

Technische Universität München
Physics department



Bachelor's thesis

Development of a Bayesian Analysis Toolkit

A new error band display style for distributions of numbers of events

Mai Xiang

Advisor: Prof. Caldwell, Max-Planck-Institute for Physics

June 4th, 2012

The common way to represent experimental data in comparison with theoretical predictions shows some insufficiencies that may lead to faulty conclusions, which can be relevant for the evaluation of significance of deviations from theoretical values.

Thus a team of physicists have put some effort in developing a new kind of error bar display for distributions of numbers of events. Also, they developed a C++ code for the Bayesian Analysis Toolkit (BAT)¹, in order to make the new kind of display easily accessible to other scientists who welcome the new presentation method.

My bachelor's thesis deals with the calculation methods and the testing of the BAT code and the application of the error band display style on a physics problem.

¹For further information on BAT see [2] or <http://www.mppmu.mpg.de/bat/>.

Contents

1. Introduction	7
2. Theoretical background	11
2.1. Central and smallest interval	11
2.2. Probability distributions	14
2.2.1. Case 1 - negligible uncertainties	14
2.2.2. Case 2 - statistical uncertainties	14
2.2.3. Case 3 - systematic uncertainties	15
2.2.4. Case 4 - systematic and statistical uncertainties	15
3. Testing the code	17
3.1. Test results	18
4. A physics example	29
5. Summary	33
A. Bibliography	34

1. Introduction

For distributions of numbers of events, several problems have been found using the common way to display error bars for experimental data compared to theoretical predictions. Usually these plots look e.g. like the following:

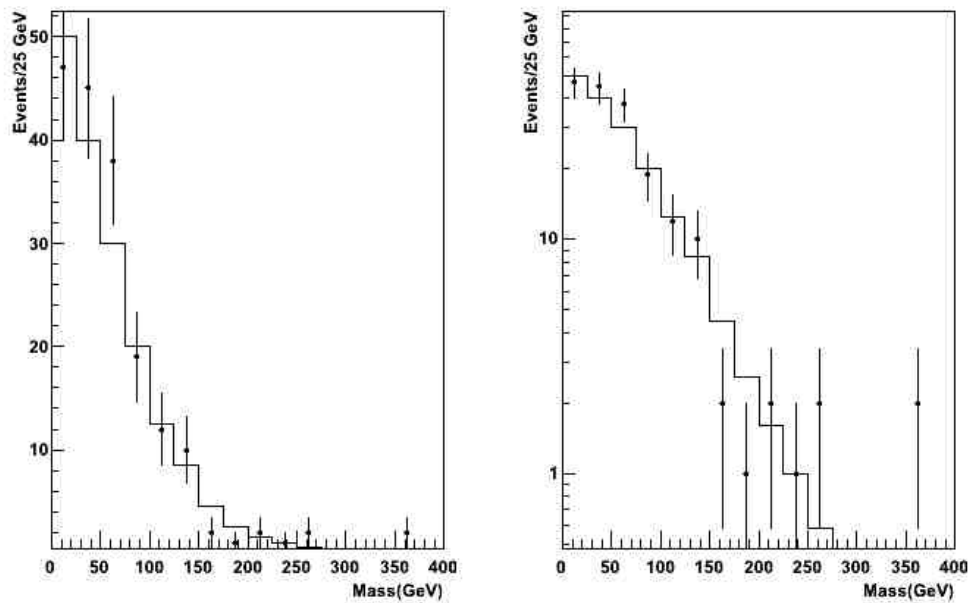


Figure 1.1.: Common way of error bar display, taken from [1]

The common data presentation style uses error bars that are centered at the observed number of events to suggest the agreement between observation and expectation values. Herein several problems can lead to incorrect conclusions concerning the acceptance of the experimental data.

The error bar indicates an uncertainty of the observed value o , which is not what is intended to be shown. The uncertainty that should be displayed is that of the

prediction made according to a model. For the comparison of experimental data to expectations this means: With the common error bar display, we are looking at how probable it is that the observed number of events could have fluctuated to the expectation. But more correctly we would need to take a look at the probability that the expectation could have yielded the number of observed events.

The length of the error bar is usually set to $\pm\sqrt{o}$ to cover a certain amount of probability ($\pm 1\sigma$) for possible values of the mean. Firstly, the probability distribution can certainly be unsymmetric. Secondly, if $o \rightarrow 0$, often, no error bar is plotted although 0 observed events also yield information about possible values of the mean.

It will be necessary to annihilate the insufficiencies of the common way to present distributions of numbers of events in order to gain more correct information from plots like these. Therefore, a new kind of data presentation has been introduced [1] and a code has been written for implementation into BAT for cases where fluctuations of predictions can be modeled with a poisson distribution. This code produces plots according to the new requirements that look like figure 1.2.

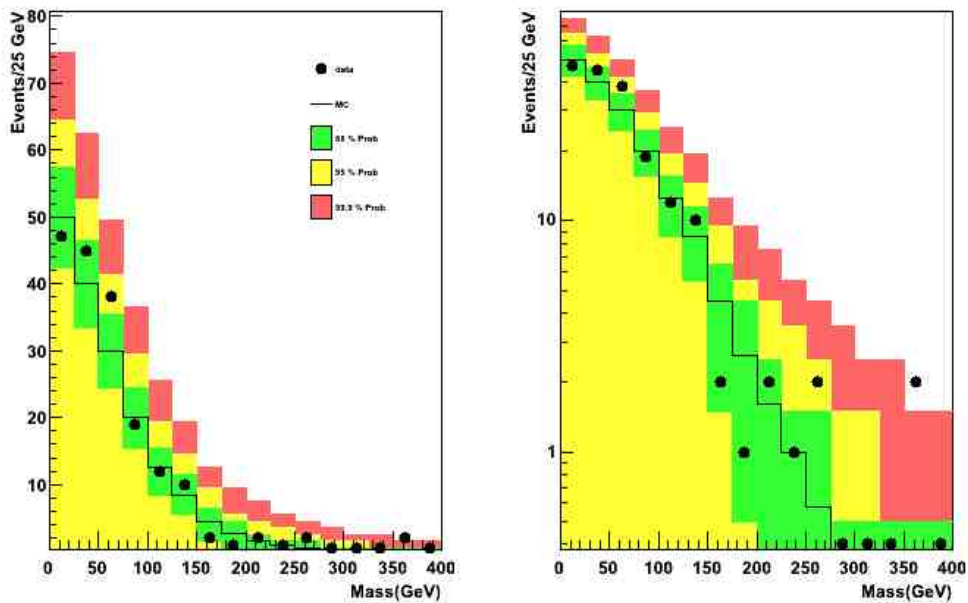


Figure 1.2.: Proposal for new data presentation style, taken from [1]

