

# S1 Coursework: Comparing multi-dimensional likelihood fit with weighted fit using sWeights for parameter estimation

Word Count: 2383

December 18, 2024

## 1 Introduction

This report will compare the statistical power of using the extended maximum likelihood fit and sWeights for weight fitting. The goal of the report is to examine how these two methods differ in performance for parameter estimation of a joint signal and background model for variable X and Y. Section 1 will introduce the Crystal Ball function and derive the normalisation constant used for the X signal model. Section 2 will define the distributions used in this exercise.

## 2 Derivation of Normalisation Constant

The Crystal Ball probability distribution describes the probability of X in terms of Z, where  $Z = (X - \mu)/\sigma$ . The probability distribution is defined as the following piecewise function:

$$P(X; \mu, \sigma, \beta, m) = N \cdot \begin{cases} e^{-Z^2/2}, & \text{for } Z > -\beta, \\ \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \left(\frac{m}{\beta} - \beta - Z\right)^{-m}, & \text{for } Z \leq -\beta \end{cases} \quad (1)$$

To solve for the inverse of the normalization constant  $N$ , Kolmogorov's second axiom of probability is used. The second axiom states that the total probability should sum to 1. The piecewise integrals of the crystal ball probability distribution sums to 1, such that:

$$1 = \int_{-\infty}^{-\infty} P(X; \mu, \sigma, \beta, m) dx$$

Note that the bounds and terms in Equation(1) are expressed in  $Z$ . To rewrite the integral in terms of  $Z$ , a change of variables is needed.

$$Z = \frac{X - \mu}{\sigma}$$

$$\frac{dZ}{dx} = \frac{1}{\sigma}$$

$$dX = \sigma dZ$$

$$1 = \sigma \int_{-\infty}^{-\infty} P(Z; \mu, \sigma, \beta, m), dz$$

Now, the the piecewise integrals from 1 can be substituted. N can be factored out as it is a constant in both integrals.

$$N^{-1} = \sigma \left( \int_{-\beta}^{\infty} e^{-Z^2/2} dz + \int_{-\infty}^{-\beta} \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \left(\frac{m}{\beta} - \beta - Z\right)^{-m} dz \right)$$

To simplify the process, there are two main integrals to solve:  $I_1$  and  $I_2$ .

$$N^{-1} = \sigma(I_1 + I_2) \quad (2)$$

The first integral,  $I_1$ , has a similar form to the Gaussian cumulative distribution function (CDF), which is defined below:

$$\Phi(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} e^{-x^2/2} dx$$

The difference between  $I_1$  and  $\Phi(\beta)$  include the scaling factor  $\frac{1}{\sqrt{2\pi}}$ , and the incorrect bounds. However, due to the symmetrical nature of the Gaussian at  $x = 0$ , flipping the bounds from  $\int_{-\infty}^{\beta}$  to  $\int_{\infty}^{-\beta}$  gives the same area under the curve, such that

$$1 - \Phi(-\beta) = \Phi(\beta)$$

Hence,  $I_1$  can be expressed as:

$$I_1 = 2\pi\Phi(\beta)$$

The second integral,  $I_2$ , can be evaluated as follows:

$$I_2 = \int_{-\infty}^{-\beta} \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \left(\frac{m}{\beta} - \beta - Z\right)^{-m} dz$$

$$I_2 = \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \int_{-\infty}^{-\beta} \left(\frac{m}{\beta} - \beta - Z\right)^{-m} dz$$

$$I_2 = \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \left[ \frac{-1}{-m+1} \left(\frac{m}{\beta} - \beta - Z\right)^{-m+1} \right]_{Z=-\infty}^{Z=-\beta}$$

As  $Z \rightarrow -\infty$

$$\left(\frac{m}{\beta} - \beta - Z\right)^{-m+1} \rightarrow 0$$

As  $Z \rightarrow -\beta$

$$\left(\frac{m}{\beta} - \beta - Z\right)^{-m+1} = \left(\frac{m}{\beta}\right)^{-m+1}$$

Thus, the evaluated integral  $I_2$  is given by:

$$I_2 = \frac{m}{\beta} e^{-\beta^2/2} \frac{1}{m-1}$$

Substituting the expressions for  $I_1$  and  $I_2$  into Equation (2), the final expression for  $N^{-1}$  in terms of  $\beta$ ,  $m$  and  $\Phi(\beta)$  can be written as :

$$N^{-1} = \sigma \left[ \sqrt{2\pi}\Phi(\beta) + \left(\frac{m}{\beta}\right)^m e^{-\beta^2/2} \frac{1}{m-1} \right]$$

