Intervals and Coverage

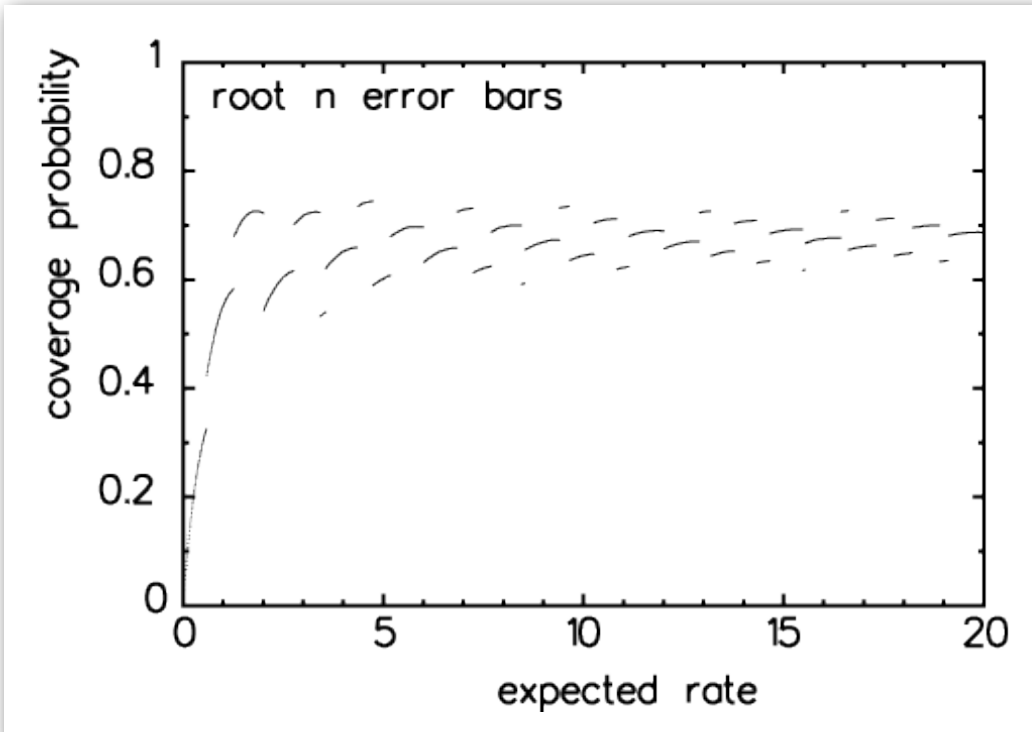
Physics 252C - Lecture 7

Prof. John Conway

coverage

- basic question: how often is the true value within the 1-sigma error bars?
- example: Poisson-distributed numbers
- use $\sqrt{n}$ error bars

what are the strange jumps?



jumps in coverage

- discrete nature of Poisson distribution causes jumps
- expected number μ is continuous
- only possible outcomes are integers: 0, 1, 2, ...
- for each integer outcome there is a given error which is the square root of that integer
- we only count that integer outcome's probability toward the coverage if it is within 1 σ of μ
- coverage probability is sum of individual outcome probabilities which are within 1 σ of μ