The error bands show the 68% (green), 95% (yellow) and 99.9% (red) probability ranges of the probability distribution for the number of observed events given the

expectation. We can then take a look at where the experimental data points (black dots) lie compared to the mean (horizontal line) and the probability intervals. This will allow us a better evaluation of the discrepancy between expected and observed numbers of events.

The code for BAT has been tested with a test code written by myself for this thesis. The test code as well as the BAT code are written in C++, since BAT is a C++ project. While the BAT code can be run in ROOT, the test code is a small simple program that can be compiled and run from anywhere independently from ROOT and BAT. It was used to compute the intervals needed for the error band display that were then compared to the intervals generated by the BAT code.

How the error bars are calculated will be described in the following chapter on the theoretical background of my work. Subsequently, the testing of the pre-written code for BAT will be depicted and the test results summarized. Finally, there will be a physics example, showing a case where this error band display could be used.

2. Theoretical background

2.1. Central and smallest interval

The error band display requires the calculation of intervals of observed numbers of events that include e.g. in our case $p = 68.0\%$, 95.0% and 99.9% probability of the probability distribution function $P(o)$. There are different sets of numbers that can be taken as the relevant interval: we provide the smallest and the central sets containing the desired probability, i.e. the smallest and central intervals.

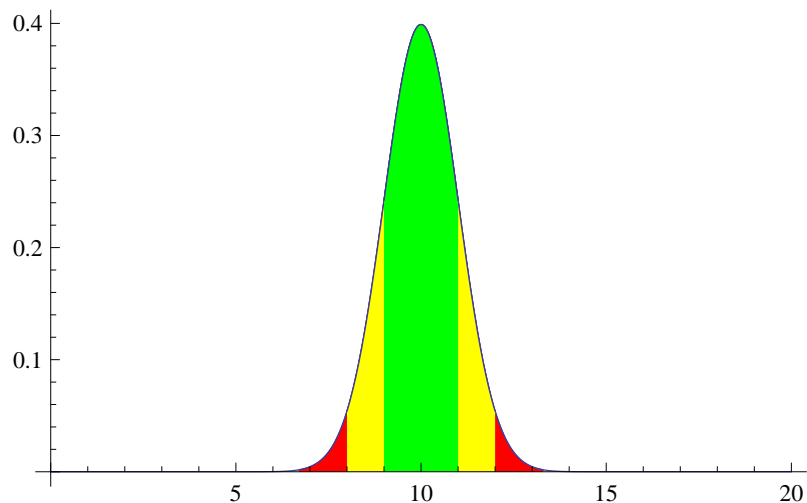


Figure 2.1.: Gaussian pdf with central intervals marked

Figure 2.1 shows an example for a probability distribution function, for which the smallest and central intervals could be calculated. The highlighted areas show the 68% (green), 95% (yellow) and 99.9% (red) central intervals for the simple Gaussian distribution. The calculation methods for both kinds of intervals are described in the following.

Smallest interval

The smallest interval (for a unimodal continuous probability distribution) can be determined by

$$p = \int_{x_1}^{x_2} P(x)dx \quad \text{and} \quad P(x_1) = P(x_2), \quad (2.1)$$

where p is the desired cumulative probability included in the interval, $[x_1, x_2]$ the smallest interval and $P(x)$ is a probability distribution function for the observed number of events given the expectation.

For a discrete case,

$$\sum_{o=o_{min}}^{o_{max}} P(o) \geq p, \quad (2.2)$$

one has to find the correspondent interval by following a simple algorithm:

- Look for the mode (the value with the highest probability), which is the first element of the interval:

$$\mathcal{O} = \{o^*\} = [o^*, o^*].$$

- Check if the cumulative probability is at least the requested probability: $P(\mathcal{O}) \geq p$. If this is the case, then we are done here. Otherwise go on.
- Add the next element to the interval by checking which of the cumulative probabilities of $\mathcal{O}_{\text{down}} = [o_{min} - 1, o_{max}]$ or of $\mathcal{O}_{\text{up}} = [o_{min}, o_{max} + 1]$, where o_{min} is the lower boundary and o_{max} is the upper boundary of the interval \mathcal{O} , is greater than the other. If $P_{\text{down/up}} > P_{\text{up/down}}$ - and $P_{\text{down/up}}$ are the cumulative probabilities of $\mathcal{O}_{\text{down/up}}$ - take $\mathcal{O}_{\text{down/up}}$ as the new interval \mathcal{O} . If the cumulative probabilities are the same, both $o_{min} - 1$ and $o_{max} + 1$ are taken into the interval.
- Repeat the second and if necessary the third step.

Central interval

To find the central interval, the following formula serves as a definition:

$$\frac{1-p}{2} = \int_{x_{min}}^{x_1} P(x)dx = \int_{x_2}^{x_{max}} P(x)dx \quad (2.3)$$

The central interval then is $[x_1, x_2]$ and it also contains the probability p .

