^{\eta_{\rm obs}} P(\eta|H_0)d\eta$$

$$p = 0.007$$

Discovery or not ?



Data = collection of energy intervals with n_i in each

Analyze energy spectrum and decide if there is evidence for a signal.
Counting experiment – Poisson statistics.

Double Beta Decay Example

Prior:

The existing limits are $T_{1/2}>4\ 10^{25}$ yr; a positive claim for a signal exists at the level $T_{1/2}=1.2\ 10^{25}$ yr; my favorite theorist believes strongly that neutrinos are Majorana particles, but he wont tell me the neutrino mass; the theorist at a neighboring university says that he believes strongly in Leptogenesis, and in that context the neutrino is a Majorana particle but it must be very light, such that neutrinoless double beta decay is unobservable,...

Two models:

H Data comes from background processes only.
Background rate uncertain

$\bar{H}$ Data comes from signal + background processes
Background and signal rate uncertain

DBD example

prior to specify uncertainty in background

$$P(Data\,|\,H) = \int P(Data\,|\,B)P_0(B)dB$$

$$P(Data\,|\,\overline{H}) = \int P(Data\,|\,S,B)P_0(S)P_0(B)dBdS$$

prior for signal

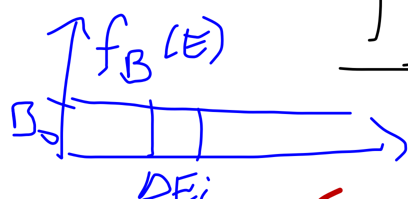
$n_i =$ observed number of events in bin i

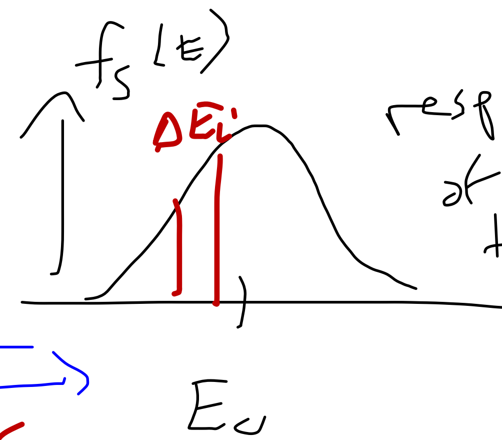
$\lambda_i =$ expected number of events in bin i

response function

$$\lambda_i = S \int_{\Delta E_i} f_S(E)dE + B \int_{\Delta E_i} f_B(E)dE$$

total signal strength





response of detector to input E_0

$\int f_S(E)\,dE = 1$

Where f_S and f_B are the normalized signal and background probability densities as functions of energy.

DBD example

total # energy intervals

S=0 for Background only hypothesis

then

$$P(Data\,|\,B) = \prod_{i=1}^{N} \frac{\lambda_i(0,B)^{n_i}}{n_i!} e^{-\lambda_i(0,B)}$$

$$P(Data\,|\,S,B) = \prod_{i=1}^{N} \frac{\lambda_i(S,B)^{n_i}}{n_i!} e^{-\lambda_i(S,B)}$$

To determine parameter values or set limits, we need

$$P(S,B\,|\,Data) = \frac{P(Data\,|\,S,B)P_0(S)P_0(B)}{\int P(Data\,|\,S,B)P_0(S)P_0(B)\,dS\,dB}$$

and then marginalize

$$P(S\,|\,Data) = \int P(S,B\,|\,Data)\,dB$$

e.g., 90% probability upper limit, S_{90} from solving

$$\int_0^{S_{90}} P(S\,|\,Data)\,dS = 0.90$$

GERDA example

Assumptions for GERDA:

$$P_0(H) = P_0(\overline{H}) = 1/2$$

$$P_0(S) = \frac{1}{S_{\max}} \quad 0 \le S \le S_{\max} \quad P_0(S) = 0 \text{ otherwise}$$

$$P_0(B) = \frac{e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}}}{\int_0^\infty e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}} dB} \quad B \ge 0; \quad P_0(B) = 0 \; B < 0$$