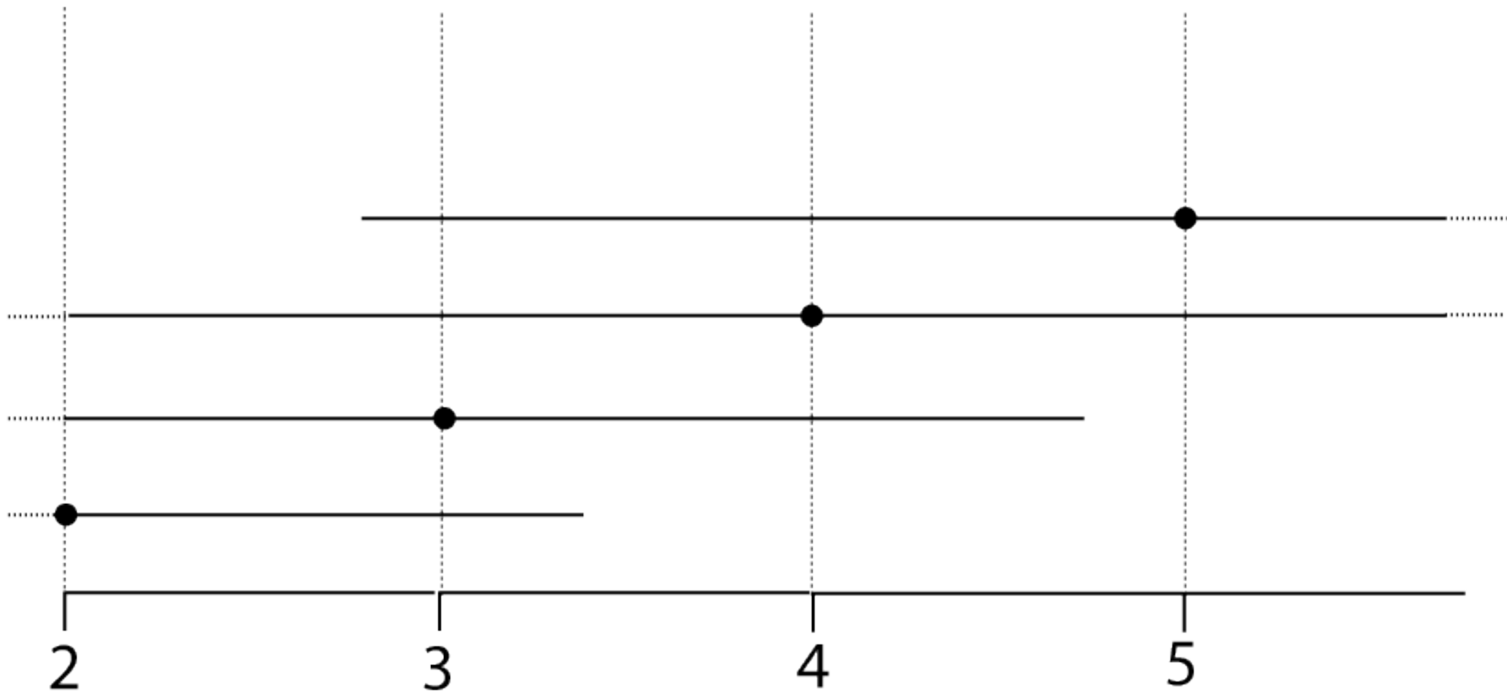
code for coverage

- simple to calculate coverage:

```
for( int i = 0; i<1000; i++)
 {
   x = 0.01 * i;
   p = 0.;

   for(int n = 0; n < 40; n++)
    {
      double xn = n;
      s = sqrt(xn);
      if(fabs(xn-x) < s) p = p + poiss_prob(x,n);
    }

   cout << x << "  " << p << endl;

 }
```

graphical attempt at explanation of jumps

- as μ increases, can suddenly add a new integer outcome’s probability, or drop one...

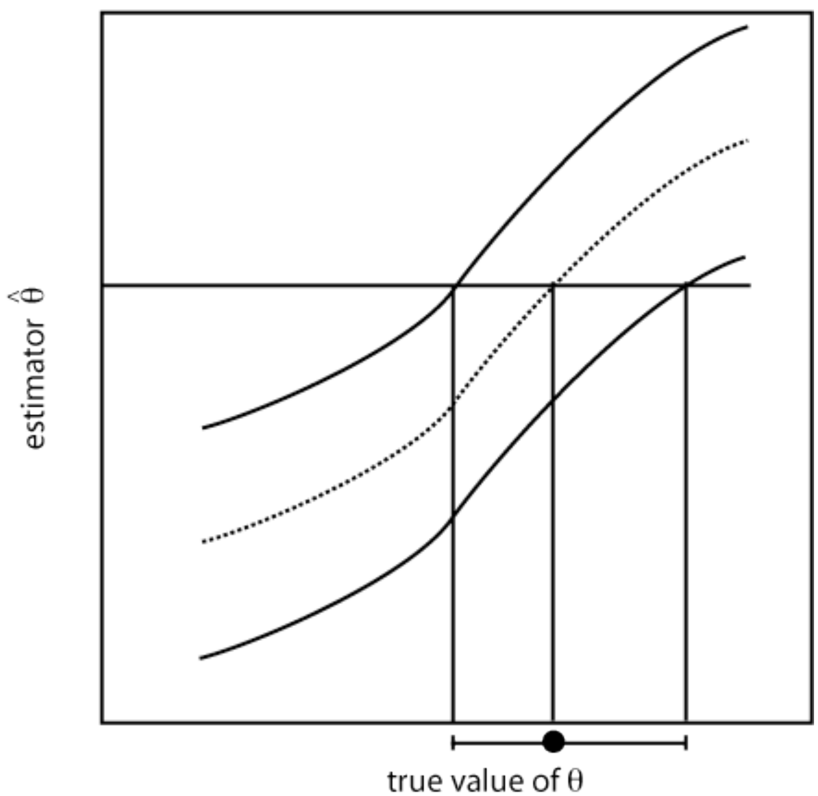


why do we care about coverage?