Moving to a discrete probability distribution, we have to start at $o_{min} = 0$ and set $o_{min} = o_{min} + 1$ until

$$\sum_{o=0}^{o_{min}} P(o) > \frac{1-p}{2}. \quad (2.4)$$

Taking $o_{min} = o_{min} - 1$, then o_{min} is the lower boundary of our central interval. Starting with $o_{max} = \infty$, we have to do the analogous procedure going downwards according to

$$\sum_{o=o_{max}}^{\infty} P(o) > \frac{1-p}{2}. \quad (2.5)$$

Then we set $o_{max} = o_{max} + 1$.

We can avoid the second sum going up to infinity by going upwards from o_{min} according to

$$\sum_{o=o_{min}}^{o_{max}} P(o) \geq p. \quad (2.6)$$

o_{max} then is the upper boundary of the interval.

2.2. Probability distributions

Here, I would like to give a short introduction to the different cases and their probability distribution functions that were used for the calculation of error bands. Further information can be found in [1].

2.2.1. Case 1 - negligible uncertainties

If the expectation values or the means of the Poisson distributions are known and the uncertainty of these values are negligibly small, the probability distribution for the observed number of events o given a certain mean number of events ν can be described by

$$P(o|\nu) = \frac{e^{-\nu}\nu^o}{o!}. \quad (2.7)$$

Here, we are looking at rare events (small probabilities) and a large number of trials, which is why we are taking a Poisson distribution function. The function above gives us the probability that some o events were observed given the mean number of events ν . The mode then is $o^* = \lfloor \nu \rfloor$, which is the largest integer smaller or equal to ν .

Taking this probability distribution, we can now calculate the smallest and central intervals for this case.

2.2.2. Case 2 - statistical uncertainties

The second case includes a statistical error on the prediction that comes from a Monte Carlo simulation. We call the MC sample result n and the scaling factor s , which the MC prediction has to be divided by to calculate the mean for the data. The probability distribution function (pdf) can be computed using the Law of Total Probability

$$P(o|n, s) = \int P(o|\nu)P(\nu|n, s)d\nu \quad (2.8)$$

with $P(o|\nu) = \frac{e^{-\nu}\nu^o}{o!}$ and $P(\nu|n, s) = \frac{4^n n!}{\sqrt{\pi}(2n)!} s(s\nu)^{n-1/2} e^{-s\nu}$ [1].

The distribution $P(\nu|n, s)$ can be derived with $P(\lambda|n) = \frac{4^n n!}{\sqrt{\pi}(2n)!} \lambda^{n-1/2} e^{-\lambda}$ and $P(\nu|n, s) = P(\lambda|n)d\lambda/d\nu$.

The resulting probability distribution for o , written as a recursive formula that can be used directly in the program code, is

$$P(o+1|n, s) = P(o|n, s) \cdot \frac{(2n+2o+2)(2n+2o+1)}{4(o+1)(1+s)(n+o+1)} \quad (2.9)$$

with $P(o=0|n, s) = \left(\frac{s}{1+s}\right)^{n+1/2}$.

2.2.3. Case 3 - systematic uncertainties

The third case includes a systematic uncertainty on the prediction. The probability distribution can be calculated using

$$P(o|\nu_0, \sigma_\nu) = \frac{\int P(o|\nu)P(\nu|\nu_0, \sigma_\nu)d\nu}{\int P(\nu|\nu_0, \sigma_\nu)d\nu} \quad (2.10)$$

and thus getting

$$P(o|\nu_0, \sigma_\nu) = \frac{\frac{1}{\sqrt{2\pi}\sigma_\nu o!} \int_0^\infty e^{-\nu} \cdot \nu^o \cdot e^{-\frac{(\nu-\nu_0)^2}{2\sigma_\nu^2}} d\nu}{\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_\nu} e^{-\frac{(\nu-\nu_0)^2}{2\sigma_\nu^2}} d\nu}$$

\Rightarrow

$$P(o|\nu_0, \sigma_\nu) = \frac{\int_0^\infty e^{-\nu} \cdot \nu^o e^{-\frac{(\nu-\nu_0)^2}{2\sigma_\nu^2}} d\nu}{o! \int_0^\infty e^{-\frac{(\nu-\nu_0)^2}{2\sigma_\nu^2}} d\nu}. \quad (2.11)$$

$P(o|\nu)$ is of course another Poisson distribution and $P(\nu|\nu_0, \sigma_\nu)$ is a Gaussian distribution with mean ν_0 and standard deviation σ_ν .

The pdf has to be divided by $\int P(\nu|\nu_0, \sigma_\nu)d\nu$ to normalize the total probability to 1.

The integral over ν has to be solved numerically in the code. There are different possibilities for the numerical integration. Which ones were used in the code for BAT and my test code are mentioned in the next chapter.

2.2.4. Case 4 - systematic and statistical uncertainties

The last case allows systematic and statistical uncertainties. We assume the systematic uncertainty lies on the scaling factor s and the distribution of s can be modeled as a Gaussian, so that

$$P(s|s_0, \sigma_s) = \frac{1}{\sqrt{2\pi}\sigma_s} e^{-\frac{(s-s_0)^2}{2\sigma_s^2}} \quad (2.12)$$

with mean s_0 and standard deviation σ_s .

The pdf for o comes from

$$P(o|n, s, \sigma_s) = \int P(o|\nu) \left[\int P(\nu|n, s)P(s|s_0, \sigma_s)ds \right] d\nu \quad (2.13)$$

with $P(\nu|n, s) = \frac{4^n n!}{\sqrt{\pi}(2n)!} s(s\nu)^{n-1/2} e^{-s\nu}$ (cf. case 2) and the Poisson distribution $P(o|\nu) = \frac{e^{-\nu} \nu^o}{o!}$.

This results in the pdf for o

$$P(o|n, s_0, \sigma_s) = \int \frac{e^{-\nu} \nu^o}{o!} \left[\int \frac{4^n n!}{\sqrt{\pi}(2n)!} s(s\nu)^{n-1/2} e^{-s\nu} \frac{1}{\sqrt{2\pi}\sigma_s} e^{-\frac{(s-s_0)^2}{2\sigma_s^2}} ds \right] d\nu. \quad (2.14)$$

These integrations are also to be solved numerically.

