Thoughts on asymmetry

Reminder: Ancient history

- Jerzy Neyman* created confidence intervals
- You can read his [original paper here]

X—Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability

By J. NEYMAN

Reader in Statistics, University College, London

(Communicated by H. JEFFREYS, F.R.S.—Received 20 November, 1936—Read 17 June, 1937)

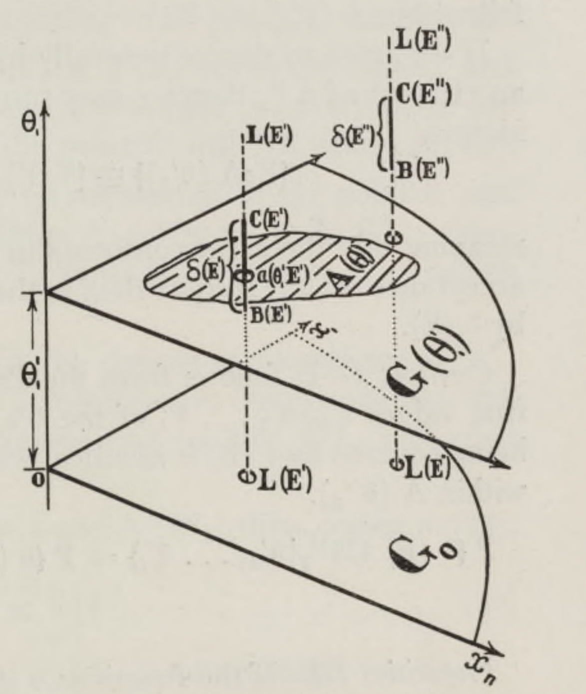


FIG. 1—The general space G.

- * I don’t think he was one of the UCL statisticians who was also an eugencist

Reminder: Confidence intervals

Is this the Bayesian Approach?

- This short summary of how to create confidence intervals is (hideously paraphrased by me) from the Feldman-Cousins paper
- If we have a 1D system with some parameter μ then to make a α confidence interval:
  - For each μ we draw the acceptance interval (horizontal line) $[x_1, x_2]$ that contains α of the probability, i.e. $P\left(x \in [x_1, x_2] \,|\, \mu\right) = \alpha$
- For a given measurement x_o the confidence interval $[\mu_1, \mu_2]$ is the union of the all the values of μ for which the vertical line at x_o intercepts the acceptance intervals.

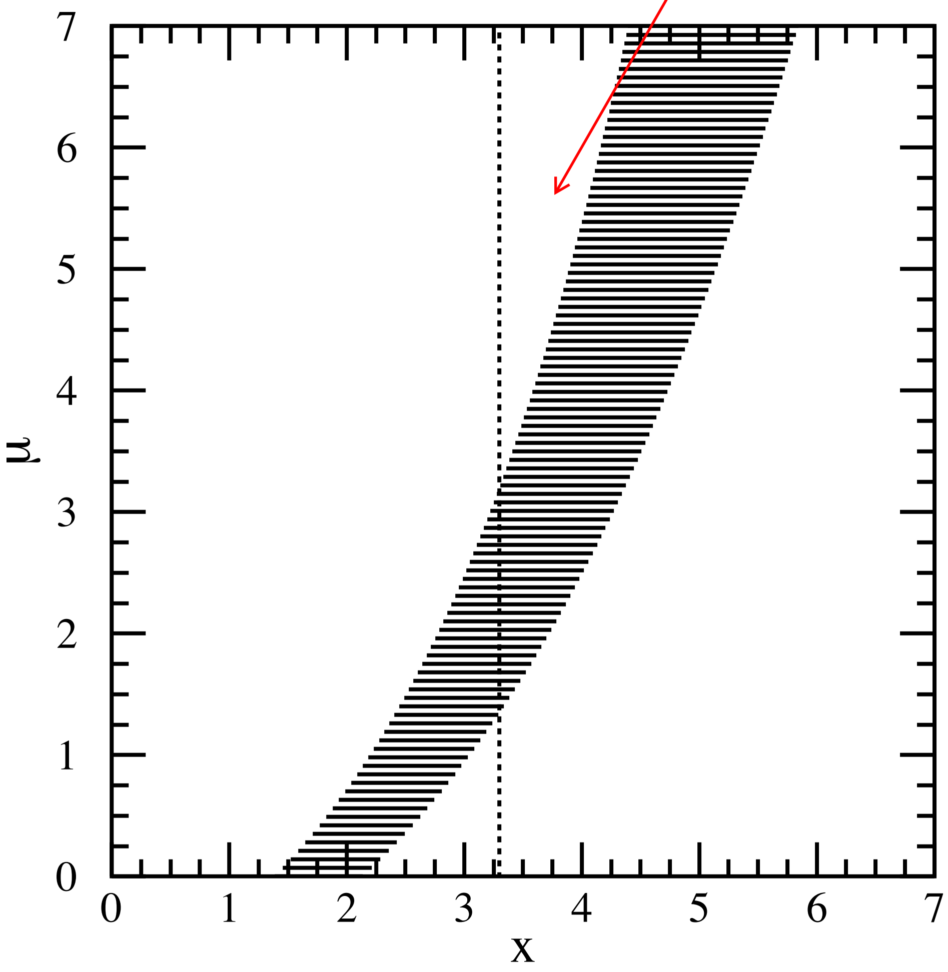
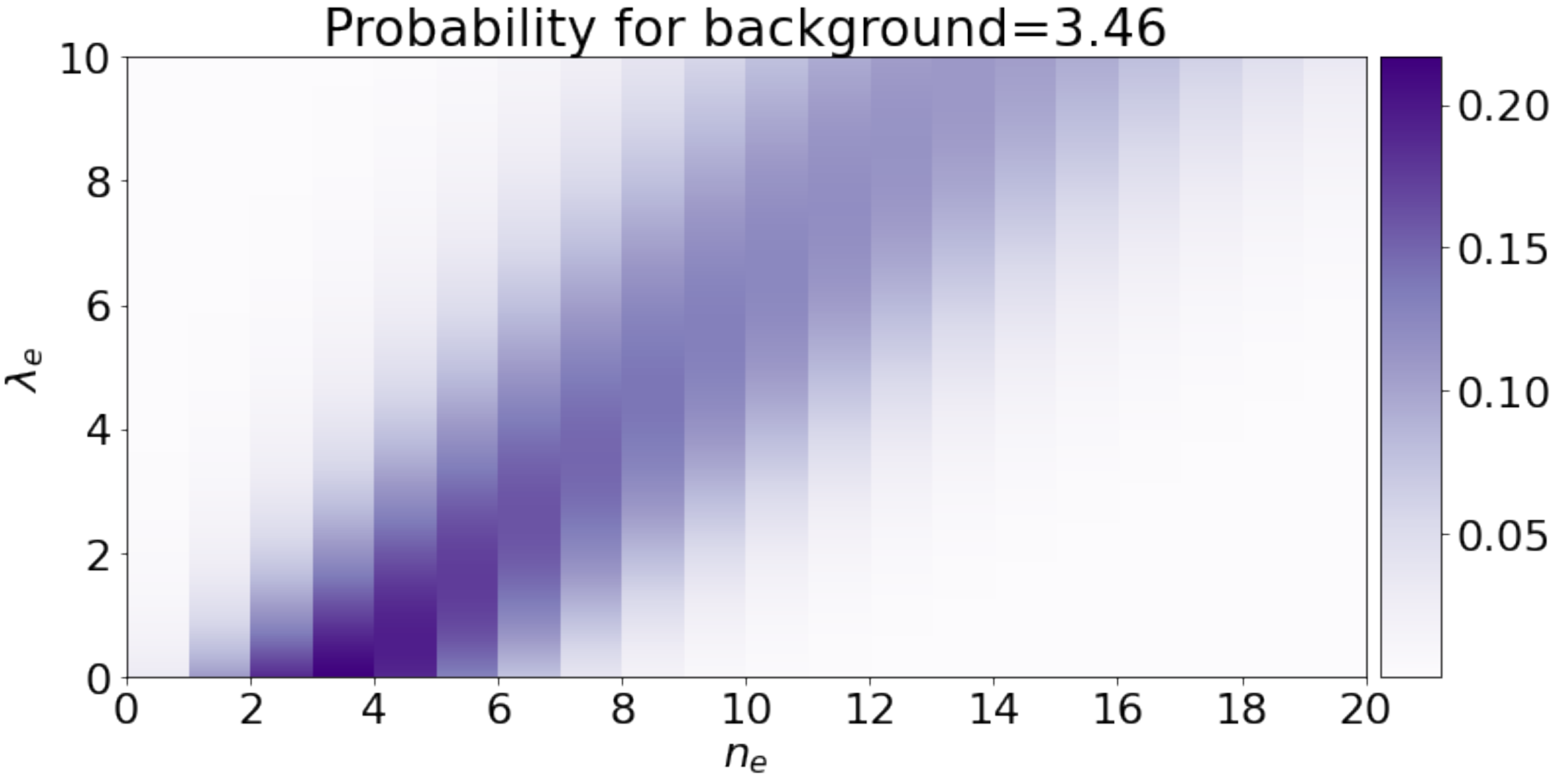


Figure 1 from [“A Unified Approach to the Classical Statistical Analysis of Small Signals”, Feldman & Cousins, 1998]

Reminder: Poisson Confidence Intervals

- Now imagine we have some number of unknown signal λ_e with a known expected background, $b = 3.46$
- Using the Poisson distribution for each λ_e (real not integer) we can calculate the probability of observing n_e events $P(n_e\,|\,\lambda_e, b)$
- So each row of the right-hand image is a discrete probability distribution



So the expression is:
Is the background added to lambda?

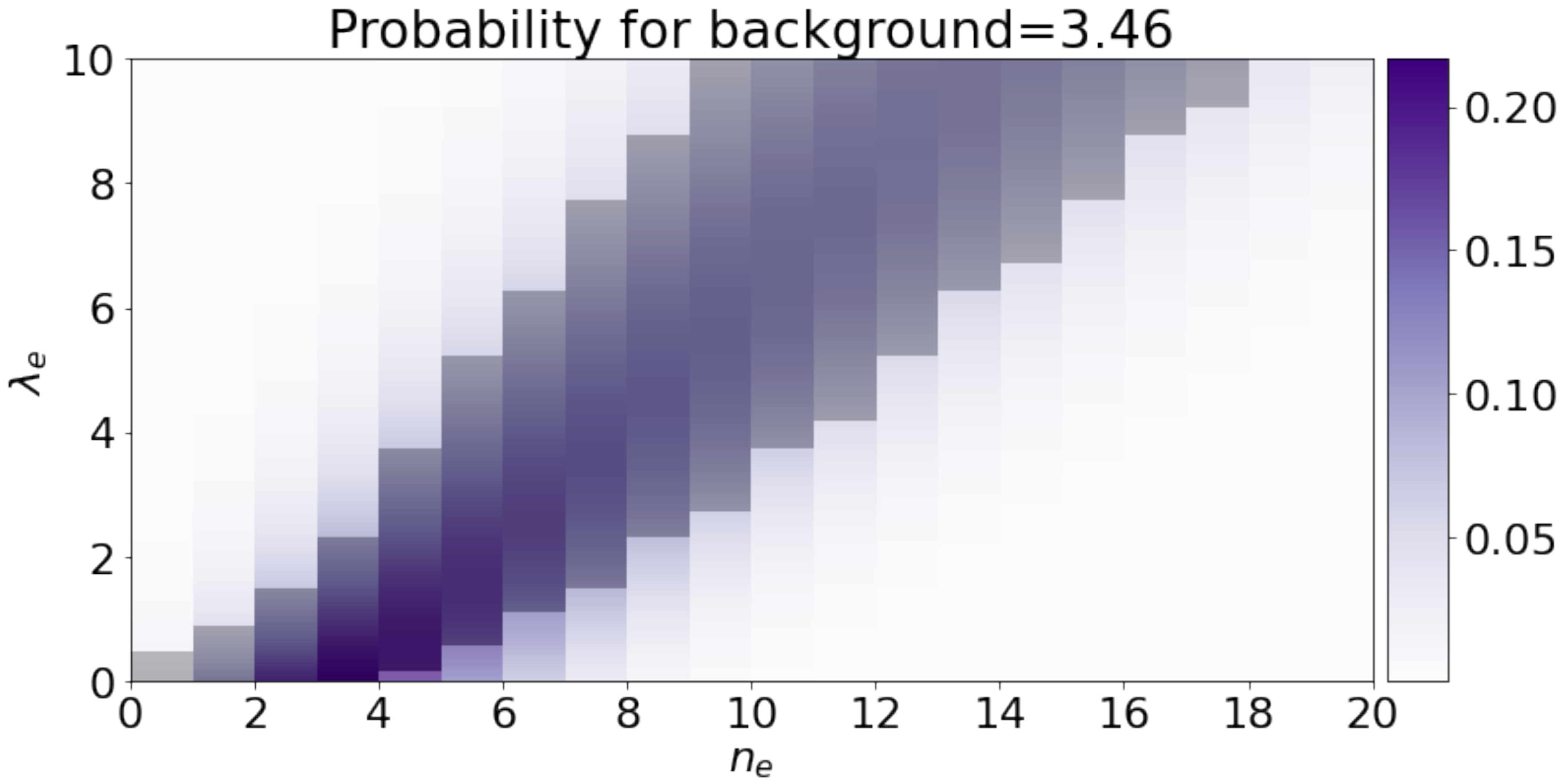
$$P(n_e|\lambda_e, b) = \frac{e^{-(\lambda_e + b)} \cdot (\lambda_e + b)^{n_e}}{n_e!}$$