- we wish to estimate the values of parameters
- in addition we specify a range in which the true value might lie: the uncertainty in the parameter
- we can alternatively (or equivalently) quote a confidence interval
- typically we demand that a confidence interval contain the true value with some known probability, like 68.27% or 95%
- as we now see, this can be problematic in the region of small statistics ("Poisson regime")
- coverage is a manifestly frequentist concept!

Neyman construction

- relatively recent - 1937 !
- the idea is that we may not know the true value of our parameter θ, but if we do know what the distribution of our estimator is, we can construct a “confidence belt”



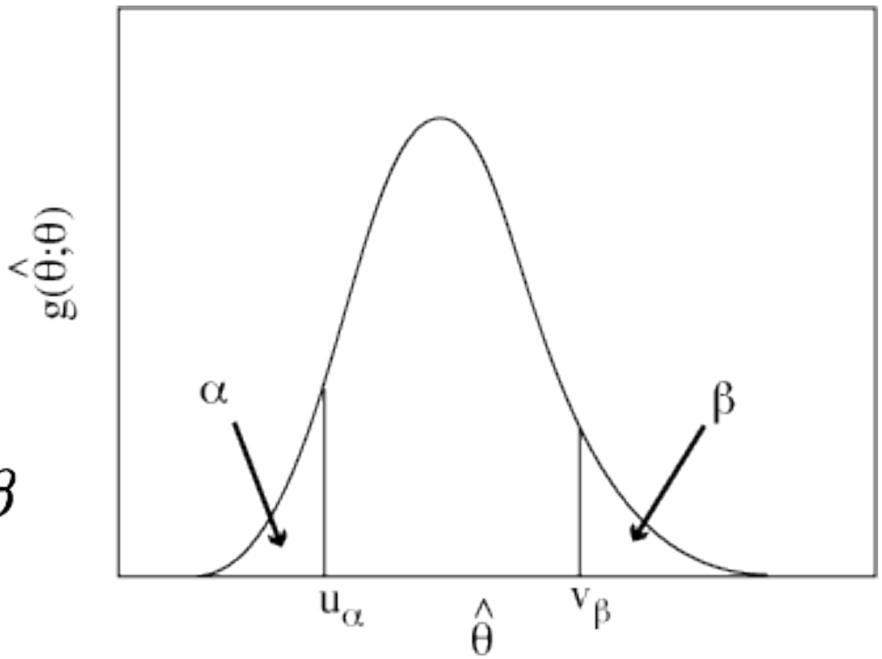
using distribution of estimator, take slice through belt to determine interval

Neyman construction

- consider vertical "slice" through belt; described by distribution $g(\hat{\theta};\theta)$

$$\begin{aligned}\mathcal{P}(\hat{\theta} < u_\alpha;\theta) &= \alpha \\ \mathcal{P}(\hat{\theta} < v_\beta;\theta) &= \beta\end{aligned}$$

$$\mathcal{P}(u_\alpha < \hat{\theta} < v_\beta) = 1 - \alpha - \beta$$



next find inverses:

$$a(\hat{\theta}) = u_\alpha^{-1}(\hat{\theta})$$
$$b(\hat{\theta}) = v_\beta^{-1}(\hat{\theta})$$

Neyman construction

- then, the key to understanding the Neyman construction is this:

$$\begin{aligned} \hat{\theta} \geq u_\alpha(\theta) \quad &\Rightarrow \quad a(\hat{\theta}) \geq \theta \\ \hat{\theta} \leq v_\beta(\theta) \quad &\Rightarrow \quad b(\hat{\theta}) \leq \theta \end{aligned}$$

- therefore the probabilities are the same

$$\mathcal{P}(a(\hat{\theta}) \geq \theta) = \alpha$$
$$\mathcal{P}(b(\hat{\theta}) \leq \theta) = \beta$$

- and thus

$$\mathcal{P}(a(\hat{\theta}) \leq \theta \leq b(\hat{\theta})) = 1 - \alpha - \beta$$