3. Testing the code

The main part of my work was to test the code for BAT, i.e. to check the error bands computed by the code for all cases.

In order to do that, I wrote my own routine to check the central and smallest intervals of a broad variety of different numbers in all four cases mentioned above. The test code calculates both kinds of intervals with 68%, 95% and 99.9% cumulative probability included.

At this point, the BAT code has not been implemented into BAT itself yet, but it can be loaded and used in a ROOT session.

For the first two case tests, only little problems were found. For example, there was no way to have the code calculate error bands for cases where a statistical uncertainty exists and the prediction comes from a Monte Carlo simulation with scaling factor $s = 1$. This has been included after a discussion with the code author.

The third case is missing in the BAT code, but it will be added as soon as possible. The case with both systematic and statistical uncertainties is there, but the numerical integration is done with a very simple approximation. The author will change that to using a numerical integration routine from ROOT for more precision.

3.1. Test results

Case 1

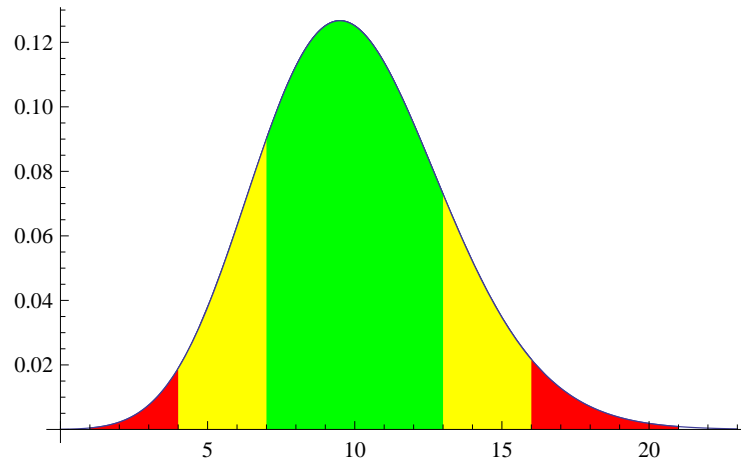


Figure 3.1.: Case 1 plot of $P(o|\nu)$ for $\nu = 10$ with smallest intervals

Figure 3.1 shows a pdf for the case with negligible uncertainty. It is a simple Poisson distribution with mean $\nu = 10$ here. The larger ν gets the more similar to a Gaussian distribution, i.e. the more symmetric the pdf becomes.

For the first case, the intervals calculated by the BAT code show no deviations from the test code except for large numbers for the mean. The BAT code uses a Gaussian distribution for values of the mean larger than 50, whereas my own code does not include such an approximation for large numbers. Therefore small differences are usual. As one can see in table 3.1, the differences for the lower and upper boundary of the intervals are small enough to be explicable with the differences caused by the usage of a Gaussian compared to a Poisson distribution function.

case	ν	$p[\%]$	test code		BAT code	
			$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$	$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$
1	0.1	68	(0,0)	(0,0)	(0,0)	(0,0)
		95	(0,1)	(0,1)	(0,1)	(0,1)
		99.9	(0,2)	(0,2)	(0,2)	(0,2)
	1	68	(0,2)	(0,1)	(0,2)	(0,1)
		95	(0,3)	(0,3)	(0,3)	(0,3)
		99.9	(0,6)	(0,5)	(0,6)	(0,5)
	10	68	(7,13)	(7,13)	(7,13)	(7,13)
		95	(4,17)	(4,16)	(4,17)	(4,16)
		99.9	(2,22)	(1,21)	(2,22)	(1,21)
	100	68	(90,110)	(90,109)	(90,109)	(90,109)
		95	(81,120)	(81,120)	(80,119)	(80,119)
		99.9	(69,134)	(69,134)	(67,132)	(67,132)
	1000	68	(969,1031)	(969,1031)	(969,1030)	(969,1030)
		95	(938,1062)	(939,1062)	(938,1061)	(938,1061)
		99.9	(898,1106)	(897,1105)	(896,1103)	(896,1103)

Table 3.1.: Test results for case 1

Case 2

For cases with non-negligible statistical uncertainties on the prediction based on a Monte Carlo simulation, the second case was designed. Figure 3.2 shows the pdf for $n = 10$ and $s = 1$. Comparing it to the graph in figure 3.1 for the first case, one can see the influence of the statistical uncertainty. The distribution is broader and the difference to a Gaussian distribution is more apparent - it is less symmetric than the function in figure 3.1. The maximum value is not as high as before, but it also lies at $o = 10$.

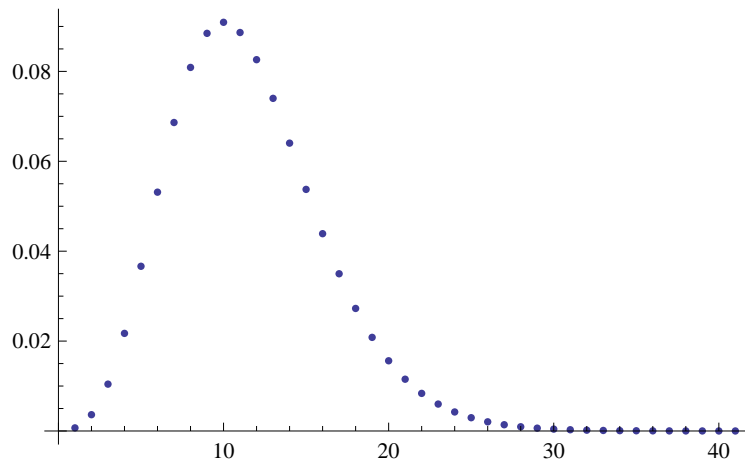


Figure 3.2.: Case 2 plot of $P(o|n, s)$ for $n = 10$ and $s = 1$

The second case test results can be found in table 3.2. The slight differences for large values of the prediction - here $n = 1000$ and $s = 10$ - are again the result of the Gaussian approximation used in the BAT code.