Or

$$P(n_e|\lambda_e, b) = \frac{e^{-\lambda_e} \cdot \lambda_e^{\,n_e}}{n_e!} + b$$

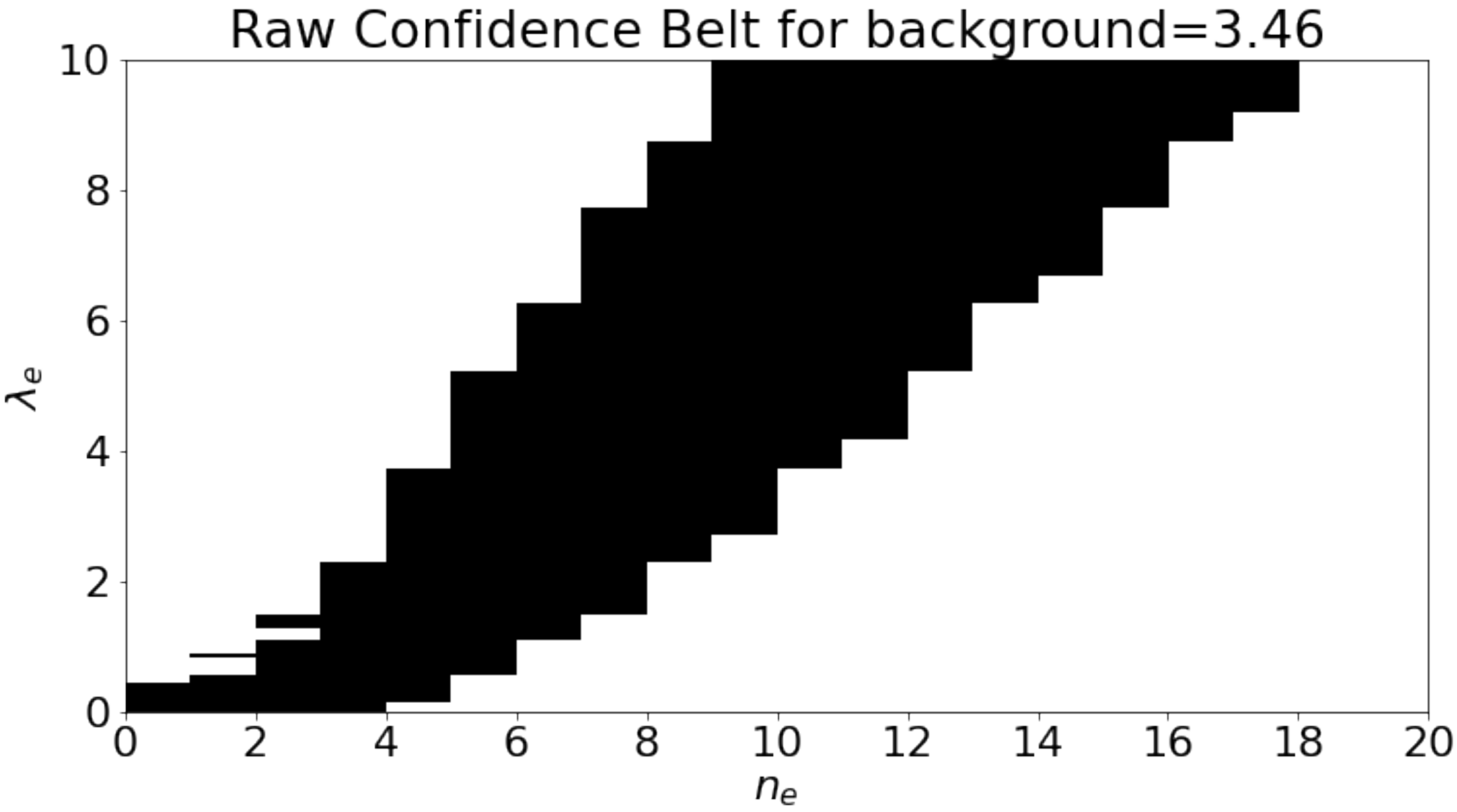
Reminder: Poisson Confidence Intervals

- Now we can use the Feldman-Cousins unified approach to draw horizontal lines on the plot that cover 68% of $P(n_e | \lambda_e, b)$



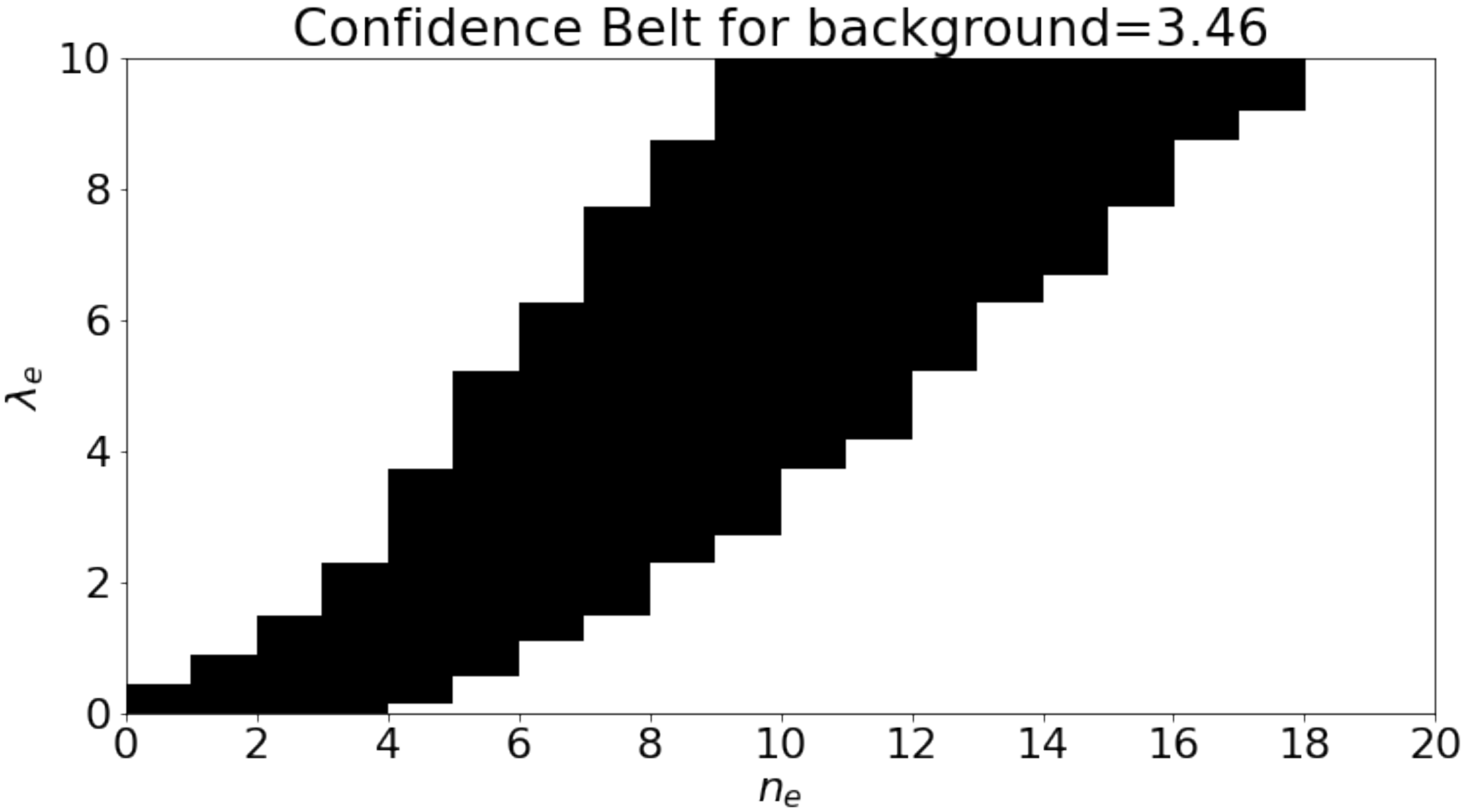
Reminder: Poisson Confidence Intervals

- These lines form our confidence belt such that when we measure n_e events we can draw a vertical line at determine our confidence interval on λ_e
- As noted in the F-C paper the discrete nature of n_e leads to some pathologies like this. Where we have to fill in the white gaps in the black regions.