Neyman construction

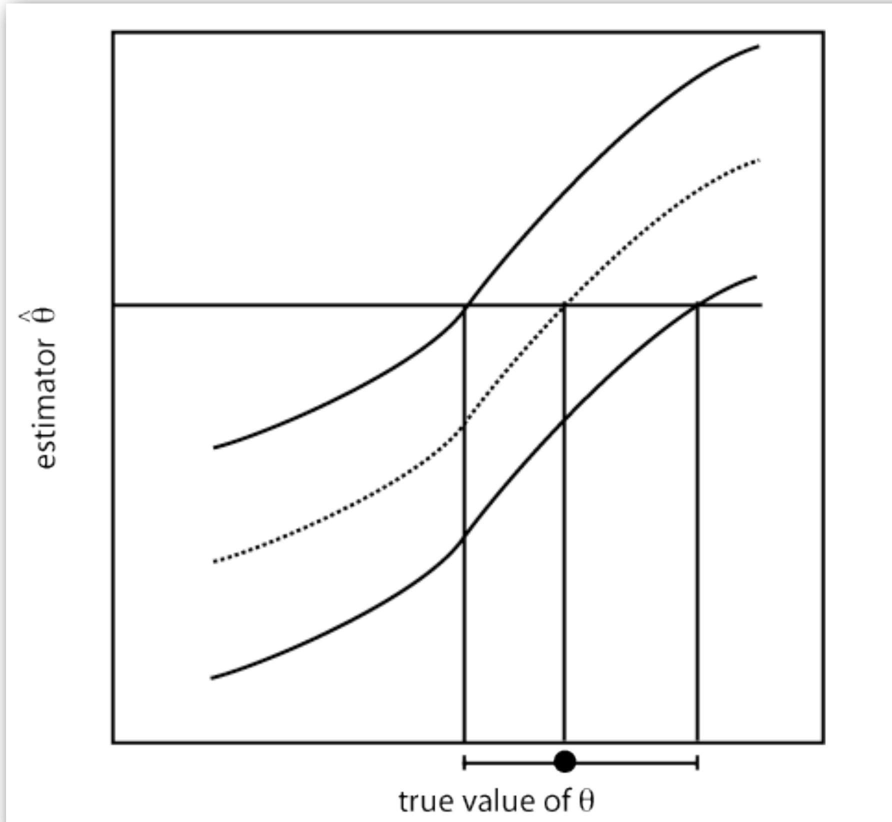
- we call the region

$$a(\hat{\theta}) \le \theta \le b(\hat{\theta})$$

  the confidence interval for the true value of θ given the estimator (and its pdf)