The influence of the statistical uncertainty is very obvious when looking at e.g. $n = 10$, $s = 1$ and the 99.9% intervals and comparing these to the 99.9% intervals for case 1. The intervals go up to about 30, while in the first case, the upper boundaries go just above 20.

On the whole, one can see the agreement between test code results and BAT code output. After having checked these cases, we can prepare to implement the code in BAT.

case	n	s	p [%]	test code		BAT code	
				$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$	$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$
2	10	1	68	(6,15)	(6,14)	(6,15)	(6,14)
			95	(3,21)	(3,20)	(3,21)	(3,20)
			99.9	(0,31)	(0,29)	(0,31)	(0,29)
	10	10	68	(0,2)	(0,1)	(0,2)	(0,1)
			95	(0,4)	(0,3)	(0,4)	(0,3)
			99.9	(0,6)	(0,6)	(0,6)	(0,6)
	100	10	68	(7,13)	(7,13)	(7,13)	(7,13)
			95	(4,17)	(4,16)	(4,17)	(4,16)
			99.9	(1,23)	(1,22)	(1,23)	(1,22)
	1000	10	68	(90,110)	(90,110)	(90,109)	(90,109)
			95	(80,121)	(80,121)	(80,119)	(80,119)
			99.9	(68,136)	(68,136)	(67,132)	(67,132)

Table 3.2.: Test results for case 2

Case 3

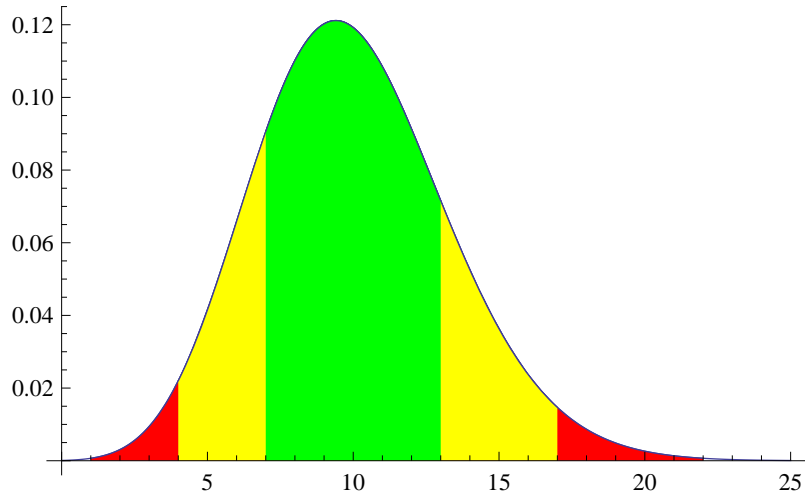


Figure 3.3.: Case 3 plot of $P(o|\nu_0, \sigma_\nu)$ for $\nu_0 = 10$ and $\sigma_\nu = 10\% \cdot \nu_0$ with central intervals

Case three includes (only) a systematic uncertainty on the prediction. Figure 3.3 shows the pdf $P(o|\nu_0, \sigma_\nu)$ for $\nu_0 = 10$ and 10% uncertainty on the prediction ν_0 .

The curve looks very much like the one from case one. The influence of the systematic uncertainty seems to be limited for small numbers of ν_0 . Large values of the mean have much broader probability distributions, which will be taken a look at in the following.

This case has not been included in the BAT code yet, and thus, only the results from the test code are presented in table 3.3. The case will be implemented into the BAT code as soon as possible and before implementation into BAT itself.

The range of the computed intervals for $\nu_0 = 100$ and $\sigma_\nu = 30\% \cdot \nu_0$ might look surprisingly large at first, but examining the probability distribution function, the intervals turn out to be correct. The pdf has been checked with Wolfram Mathematica[®] and is shown in 3.4. You can see that the probability distribution is widely spread around 100, which is logical of course since the width of the curve is directly proportional to σ_ν .

Comparing the distribution to the case one distribution $P(o|\nu)$ with $\nu = 100$ in figure 3.5, one can see the big difference in the width of the curve. Case one has a higher and more narrow peak. This explains why the intervals for case three have to be much more widely ranged than for case one.

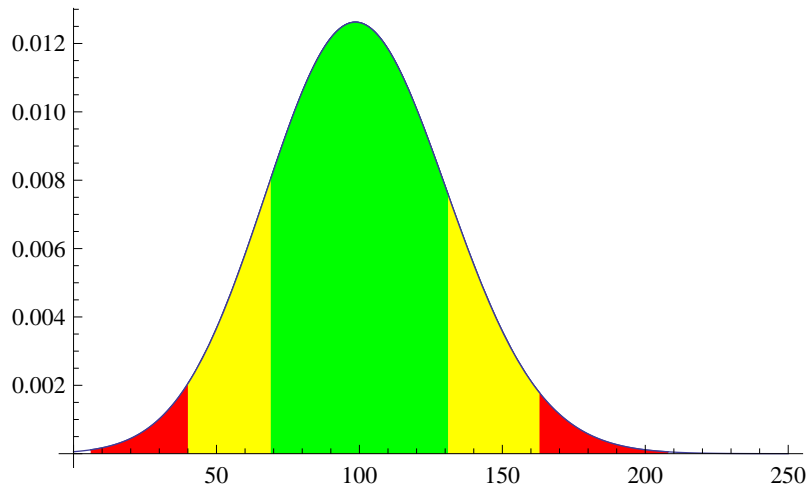


Figure 3.4.: Case 3 plot of $P(o|\nu_0, \sigma_\nu)$ for $\nu_0 = 100$ and $\sigma_\nu = 30\% \cdot \nu_0$ with central intervals