Reminder: Poisson Confidence Intervals

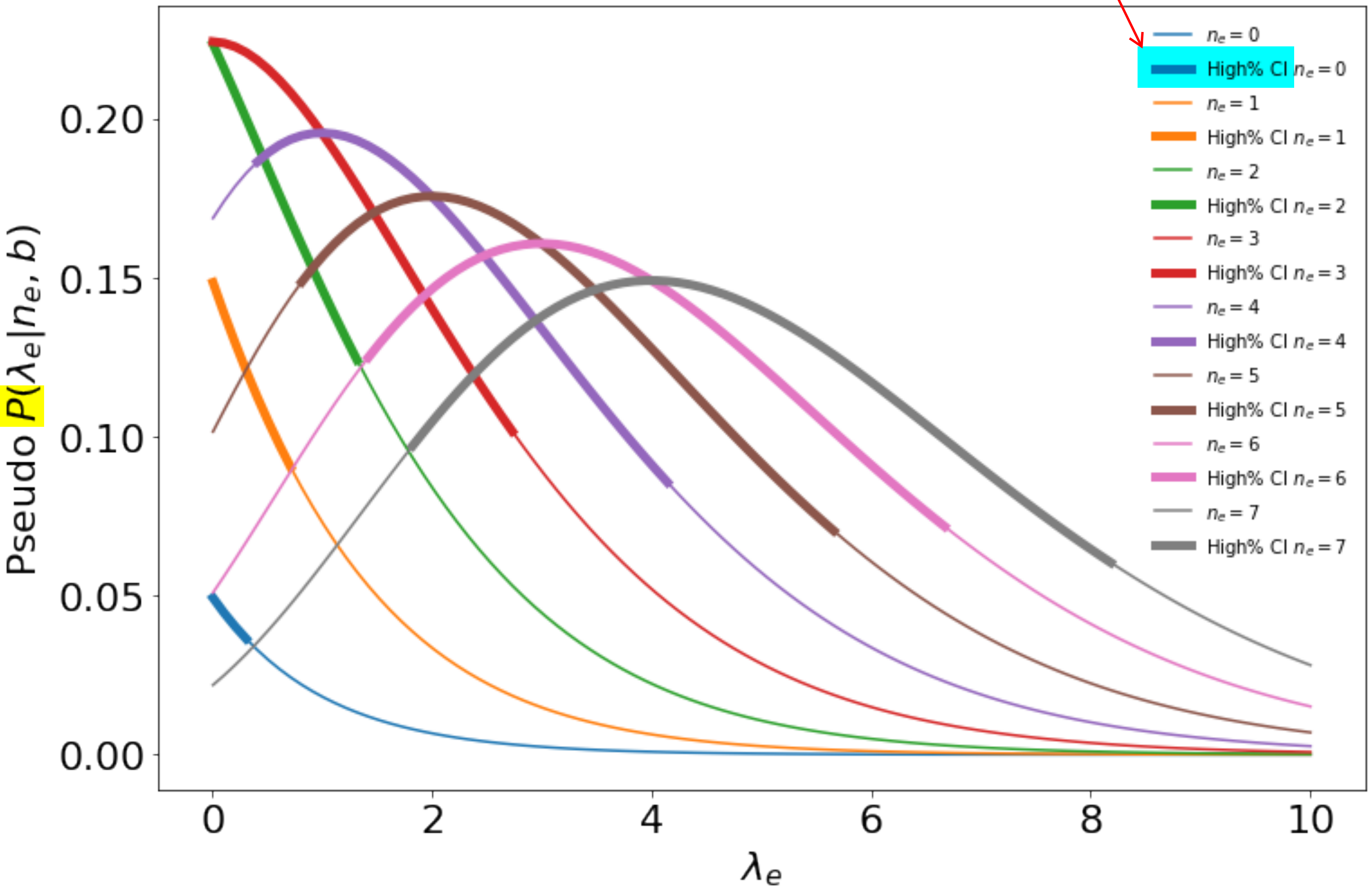
- These lines form our confidence belt such that when we measure n_e events we can draw a vertical line at determine our confidence interval on λ_e



Vertical slices through the λ_e vs n_e probability distribution?

- What happens if we take the vertical slice along some measured number of events and plot $P(n_e | \lambda_e, b)$, obviously it is tempting to interpret this as $P(\lambda_e | n_e, b)$
- But if we want to be strictly frequentist we are not allowed to do this.

What is this?
68% CL



Refer to P5

Aside: Bayesian Relative Belief Updating Ratio

- The Poisson distribution for getting n_o when we expect a signal of λ_e and background of λ_b is $f(n_o \mid \lambda_e, \lambda_b) = \dfrac{e^{-(\lambda_e + \lambda_b)} \left(\lambda_e + \lambda_b\right)^{n_o}}{n_o!}$

What is the relation
between lambda_b and b?

- Starting from a prior we can generate a relative belief updating ratio defined as

$$\boxed{\mathscr{R}(\lambda_e; n_o, \lambda_b) = \frac{f\left(n_o \mid \lambda_e, \lambda_b\right)}{f\left(n_o \mid \lambda_e = 0, \lambda_b\right)}}$$

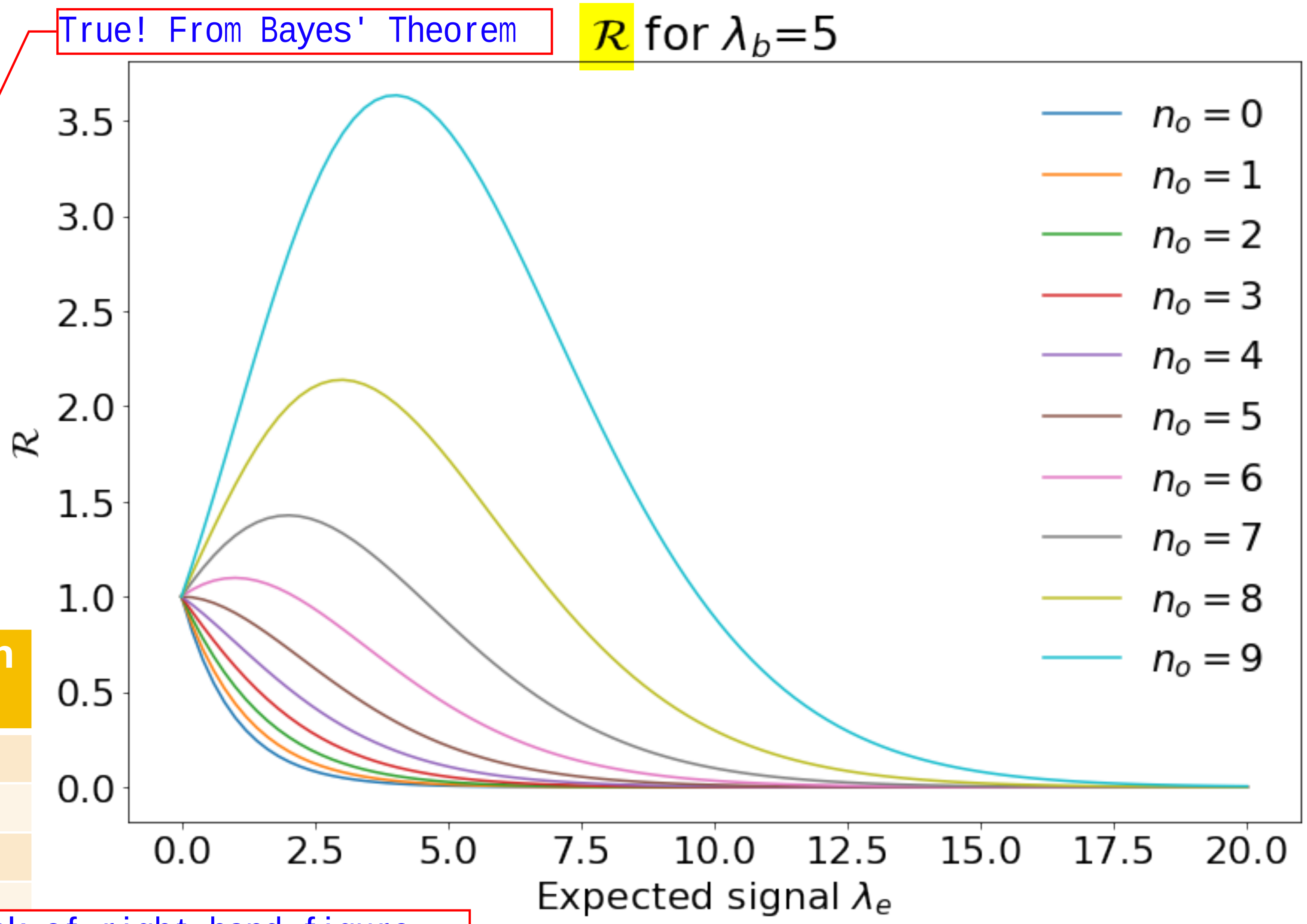
- Which for us is

$$\mathscr{R}(\lambda_e; n_o, \lambda_b) = e^{-\lambda_e} \left(1 + \frac{\lambda_e}{\lambda_b}\right)^{n_o}$$

Aside: Relative Belief Updating Ratio

- The relative belief updating ratio has a silly name, but it is very simple to use.
- If we take a flat prior in λ_e then $\mathscr{R}$ is proportional to the posterior probability
- So the peak value of $\mathscr{R}$ is the most preferred value
- The $\mathscr{R}$ curves are a normalisation away from the curves which were not $P(\lambda_e \mid n_e, b)$

True! From Bayes' Theorem

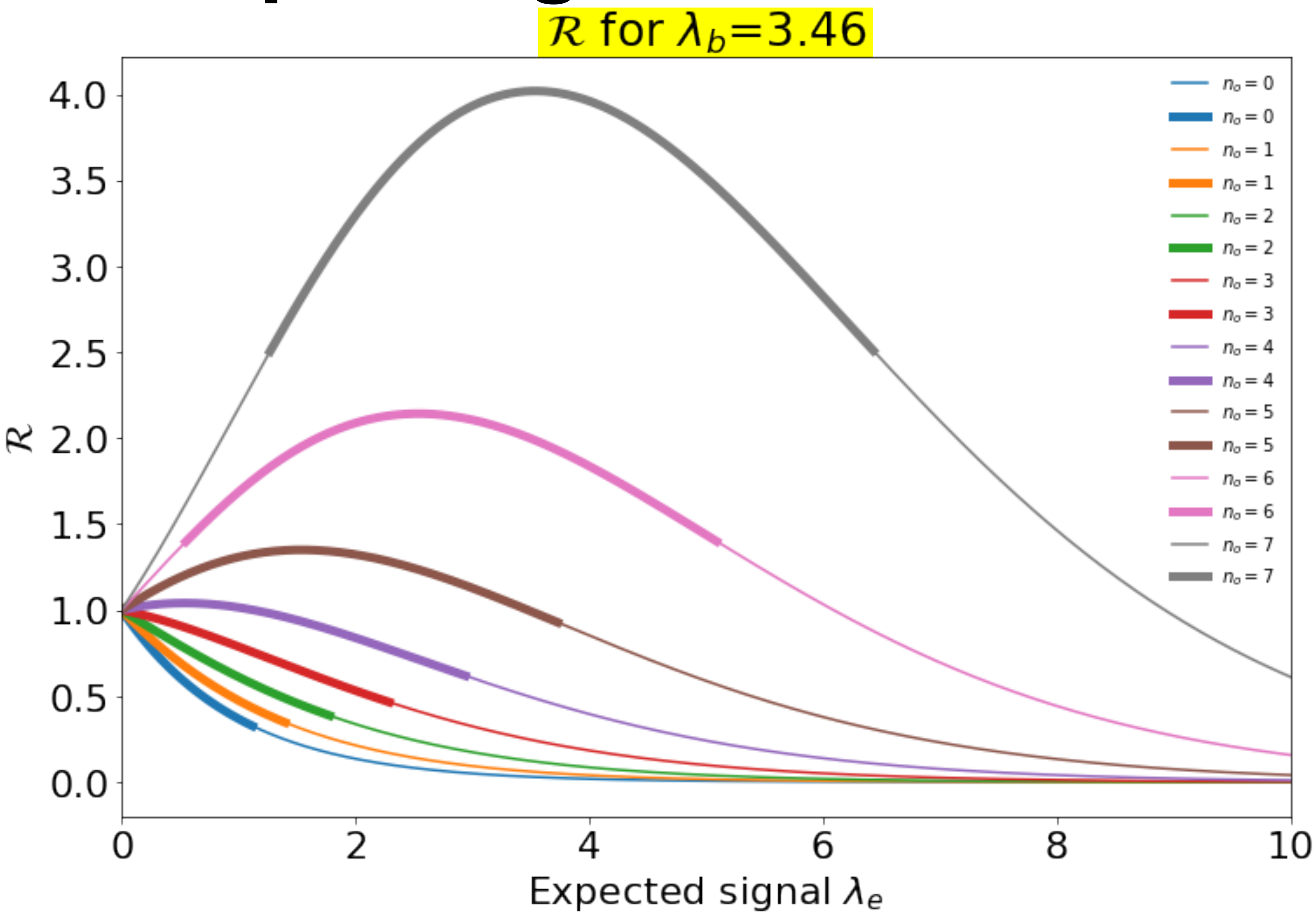


Observed	Expected background	Signal Prediction
0	5	0
5	5	0
6	5	1.01
7	5	2.02

Peak of right-hand figure.