The terms could be rearrange to obtain the final form given:

$$N^{-1} = \sigma \left[ \frac{m}{\beta(m-1)} e^{-\beta^2/2} + \sqrt{2\pi}\Phi(\beta) \right] \quad (3)$$

### 3 Part B: Defining a module for the crystal ball function

Equation(4) is the joint probability distribution between  $X$  and  $Y$ , such that  $X$  and  $Y$  contain independent signal and background components. The  $X$  distribution is defined in the range of  $X \in [0, 5]$ , with a signal distribution ( $g_s(x)$ ) given by the truncated crystal ball PDF. The  $X$  background distribution ( $g_b(x)$ ) is defined by a uniform distribution. The  $Y$  distribution is defined in the range of  $Y \in [0, 10]$ , with a signal distribution ( $h_s(y)$ ) given by the truncated exponential PDF. The  $Y$  background distribution ( $h_b(y)$ ) is defined by a truncated normal distribution.

$$\left\{ f(X, Y) = f.g_s(X)h_s(Y) + (1-f).g_b(X)h_b(Y). \right. \quad (4)$$

$g_s(X)$ ,  $g_b(X)$ ,  $h_s(Y)$ ,  $h_b(Y)$  along with the total signal distribution ( $s(X,Y)$ ), total background distribution ( $b(X,Y)$ ), and total joint distribution ( $f(X,Y)$ ) are defined using the scipy.stats library.

```

1 # g_s(X)
2 def signal_x_pdf(x, beta, m, mu, sigma):
3     dist = stats.crystalball(beta=beta, m=m, loc=mu, scale=sigma)
4     normalise = dist.cdf(5) - dist.cdf(0)
5     return dist.pdf(x) / normalise
6
7 # h_s(Y)
8 def signal_y_pdf(y, lambda_val):
9     dist = stats.truncexpon(b=10, scale=1 / lambda_val)
10    normalise = dist.cdf(10) - dist.cdf(0)
11    return dist.pdf(y) / normalise
12
13 # g_b(X)
14 def background_x_pdf(x):
15     return stats.uniform.pdf(x, loc=0, scale=5)
16
17 # h_b(Y)
18 def background_y_pdf(y, mu_b, sigma_b):
19     a = -mu_b / sigma_b
20     b = (10 - mu_b) / sigma_b
21     dist = stats.truncnorm(a=a, b=b, loc=mu_b, scale=sigma_b)
22     return dist.pdf(y)
23
24 # s(X,Y)
25 def signal_pdf(x, y, beta, m, mu, sigma, lambda_val):
26     return signal_x_pdf(x, beta, m, mu, sigma) * signal_y_pdf(y, lambda_val)
27
28 # b(X,Y)
29 def background_pdf(x, y, mu_b, sigma_b):
30     return background_x_pdf(x) * background_y_pdf(y, mu_b, sigma_b)
31
32 # f(X,Y)
33 def joint_pdf(x, y, f, beta, m, mu, sigma, lambda_val, mu_b, sigma_b):
34     return f * signal_pdf(x, y, beta, m, mu, sigma, lambda_val) + \
35         (1 - f) * background_pdf(x, y, mu_b, sigma_b)

```

Listing 1: PDF definitions

To ensure the PDFs are normalised to 1 between the ranges of  $X \in [0, 5]$  and  $Y \in [0, 10]$ , `scipy.integrate` is used. The PDFs are first integrated using the parameter values in Table 1. Subsequently, three more sets of values are used to validate whether the PDF normalised to 1 under different parameter values. All PDFs integrate to 1, up to 4 decimal place precision.

Table 1: Parameter Values Table

Parameter	Value
$\mu$	3.0
$\sigma$	0.3
$\beta$	1.0
$m$	1.4
$f$	0.6
$\lambda$	0.3
$\mu_b$	0
$\sigma_b$	2.5

## 4 Part C: Examine the 1D distributions

To examine the 1D distributions of X and Y and their respective signal and background contributions, X and Y marginal PDFs are redefined as follows:

```
1 # X PDF
2 def x_pdf(x, beta, m, mu, sigma,f):
3     return f * signal_x_pdf(x, beta, m, mu, sigma) + (1-f) * background_x_pdf(x)
4
5 # Y PDF
6
7 def y_pdf(y, lambda_val, mu_b, sigma_b, f):
8     return f * signal_y_pdf(y,lambda_val) + (1-f) * background_y_pdf(y,mu_b,sigma_b)
```

Listing 2: Marginal Distributions

When evaluated with the given parameter values in Table 1, the marginal distributions for X and Y in figure 1 aligns with the expected shape of the distributions. The X marginal distribution is composed of a Gaussian core and Power-law tail distribution, with a uniform background. The Y marginal distribution is composed of a truncated exponential and truncated normal distribution.