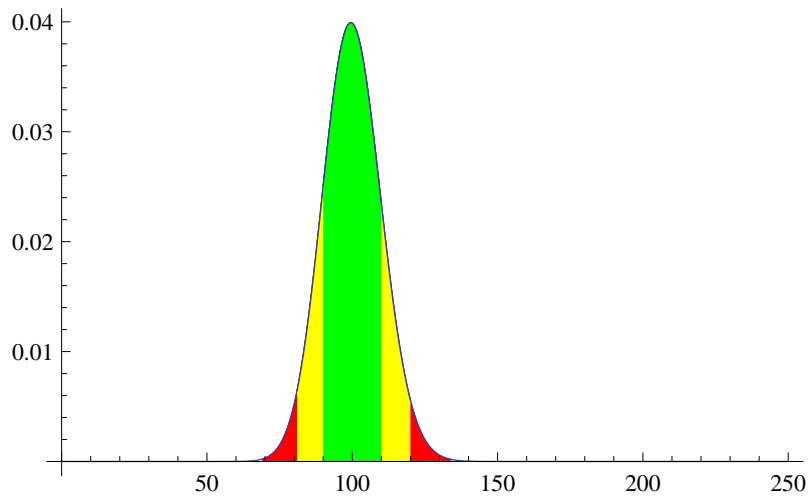


Figure 3.5.: Case 1 plot of $P(o|\nu)$ for $\nu = 100$ with central intervals

case	ν_0	$\sigma_\nu[\% \cdot \nu_0]$	$p[\%]$	test code	
				$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$
3	0.1	10	68	(0,0)	(0,0)
			95	(0,1)	(0,1)
			99.9	(0,2)	(0,2)
	0.1	30	68	(0,0)	(0,0)
			95	(0,1)	(0,1)
			99.9	(0,2)	(0,2)
	10	10	68	(7,13)	(7,13)
			95	(4,17)	(4,16)
			99.9	(1,22)	(1,22)
100	30	68	(69,131)	(68,130)	
		95	(40,163)	(39,162)	
		99.9	(6,208)	(3,204)	

Table 3.3.: Test results for case 3

Case 4

The fourth case with both systematic and statistical uncertainties is tested here.

A problem with this case is the computation speed. Both the test code and the BAT code are very slow calculating the intervals for this case.

case	n	s_0	$\sigma_s[\% \cdot s_0]$	my code		BAT code	
				$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$	$\mathcal{O}_{central}$	$\mathcal{O}_{smallest}$
4	10	1	10	(6,15)	(5,14)	(6,15)	(6,14)
				(3,21)	(3,20)	(3,21)	(3,20)
				(0,30)	(0,30)	(0,31)	(0,29)
	10	10	10	(0,2)	(0,1)	(0,2)	(0,1)
				(0,4)	(0,3)	(0,4)	(0,3)
				(0,6)	(0,6)	(0,6)	(0,6)
	10	10	30	(0,2)	(0,1)	(0,2)	(0,1)
				(0,4)	(0,4)	(0,4)	(0,4)
				(0,16)	(0,16)	(0,16)	(0,11)
	100	10	30	(6,16)	(5,14)		
				(3,27)	(2,22)		
				(1,257)	(0,139)		

Table 3.4.: Test results for case 4

The intervals for case four are summarized in table 3.4. The test code results and the BAT code results agree with each other pretty well. Small deviations can be explained by the difference in the calculation method for the numerical integrations. For the last case with $n = 100$, $s_0 = 10$ and $\sigma_s = 30\% \cdot s_0$, the BAT code failed to calculate all the probabilities needed for the cumulative probability. Some adjustments will have to be made to solve this problem. Also, for $n = 10$, $s_0 = 10$ and $\sigma_s = 30\% \cdot s_0$ the 99.9% smallest interval is too small, which is probably a result of the rounding error of the numerical integration. This problem might be solved when the integration method is changed.

The 99.9% intervals for $n = 100$, $s_0 = 10$ and $\sigma_s = 30\% \cdot s_0$ seem surprisingly large, but calculating them with Wolfram Mathematica[®] gives the same results.

A reason why the upper boundaries for the central and smallest intervals calculated by the test code differ so much could be that the central one begins at $o = 1$

and the smallest at $o = 0$. This already makes a difference, since the distribution falls to low values very fast, as can be seen in figure 3.6, and contributions from about $o = 45$ on are already below the value at 0 of $P(0|100, 10, 3) = 0.00031807$, which is visible in figure 3.7. The probability values continue to fall fast, so that the contributions of increasing o decrease quickly. Therefore, the larger o gets, the more weight $P(0|100, 10, 3)$ has in the cumulative probability and thus the bigger the difference it makes, whether or not this value is included. For the 99.9% interval, this is visible in the large interval difference.

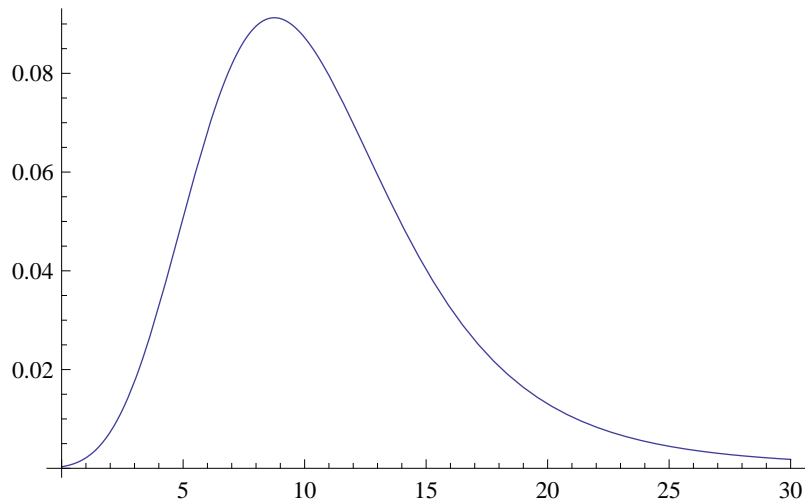


Figure 3.6.: Case 4 plot of $P(o|n, s_0, \sigma_s)$ for $n = 100$, $s_0 = 10$ and $\sigma_s = 30\% \cdot s_0$ for a range of $o = 0 - 30$