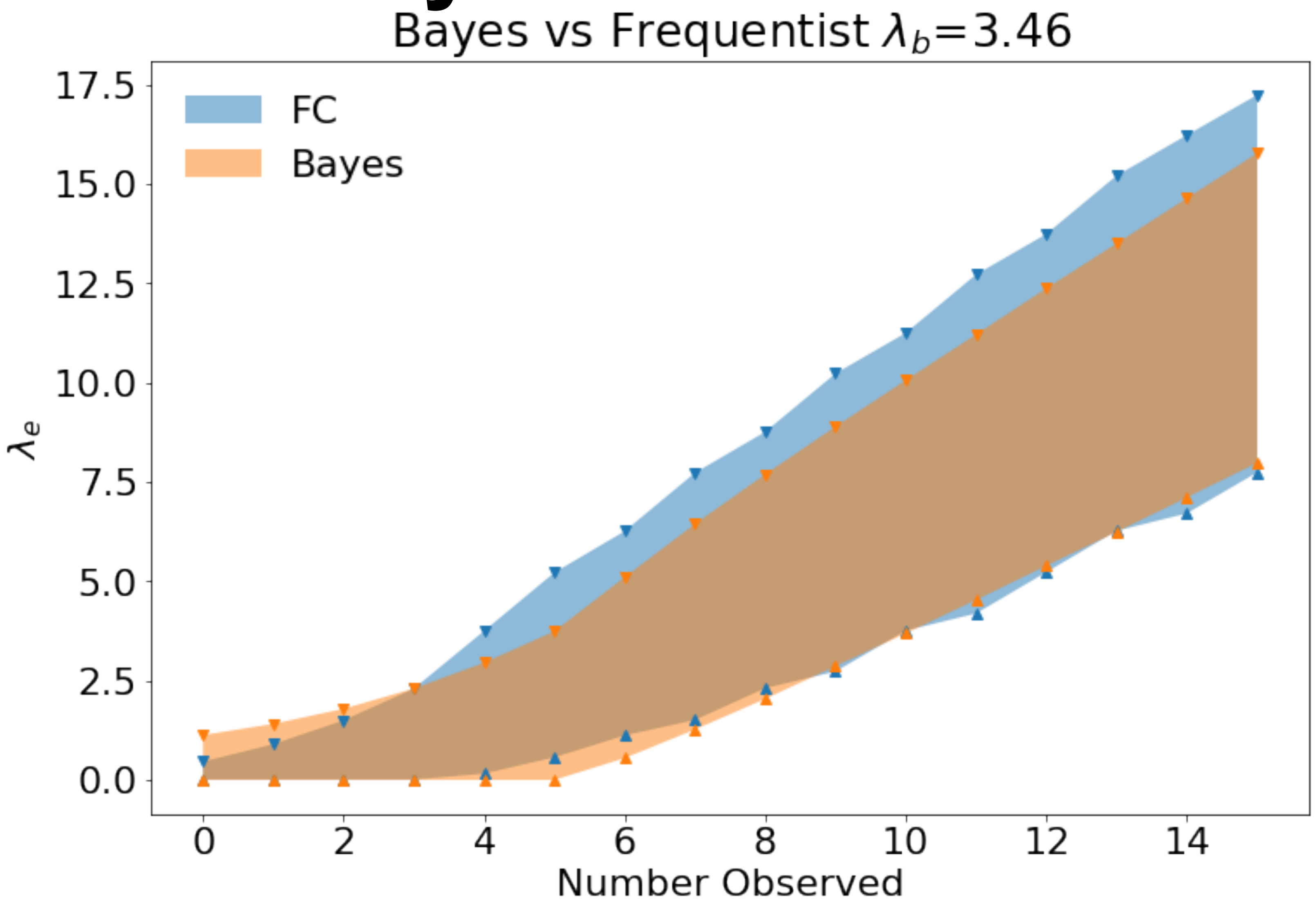
Aside: Relative Belief Updating Ratio

- The relative belief updating ratio has a silly name, but it is very simple to use.
- If we take a flat prior in λ_e then $\mathscr{R}$ is proportional to the posterior probability
- So the peak value of $\mathscr{R}$ is the most preferred value
- The $\mathscr{R}$ curves are a normalisation away from the curves which were not $P(\lambda_e \mid n_e, b)$
- If we just pretend these curves are $P(\lambda_e \mid n_e, b)$ and then select the most probable 68% we can also trivially define something that looks like confidence intervals (are they credibility intervals?)



Aside: Frequentist vs Bayesian

- If you compare the 68% frequentist (FC) vs Bayesian limits for a background of 3.46 they have similarities and differences
    - At highish number of observed events the FC limits always have more coverage (probably due to the integer maths)
    - At low number of observed events the Bayesian limit has more coverage (there is lots of discussion of this feature and its dependence on the background in the literature)



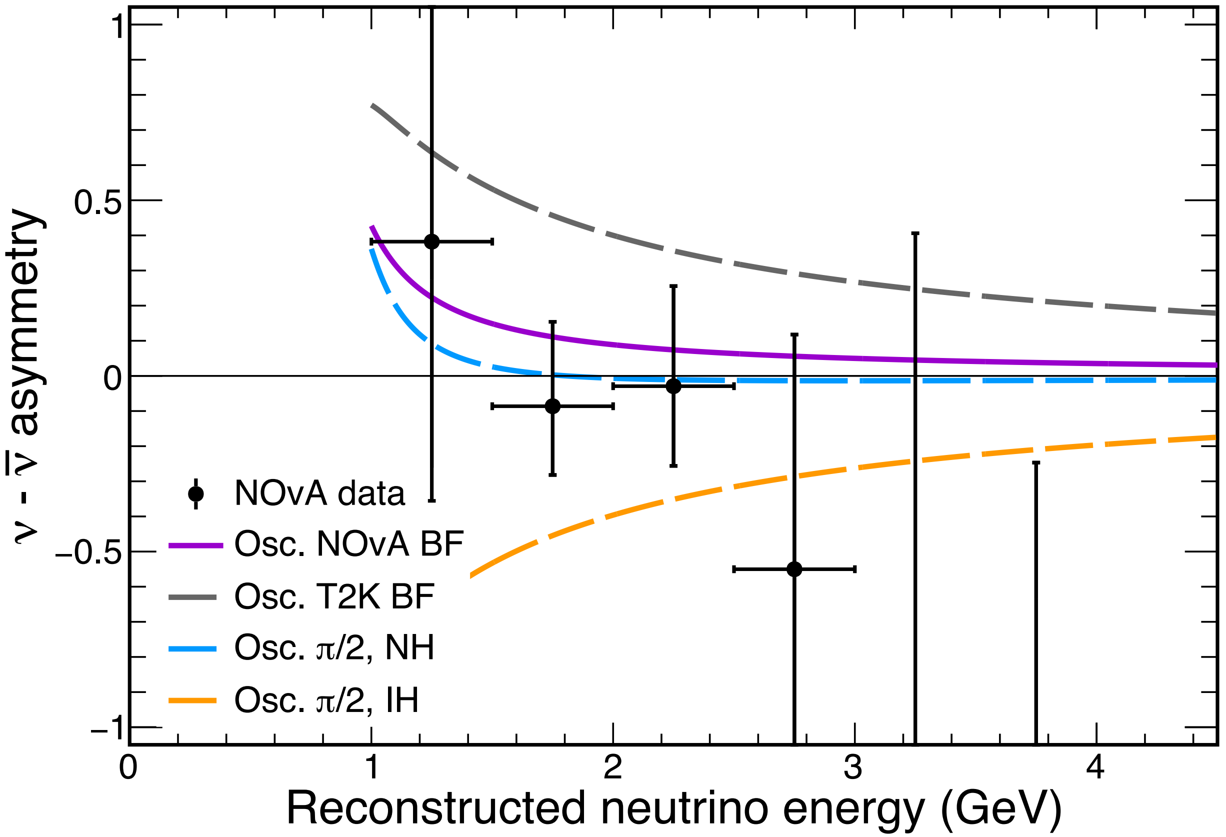
Thoughts on asymmetry

Asymmetry Error Bar

- Jon Urheim pointed out that our current error bars on the asymmetry measurement enter unphysical regions

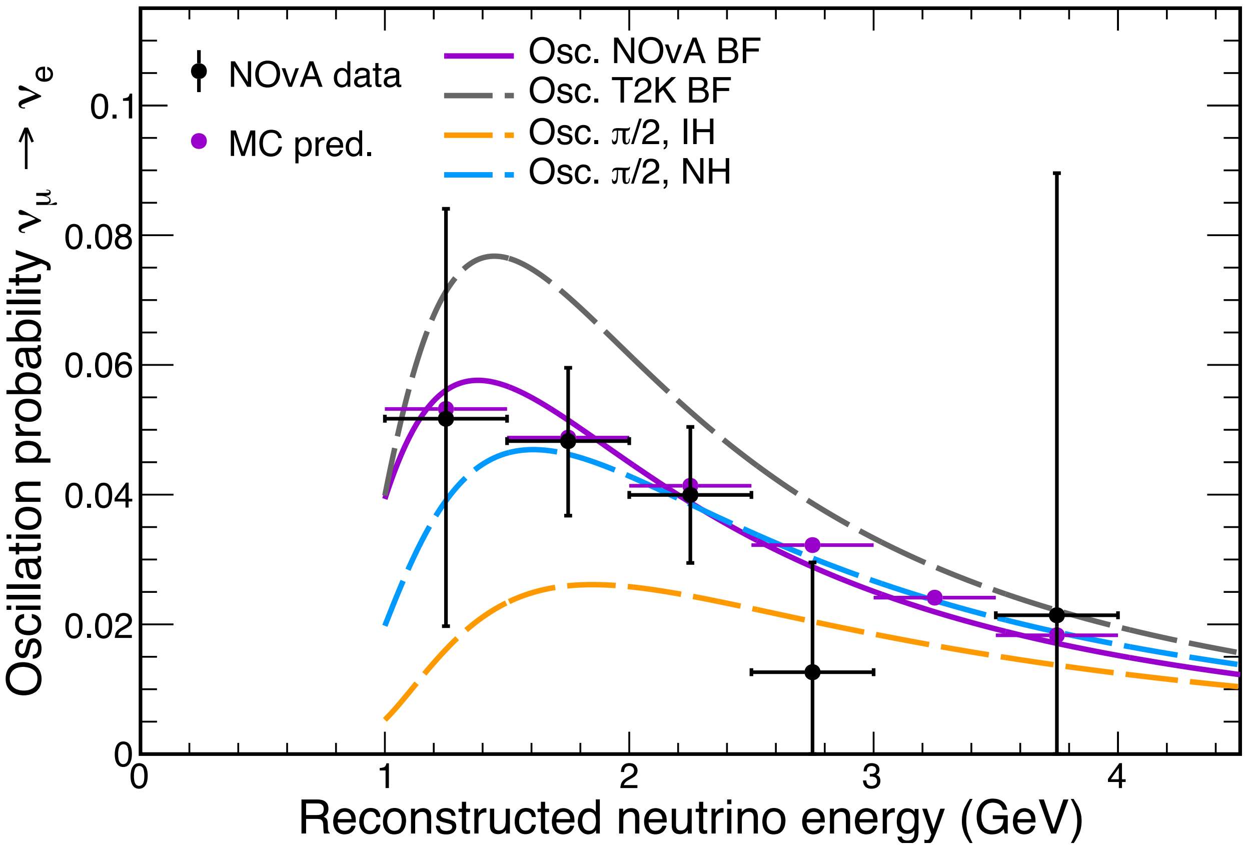
$$\mathscr{A}_{CP} = \frac{P(\nu_\mu \to \nu_e) - P(\bar{\nu}_\mu \to \bar{\nu}_e)}{P(\nu_\mu \to \nu_e) + P(\bar{\nu}_\mu \to \bar{\nu}_e)}$$

- Since $\mathscr{A}_{CP}$ is bounded to the region -1 to +1
- Why do our error bars enter these unphysical regions?