- the confidence level is the coverage probability, equal to 1-α-β

this is the frequentist approach to setting confidence intervals

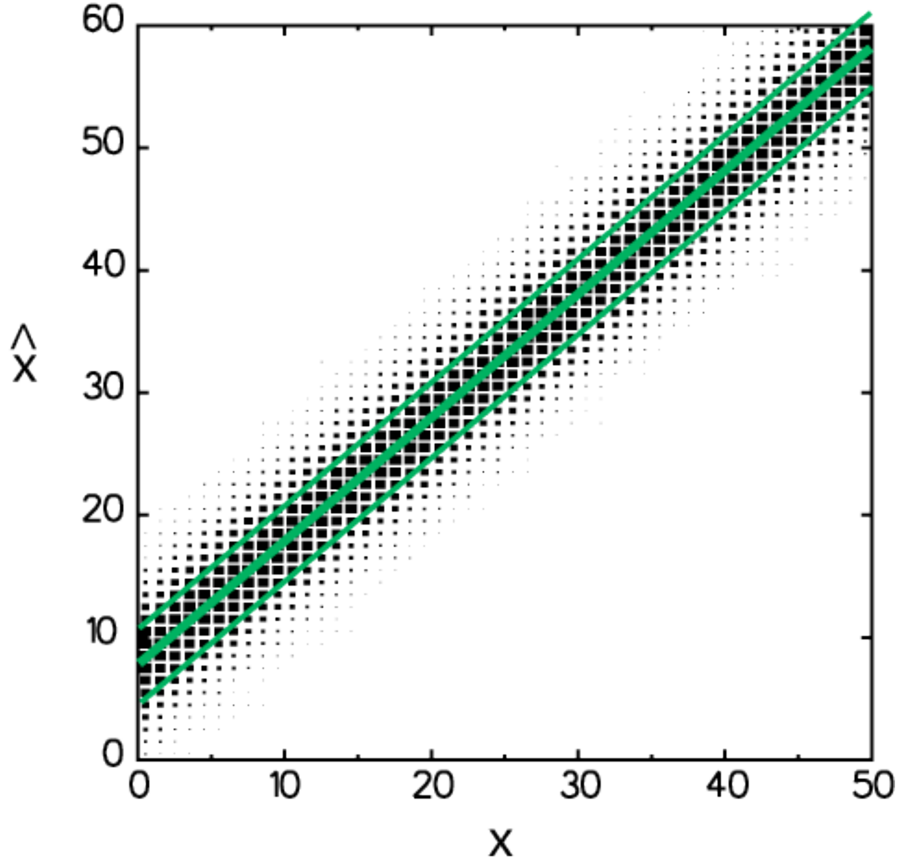


Neyman construction

- how do we perform the Neyman construction?
  - analytic calculation: clearly requires full knowledge of estimator pdf, integrability, and invertibility
  - Monte Carlo: scan in θ, estimate the u_α and v_β points for desired α and β numerically, interpolate to get inverse functions
- in the vast majority of cases we construct Neyman bands via a MC technique
- $\alpha=\beta$ corresponds to central interval
- (are these always sensible??)

example: Gaussian numbers

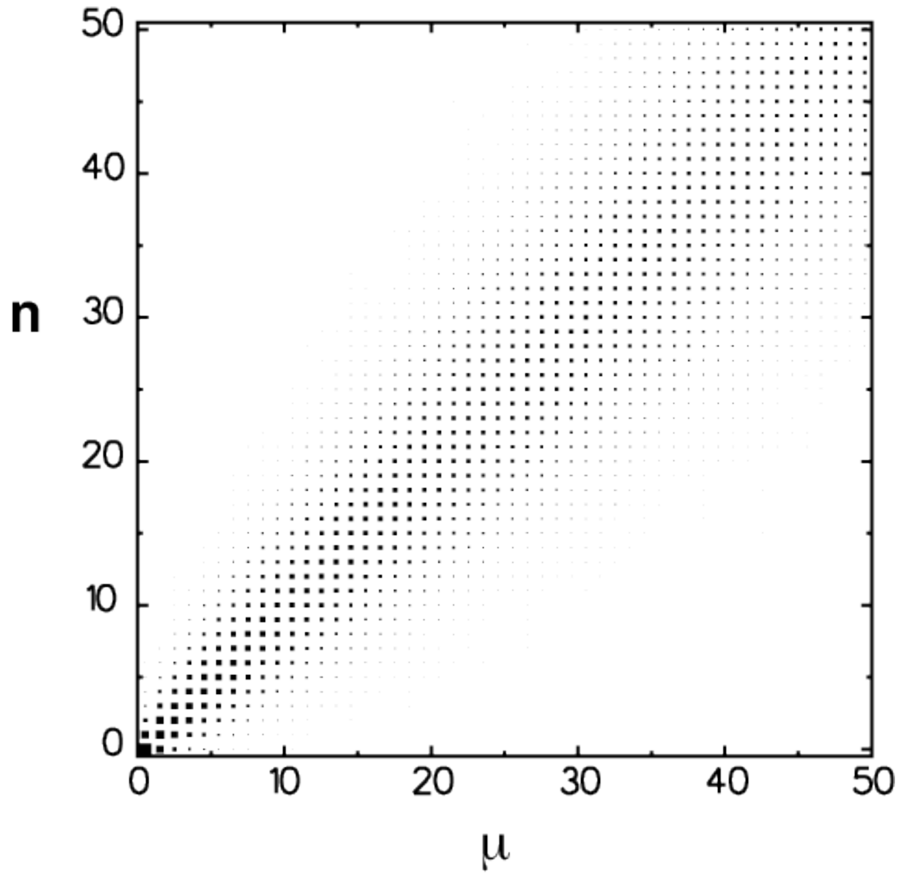
- plot estimator for x versus x



- this is a biased estimator! (which can be okay in this formulation...)

bad example: Poisson numbers

- generate Poisson-distributed random numbers for values of μ in the range (0,50)



- bad example because cannot draw confidence belt as smooth function! (u_α and v_β jump!)

confidence belts for discrete distributions

- must decide on desired coverage properties
    - always over-cover?
    - minimally over-cover?
    - average coverage?
- these are some of the toughest practical problems to deal with in statistical data analysis
- problem is "interesting" near zero...discuss
- physical boundaries in general represent problems for setting intervals

One sided intervals (limits)