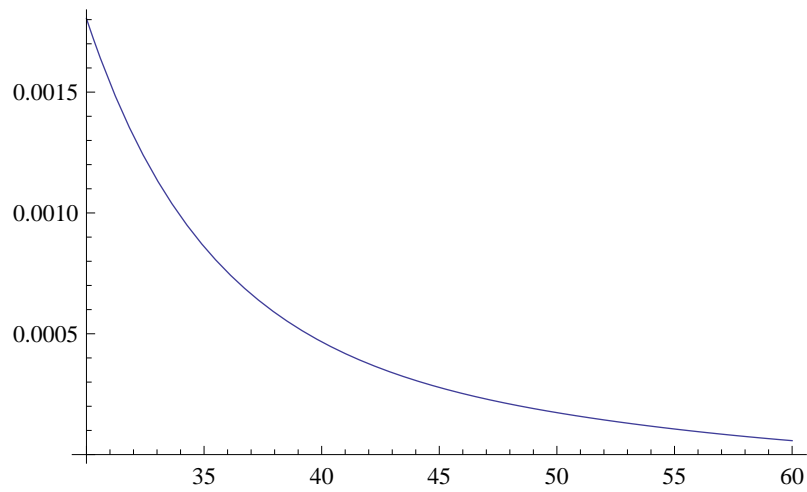


Figure 3.7.: Case 4 plot of $P(o|n, s_0, \sigma_s)$ for $n = 100$, $s_0 = 10$ and $\sigma_s = 30\% \cdot s_0$ for a range of $o = 30 - 60$

4. A physics example

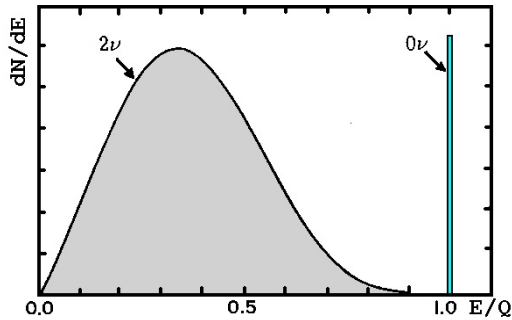


Figure 4.1.: Signal shape for double beta decay, taken from [3]

Experimental data might be processed before it will be displayed in a final histogram. In this following example, data has been put into the template fitter in BAT so that for each bin, a probability distribution has been computed. It is possible to take this output for the calculation of the error bands, i.e. the central and smallest intervals. This saves some work, since the probability distribution function values are already given and there is no need for calculating them once more.

Figure 4.3 shows output from the template fitter in BAT of some fake data set.

The physical background of this example is the current research on neutrinoless double beta decay of the germanium isotope ^{76}Ge from GERDA (GERmanium Detector Array). The neutrinoless double beta decay ($0\nu\beta\beta$ -decay) is a rare process, which is only possible if the neutrino particle is its own anti-particle, i.e. if the neutrino is a Majorana fermion. Two neutrons will be converted into protons via β^- -decay according to $n \rightarrow p + e^- + \bar{\nu}_e$, and the neutrinos will annihilate themselves because of their Majorana particle character, thus the whole process can be written as $2n \rightarrow 2p + 2e^-$. Since the $0\nu\beta\beta$ -decay is a rare event, the measured rate will be very small and the spectrum will

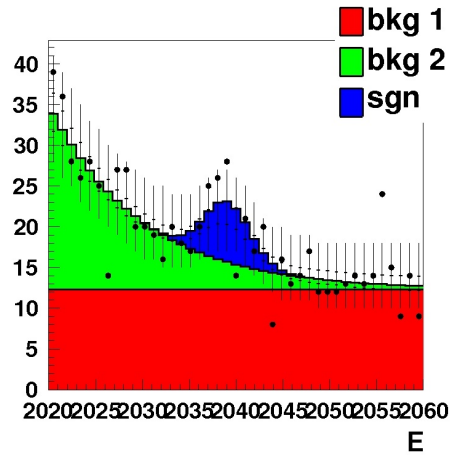


Figure 4.2.: Example for data set, taken from [3]

be dominated by background. The shape of the signal is known very well (fig. 4.1). The expected Q -value lies at 2039keV.

In figure 4.2, two different shapes of background and the signal are shown. Since the shapes are known very well, examples for data sets can be generated.

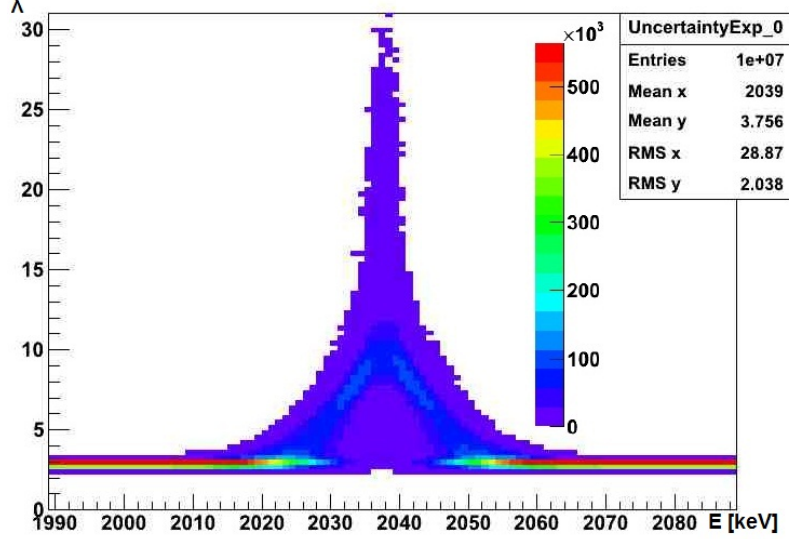


Figure 4.3.: Data run through the template fitter in BAT

In our example (see fig. 4.3), the background has been assumed to be flat. On top of the background, a Gaussian signal has been generated. The number of events per bin fluctuate according to a Poisson distribution. A fit for the data set can be generated with BAT.