Probability Error Bar

- If we look at the error bars on the probability distribution we see the same effect of error bars extending into the unphysical region of P<0
- We define the oscillation probability as (data-background) / unosc, which obviously has the possibility of being negative



Error bar calculation

- Currently to calculate the error bar we do the following:
  - Generate pseudo-experiment data based on Poisson fluctuations around the observed number of events in the real data
  - Fit the pseudo-experiment data using (random) systematically shifted data.. this determines the new background at the new best fit point.
  - Define the probability as (data-background)/unoscillated *
  - Make histograms of the event counts in each bin, oscillation and asymmetry in each bin.
    - Define the error bars from the 68% region in these histograms

*: This is the point at which we things can go into the unphysical region.

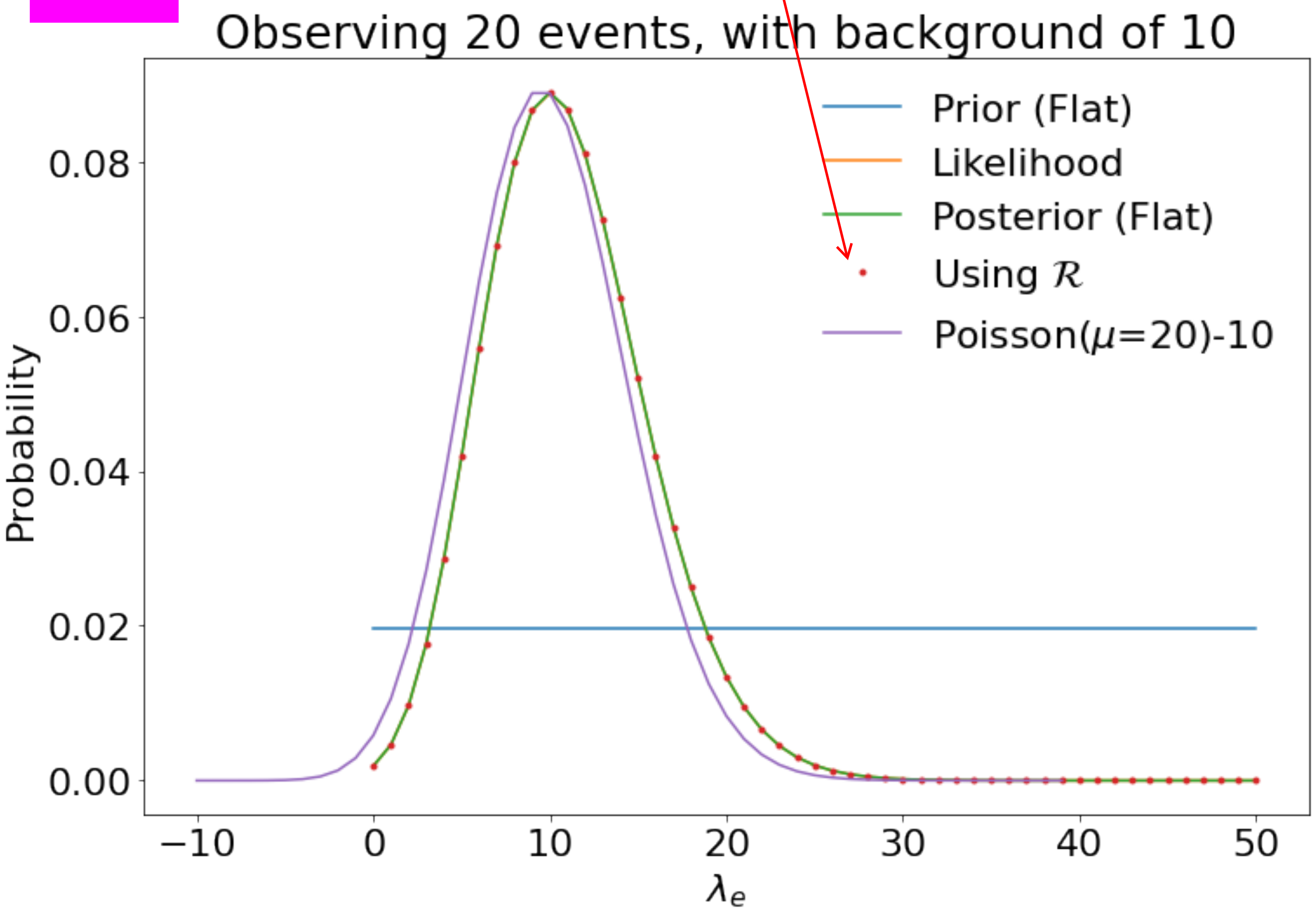
Potential Problems and Potential Solutions

- There are a couple of potential problems with the current system
  1. What do we do for the bins with zero events in the data?
  2. How do we deal with background subtraction forcing us into the unphysical region?
- I think the solutions to both of these problems is either to be a little more Frequentist or be a little more Bayesian
  1. Don't have bins with too few events!
  2. Rather than taking the observed number of events and subtracting the background we should determine the most likely λ_e in each pseudo experiment given the expected background and observed number of events (this is actually just taking the maximum of 0 and λ_e - background.

What are we doing

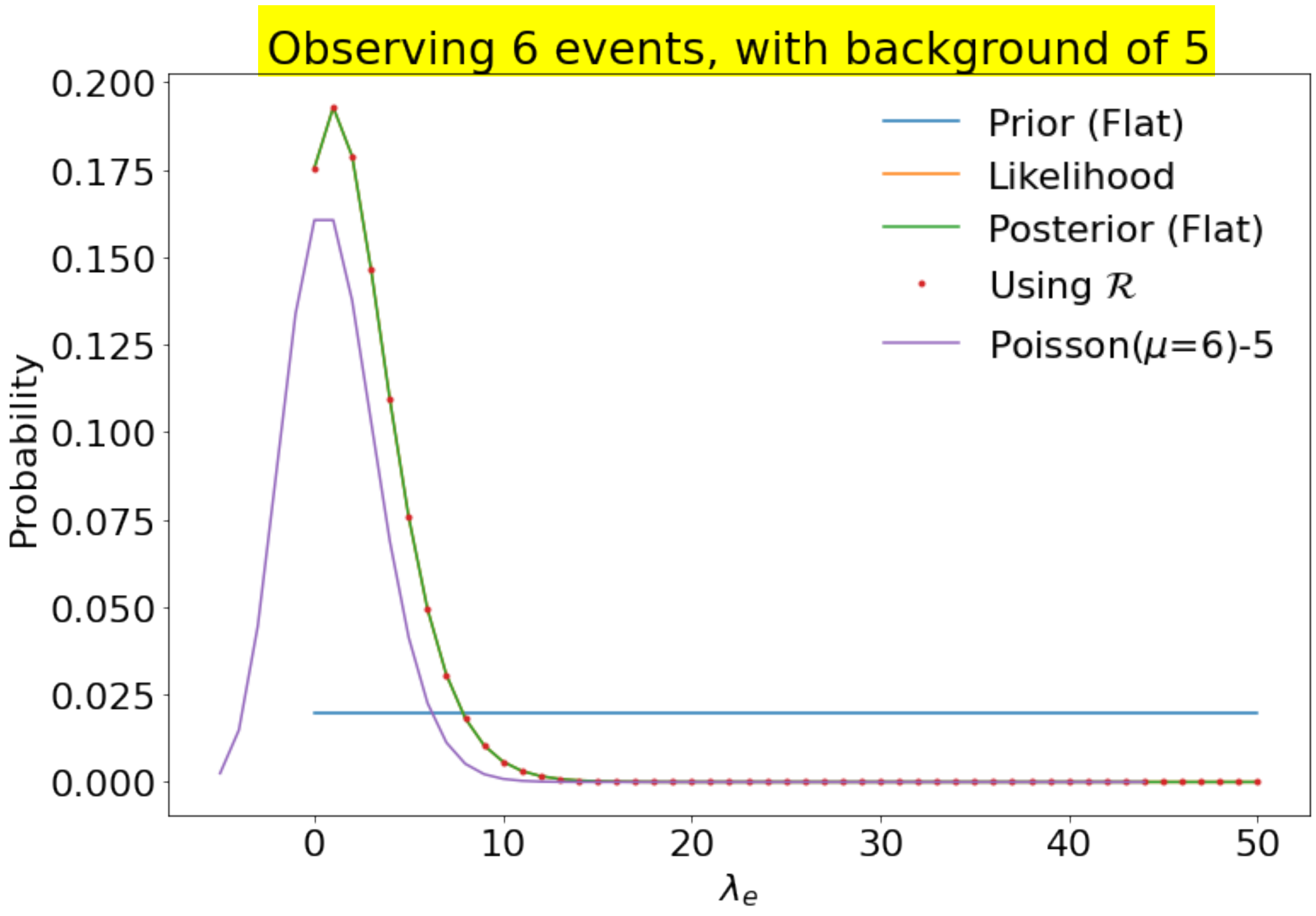
- Here is a comparison of what we are currently doing (throwing a Poisson at observed number of events and then subtracting the background) to what I think we should be doing (using $\mathscr{R}$ to determine the probability distribution for λ_e
- In the high count limit they are very similar.

Compared to Fig in P10, is R here normalized?????



What are we doing

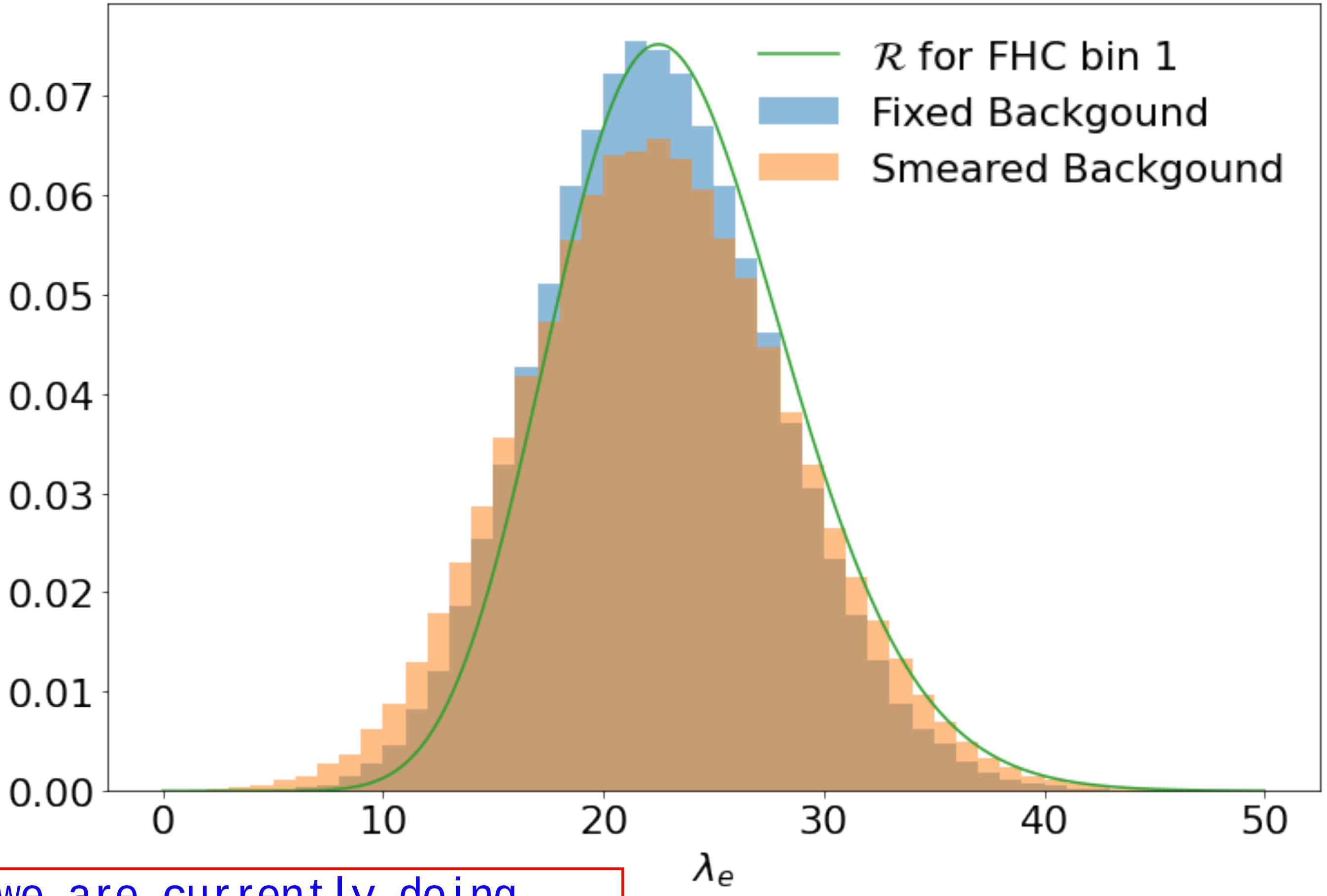
- Here is a comparison of what we are currently doing (throwing a Poisson at observed number of events and then subtracting the background) to what I think we should be doing (using $\mathscr{R}$ to determine the probability distribution for λ_e
- In the high count limit they are very similar.
- But they diverge at low counts