- we may wish to make a statement such as

$$\mathcal{P}(\theta < \theta_{95}; \hat{\theta}) > 95\%$$

- in this case θ_{95} is a 95% CL upper limit on θ
- this represents a one-sided interval
- the statement itself is manifestly Bayesian!
- a frequentist would say "given any value of $\theta > \theta_{95}$, there is less than a 5% chance of having observed the value of θ we did, or less"

$$\mathcal{P}(\hat{\theta} < \hat{\theta}_{obs}; \theta) < 5\% \;\; \forall \; \theta > \theta_{95}$$

Frequentist one-sided intervals

- readily done using Neyman construction: simply consider only the high tail of the distribution of the estimator; require $\beta = 1 - \gamma$ for desired confidence γ
- limits suffer from same discreteness problems in the Poisson regime
- simple cases can be calculated directly
- example: upper limit on μ for Poisson process given n observed events

$$\sum_{n=0}^{n_{obs}} \frac{\mu_{95}^n e^{-\mu_{95}}}{n!} = 0.05$$

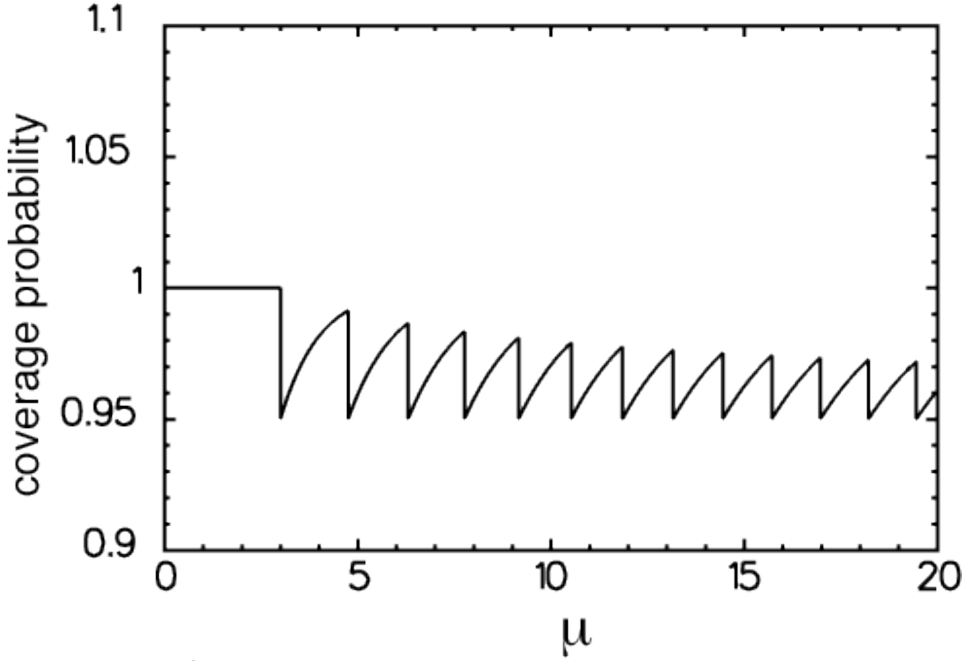
A few values for upper limits

- not bad to memorize the first few...

n_{obs}	90% CL	95% CL
0	2.30	3.00
1	3.89	4.74
2	5.32	6.30
3	6.68	7.75

coverage for upper limit on Poisson process

- “dinosaur plot” of coverage for upper limit on Poisson process:



- flat, unity near zero: why?
- why the jumps?
- overcoverage reflects > sign in calculation

Poisson process with background

- next most complicated limit problem: expect some background with rate b
- consider large n, b first - Gaussian approximation
  - 95% CL limit is at s > $1.96\sqrt{s}$
  - this explains why we optimize $s/\sqrt{b}$ for limits
- what about Poisson regime?
- great deal of history in our field with this problem
- "PDG formula": limit is at that value of the signal where we would have observed $\leq n_{obs}$ <u>and</u> have $n_{bkg} \leq n_{obs}$, all with probability 1-γ

PDG formula

$$\frac{\sum_{j=0}^{n_{obs}} (s+b)^j e^{-(s+b)}/j!}{\sum_{j=0}^{n_{obs}} b^j e^{-b}/j!} < 0.05$$

- this formula was in the 1988-1996 Review of Particle Properties Review of Statistics (not the PDG!)
- remarkably it was derived independently using a Bayesian approach by O. Helene in 1983:

  O. Helene, Nucl. Intrum. Methods Phys. Res. A 212, 319 (1983)

- now generally known as the Helene formula
- evades non-physical (negative) limits