background flat shape but uncertain level

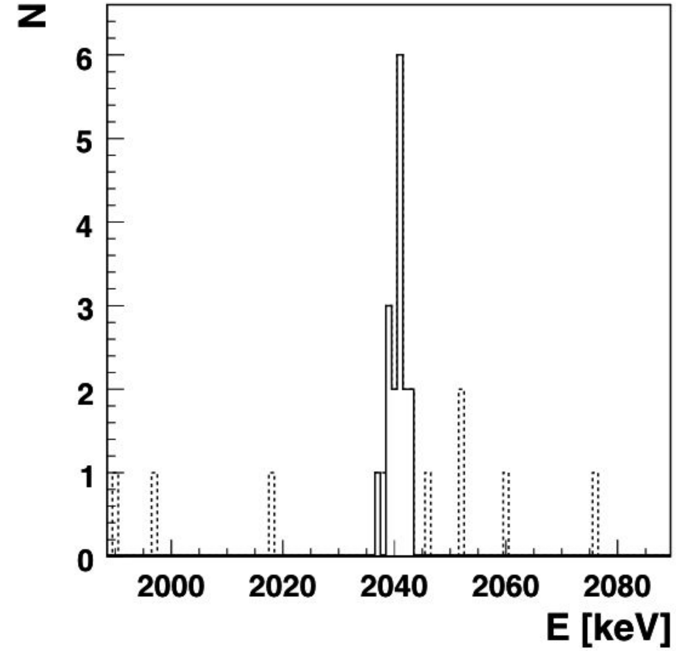
S_{max} was calculated assuming $T_{1/2}$=0.5 10^{25} yr

$\mu_B = B_0$, $\sigma_B = B_0/2$

100 keV window analyzed. B_0 total background in this window.

Example:

S_{true}=16, B_{true}=9

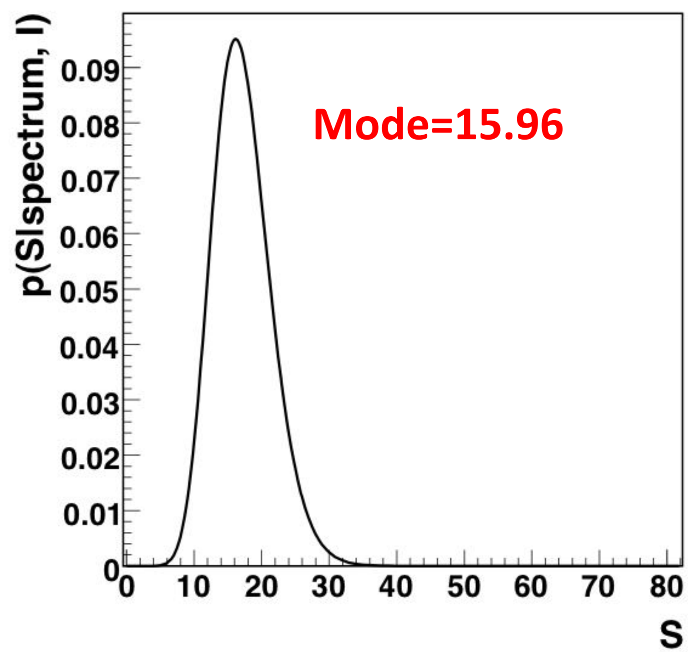


$$P(H \mid Data) = 2.2 \cdot 10^{-12}$$

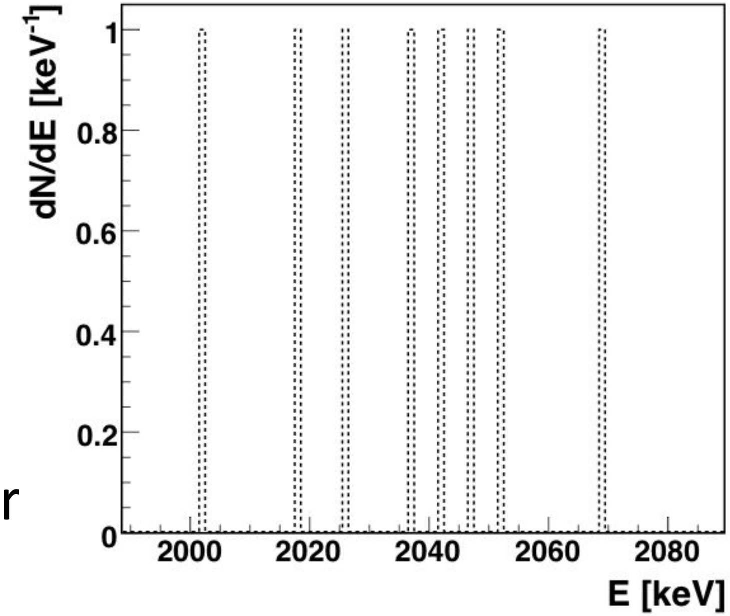
background only

Discovery !

$P(\bar{H} \mid Data) = 1 - P(H \mid data)$



Example:



$S_{true}=0$, $B_{true}=8$

10^{-3}/(kg keV yr)
Exposure 100 kg-yr

$$\boxed{P(H \mid Data) = 0.93}$$

$$P(\bar{H} \mid Data) = 0.07$$

No discovery – belief in background only enhanced

(started 50/50)

Upper limit on number of signal events:

$$\boxed{S_{90} = 3.99}$$