From P17

Practical recommendations

- The practical recommendations are
  1. In the data (and pseudo-experiments) determine the preferred value of λ_e in each bin using $\mathscr{R}$, this value can definitionally never be negative. The easiest way to do this is to histogram, max $\left(0,(n_o - \lambda_b)\right)$ as we are and take the peak non-negative bin as the preferred value. The confidence limits should be computed again using only the physical values.
  2. Don't have bins with too few events



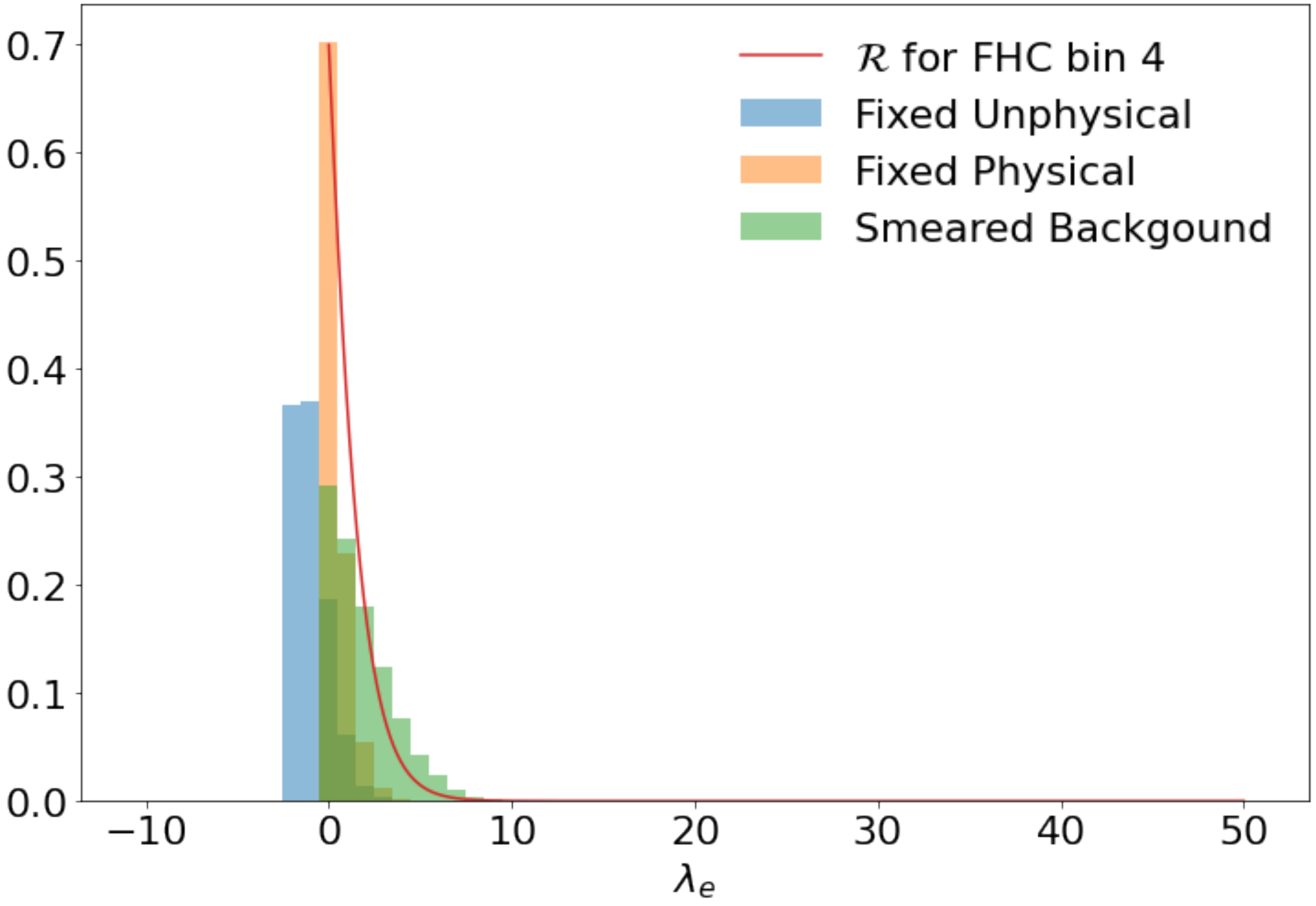
What we are currently doing

Green is $\mathscr{R}$ for $n_0 = 28, \lambda_b = 5.5$

Blue is histogram of throwing Poisson numbers with a mean of 28 and subtracting 5.5, Orange is the same random numbers but smearing the background by a normal distribution with width of 3.

Practical recommendations

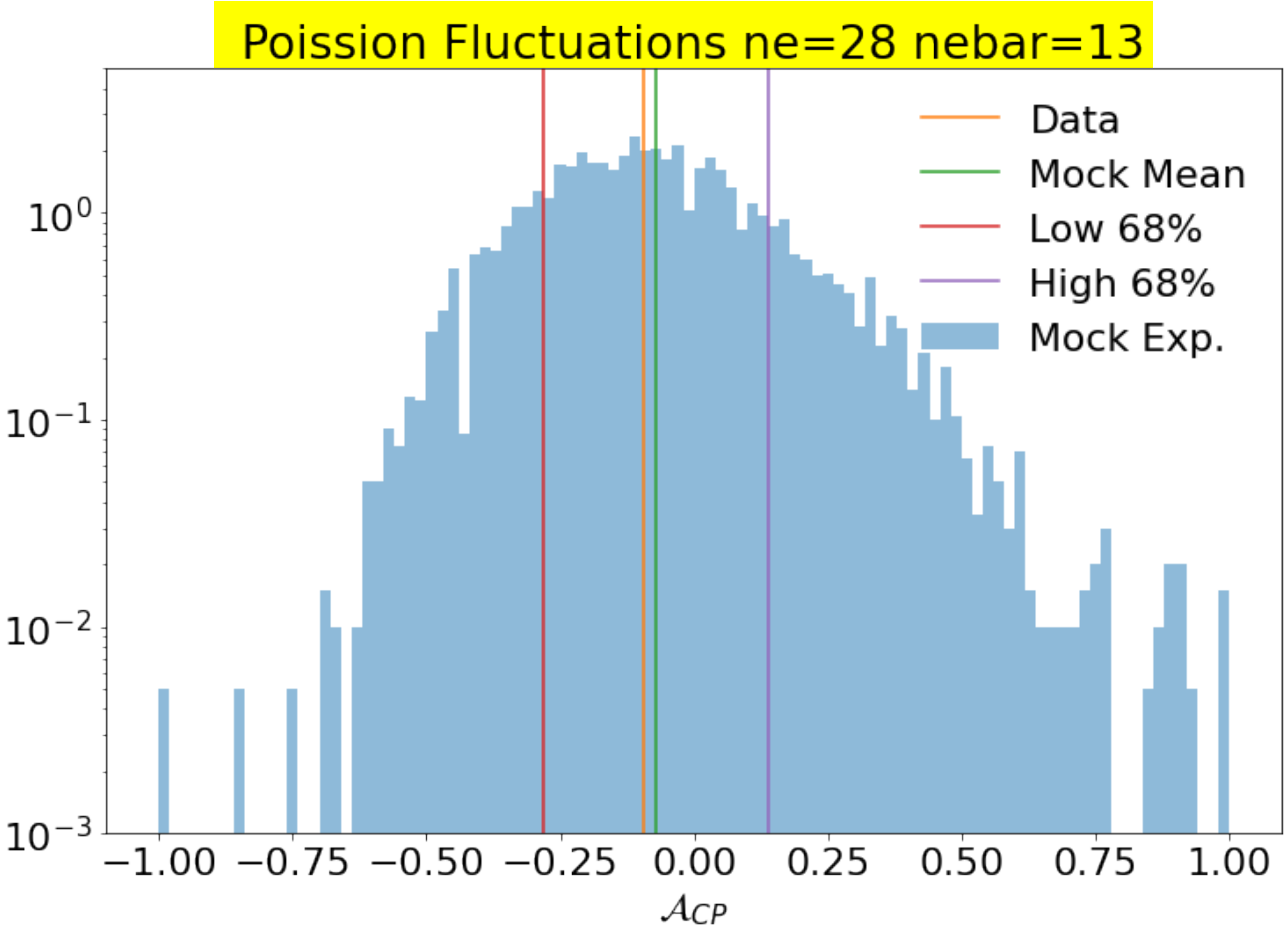
- The practical recommendations are
  1. In the data (and pseudo-experiments) determine the preferred value of λ_e in each bin using $\mathscr{R}$, this value can definitionally never be negative. The easiest way to do this is to histogram, max $\left(0,(n_o - \lambda_b)\right)$ as we are and take the peak non-negative bin as the preferred value. The confidence limits should be computed again using only the physical values.
  2. Don’t have bins with too few events



Red is $\mathscr{R}$ for $n_0 = 1, \lambda_b = 2.3$
Blue is histogram of throwing Poisson numbers with a mean of 1 and subtracting 2.3, Orange is the same but renormalised in the physical region, green is smearing the background by a normal distribution with width of 3.