The distribution has its maximum at 2039keV (Q -value for $0\nu\beta\beta$ -decay of ^{76}Ge) and sits on a background with constant number of counts. The colors hint at the probabilities: the redder a bin entry, the higher the probability for this value. The probability displayed here is $P(\nu_i|E_i)$, which is the probability to have ν_i expected numbers of events in a certain bin at E_i (or ΔE_i).

To calculate the probability distribution for the observed numbers of events, we equivalently to (2.7) take

$$P(o|E_i) = \int P(o|\nu_i)P(\nu_i|E_i)d\nu_i. \quad (4.1)$$

Thus, we can calculate $P(o)$ for each bin using

$$P(o|E_i) = \sum_{\nu_i} \frac{e^{-\nu_i} \nu_i^o}{o!} P(\nu_i|E_i). \quad (4.2)$$

This is then our probability distribution function for o and we can use the same

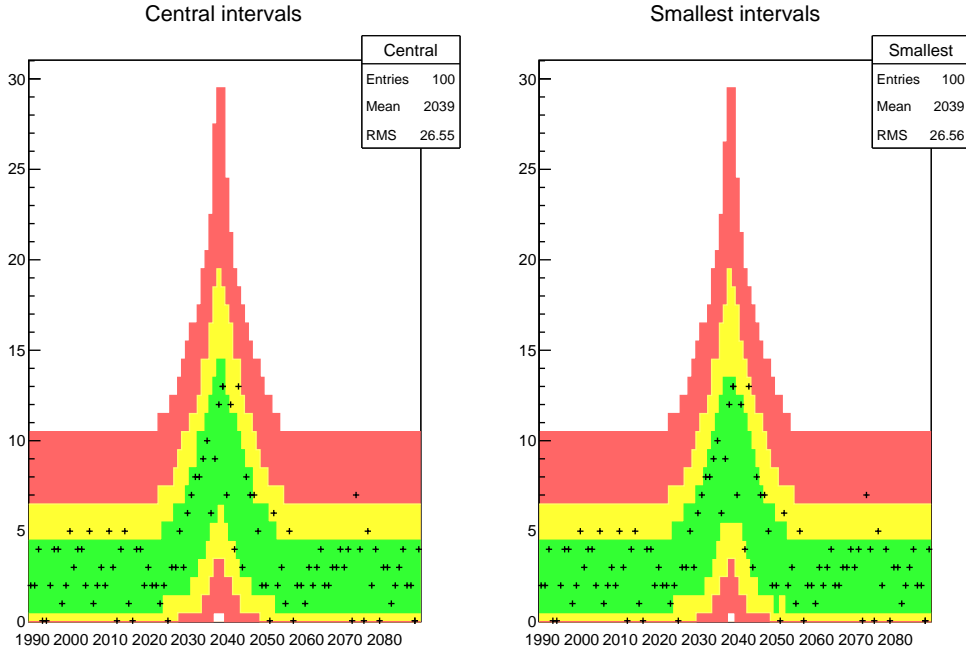


Figure 4.4.: Error band display using output from the BAT template fitter

algorithms from section (2.1) to calculate smallest and central intervals for the new kind of error bar display. This has been done and the result is figure 4.4. The original data set has been displayed in the same histogram, as one would do in a real problem.

If the data set had been a real experimental set of observed numbers, one would have been able to check the agreement between the model and the observations. We can do this as an example here.

At least 68% of the measured data points should lie within the green error band, 95% in the yellow one and 99.9% in the red one. Here, the experimental data confirms the theoretical model very well. All measured data points lie within the three error bands and most of them lie within the green one. The fluctuations from a curve that consists of a flat background and a Gaussian peak are statistically explicable.

5. Summary

The case tests verified the program codes and showed very good computation behaviour for cases with negligible and statistical uncertainties. For cases with numerical integrations to be done, the code needs to be modified for more precision and speed. Especially the last case with both statistical and systematic uncertainties needs great improvement in matters of speed. For the last case, additionally, some adjustments are needed to be done for large values of the prediction.

From our physics example we can see the usefulness of the implementation of this error band display style in BAT. When made easily available, anyone can use the error representation style for analyses. It is very intuitive with the three color bands and shows clearly how much experimental data corresponds to a theoretical model and thus creates the conditions for easy interpretation of comparisons of experiment and theory.

A. Bibliography

- [1] R. Aggarwal, A. Caldwell: Error bars for distributions of numbers of events,,
The European Physical Journal Plus **127**: 24, 2012, Springer
- [2] BAT - The Bayesian Analysis Toolkit, A. Caldwell, D. Kollar, K. Kröninger,
Computer Physics Communications 180, 2009, 2197-2209
- [3] A short introduction to the BAT hypothesis testing and model comparison
tutorial, K. Kröninger

Acknowledgements

I would like to thank

- Prof. A. Caldwell for being a great advisor to a student who is still learning a lot to become a good physicist and for the understanding he has showed during this phase in my life, which not only academically was a step forward but also personally.
- the BAT team and especially R. Aggarwal, F. Beaujean, Dr. K. Kroeninger, N. Becerici-Schmidt, who have helped me with my work. Thank you for helpful discussions, hints and other contributions.
- my parents Guanqi Xiang and Xi Sun and my sisters Ying Xiang-Sailer, Meng Xiang-Gruess and Ting Xiang, whom I can always count on supporting me with their experience and affection.
- my partner Michael Fuchs for his essential support and love during this fast-paced period of time.