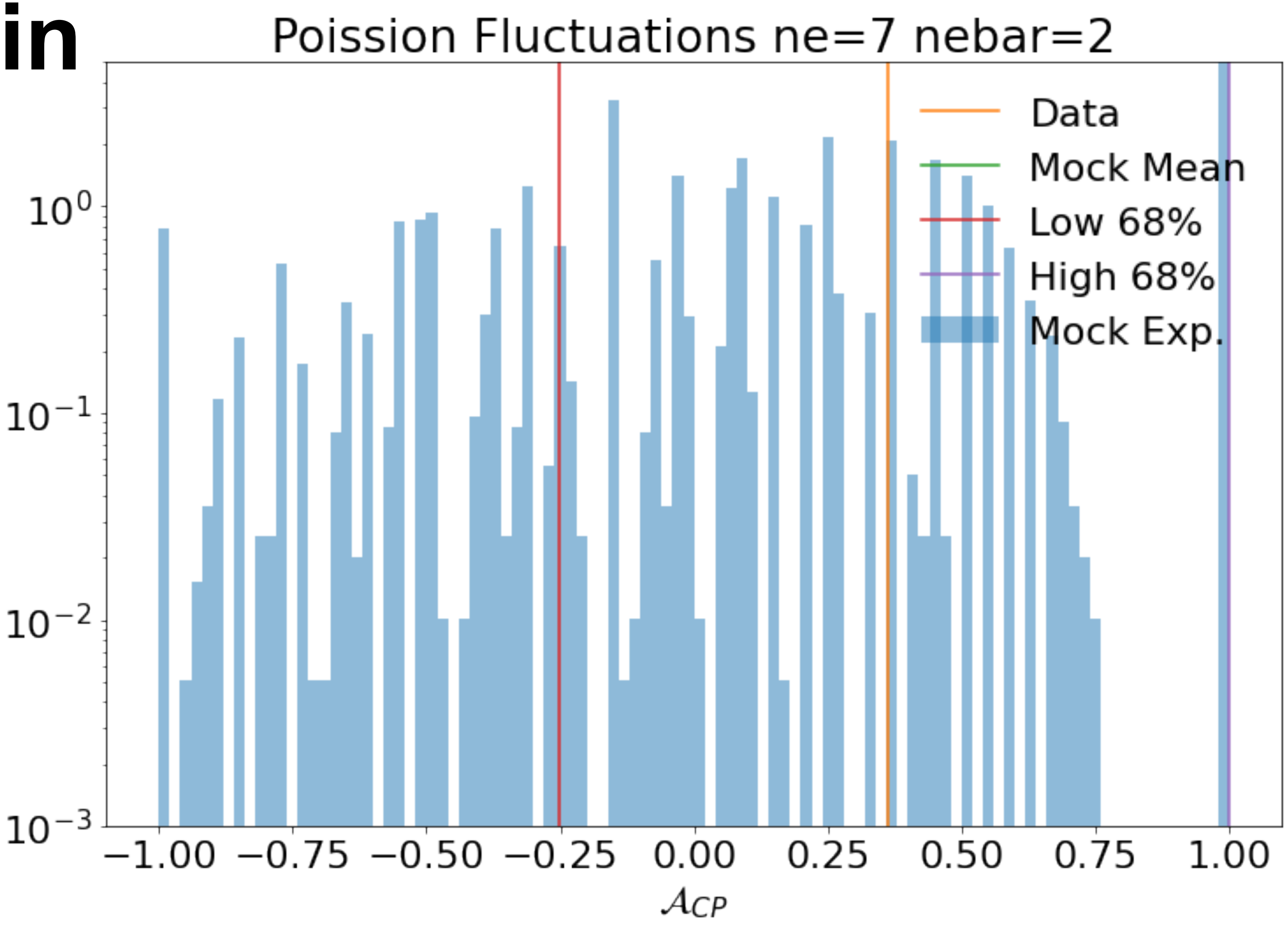
A Good Bin

- Here is what you get using our system for a particular bin
- Backgrounds: 5.5, 2.6
- Predictions: 471, 182



A Less Good Bin

- Here is what you get using our system for a particular bin
- Backgrounds: 2.4, 1.2
- Predictions: 89, 32



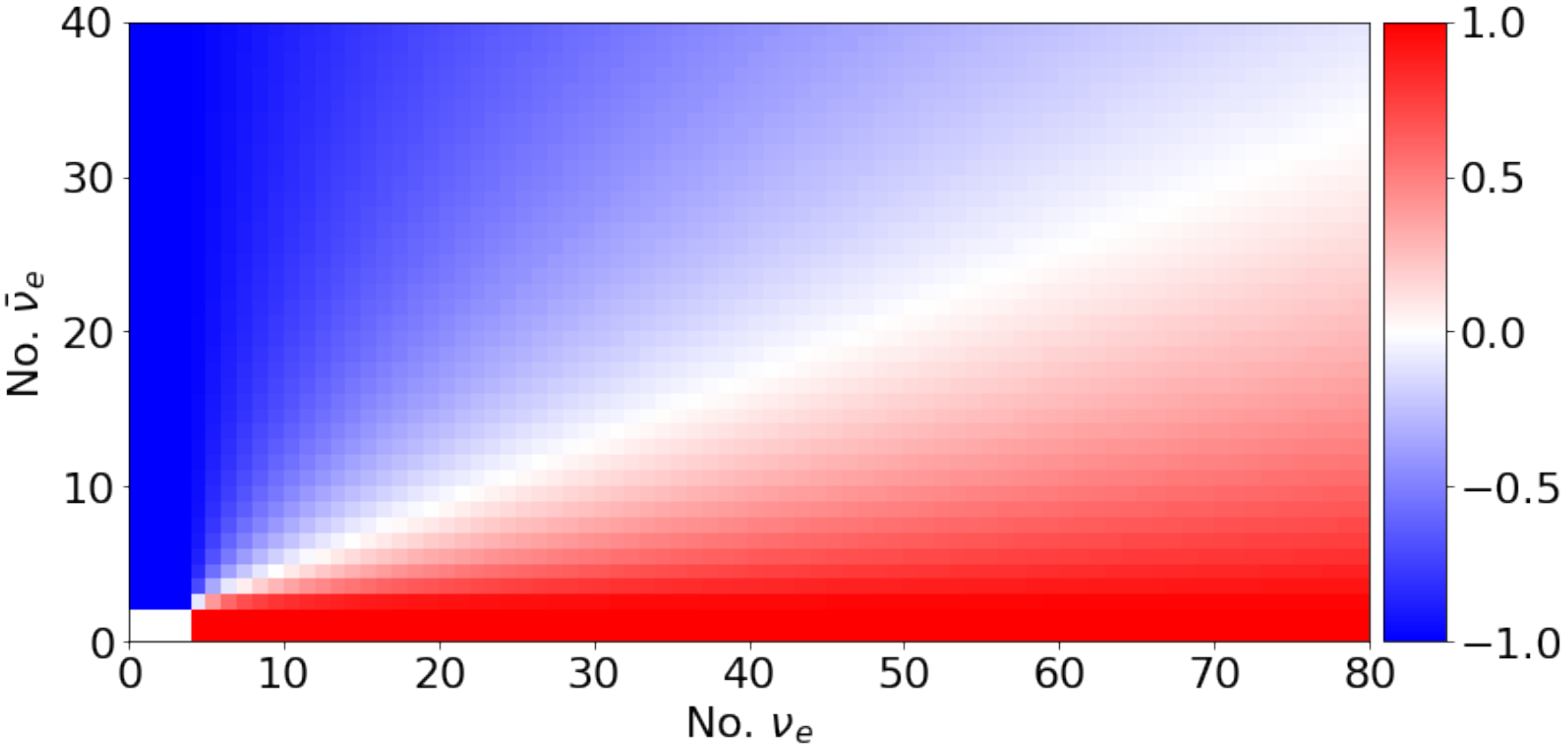
Asymmetry vs Number of Events

Total theoretical;
Not pseudo experiment..

$$P_{FHC} = \frac{n_{\nu_e} - b_{FHC}}{\mu_{FHC}}$$

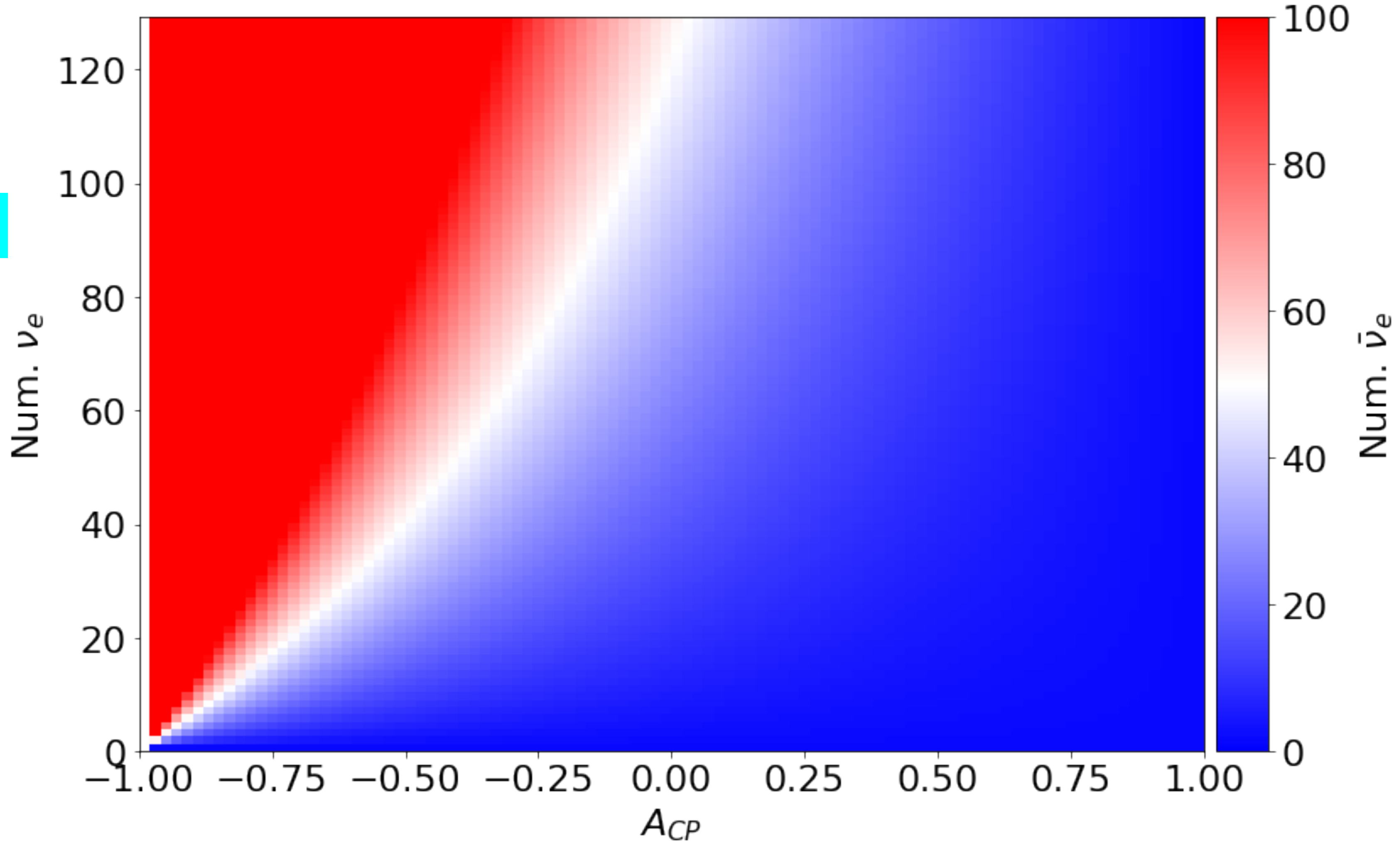
$$P_{RHC} = \frac{n_{\bar{\nu}_e} - b_{RHC}}{\mu_{RHC}}$$

$$\mathscr{A}_{CP} = \frac{P_{FHC} - P_{RHC}}{P_{FHC} + P_{RHC}}$$



Expected antineutrino vs $\mathscr{A}_{CP}$

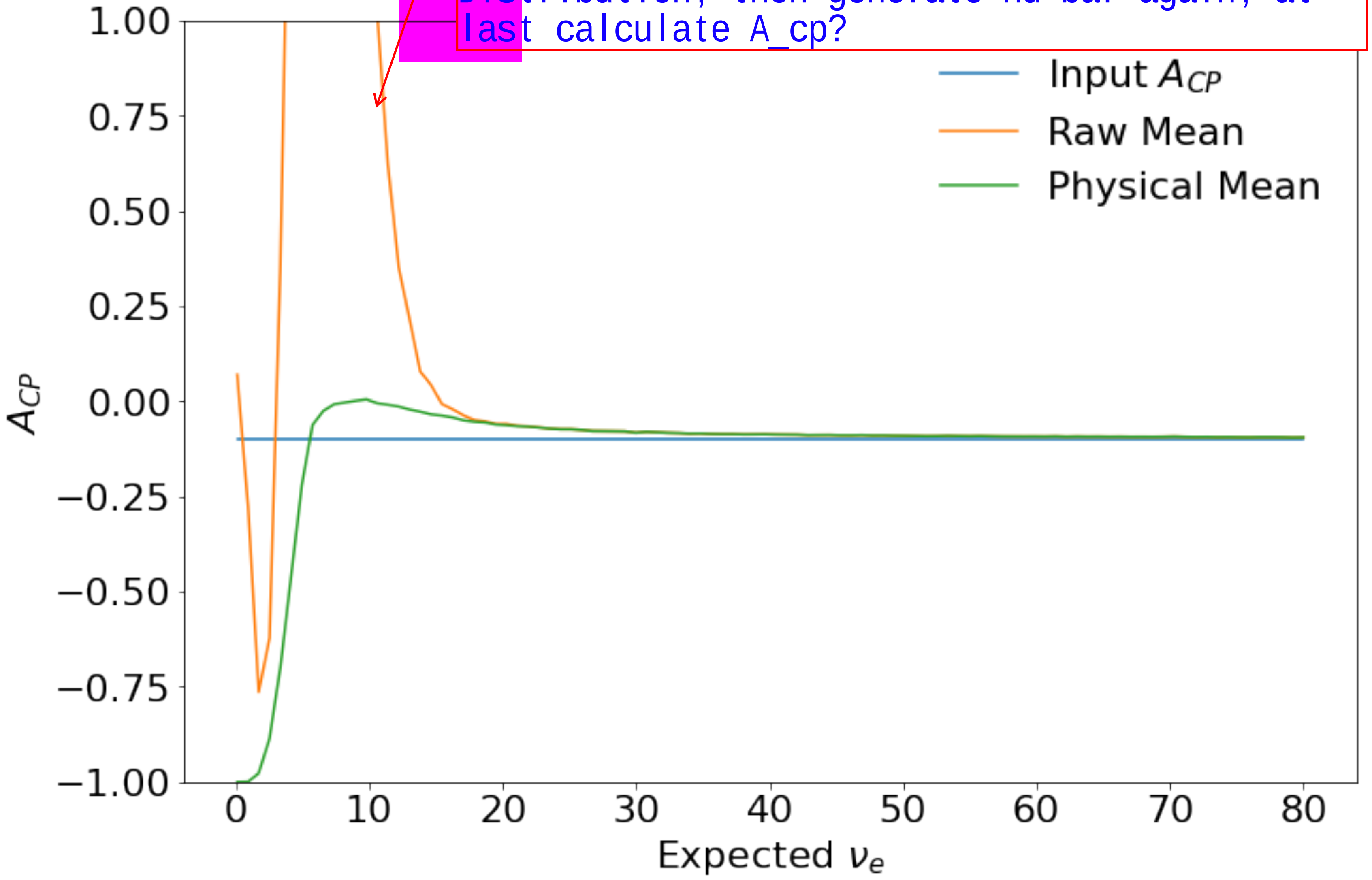
- Of course we can remake this plot looking at the expected number of antineutrinos vs true $\mathscr{A}_{CP}$
- Now we could look at how the measured $\mathscr{A}_{CP}$ relates to the input $\mathscr{A}_{CP}$ across a range of expected number of neutrinos



$$\mathscr{A}_{CP} = -0.1$$

- Now we can look at what is the mean measured $\mathscr{A}_{CP}$ for a fixed true $\mathscr{A}_{CP}$ as a function of number of expected neutrinos

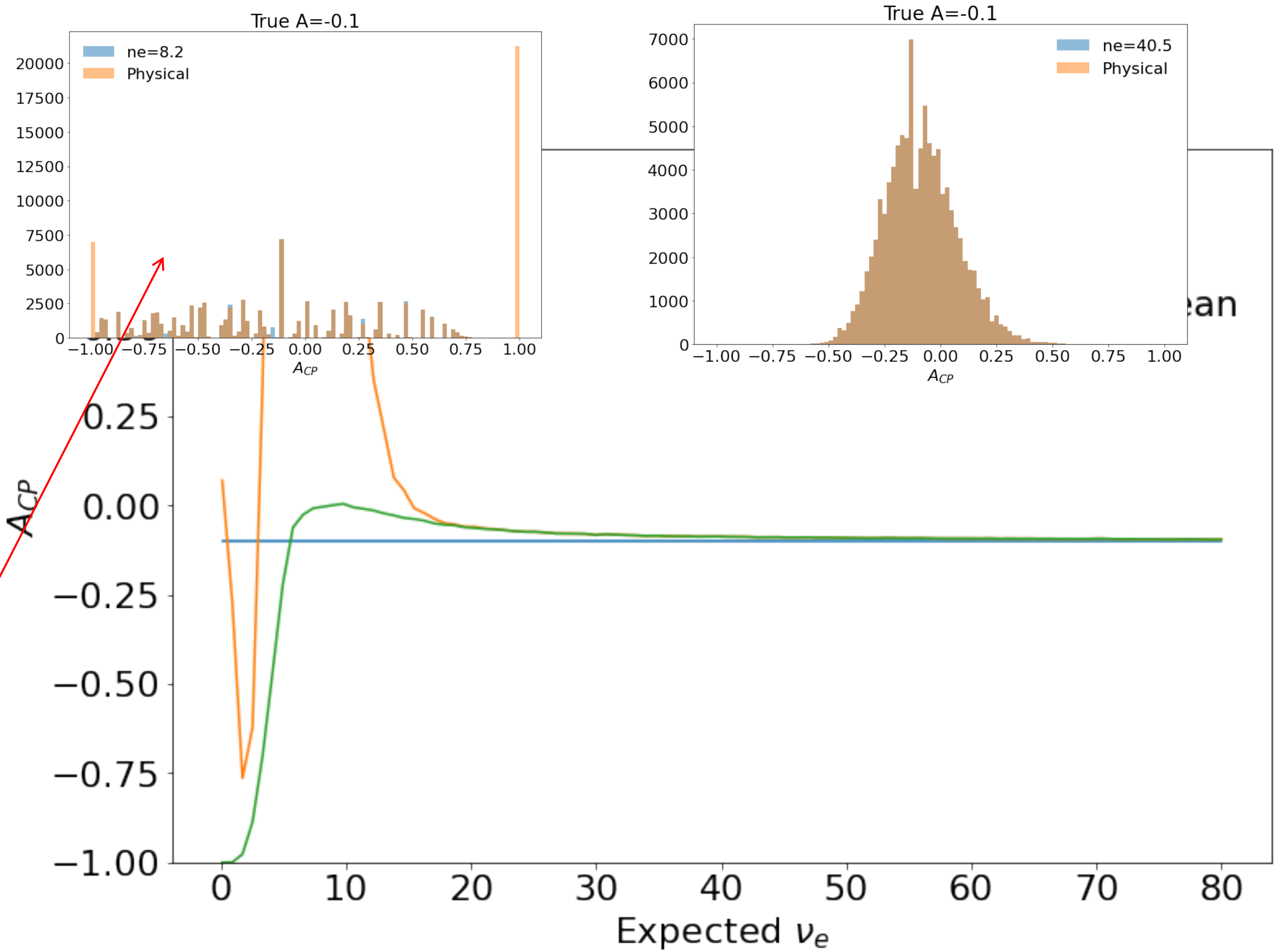
How to get this figure? For certain delta and n_v, we can calculate the num of n_u-bar, then we take num of nu-bar as lambda for a Poisson Distribution, then generate nu-bar again, at last calculate A_cp?



$$\mathscr{A}_{CP} = -0.1$$

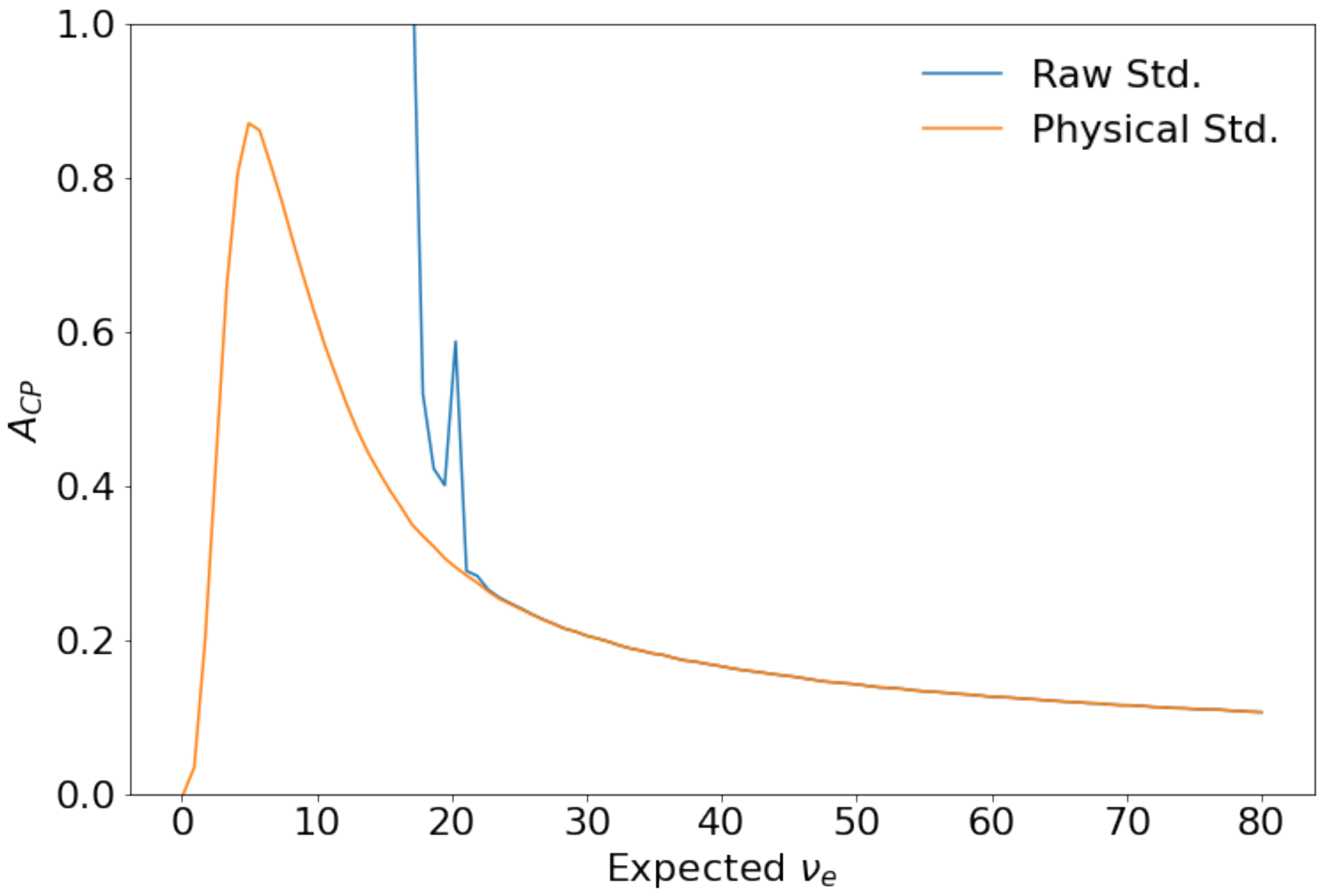
- Now we can look at what is the mean measured $\mathscr{A}_{CP}$ for a fixed true $\mathscr{A}_{CP}$ as a function of number of expected neutrinos

Dark brown
Light Brown
What is "physical"?



$\mathscr{A}_{CP} = -0.1$

- Now we can look at what is the mean measured $\mathscr{A}_{CP}$ for a fixed true $\mathscr{A}_{CP}$ as a function of number of expected neutrinos
- Note the width (here std. dev.) of these distributions can not be used (directly) to determine the error on $\mathscr{A}_{CP}$
- This is closer to the horizontal lines on the



Summary

- The way we are doing the errors on the asymmetry (if we restrict the values to be physical) is reasonable.
- What we are doing is an interesting little corner of statistics, see for example:
  - “Distribution of the ratio of two Poisson random variables”, T. F. Griffin (1992)
  - “Bayesian Estimation of Hardness Ratios: Modeling and Computations”, Park et al. (2006)