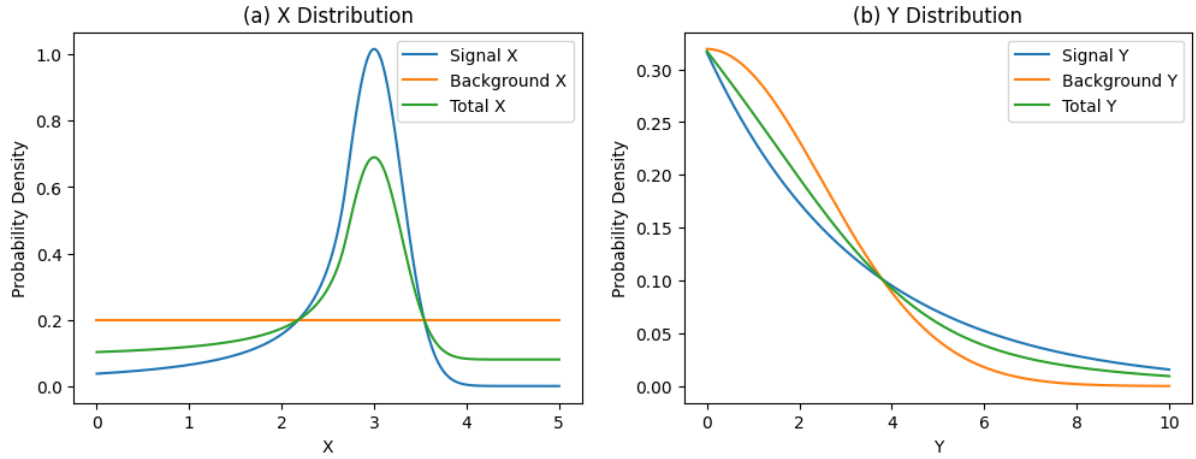


Figure 1: The Marginal distribution and normalised signal and background components for variable a) X and b) Y

The joint distribution of X and Y is shown in figure 2, which shows a maximum probability density of around 0.21 at  $x = 3$  and  $y = 0$  when the parameters in table 1 are used.

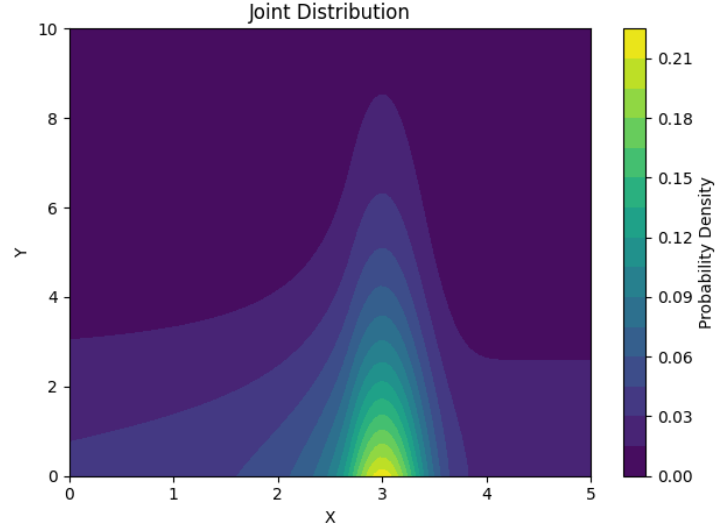


Figure 2: Joint distribution of X and Y

## 5 Part D: Extended Maximum Likelihood fit for parameter estimation

This section showcases the use of the Extended Maximum Likelihood (EML) Estimation. Generally, Maximum Likelihood Estimation aims to maximize the likelihood function with respect to the parameters, to identify values at which the observed dataset is most probable to occur. The extended component of EML incorporates the number of events to be an extra parameter estimated.

To do so, 100,000 samples are drawn from the  $f(X,Y)$  distribution using parameters in table 1. The EML method is then applied to estimate the parameters, and the results will be compared to the true, known parameter values. Firstly, the samples were drawn using an accept-reject generator. The sample distribution is plotted in figure 3, which resembles the distribution in figure 2 as expected.

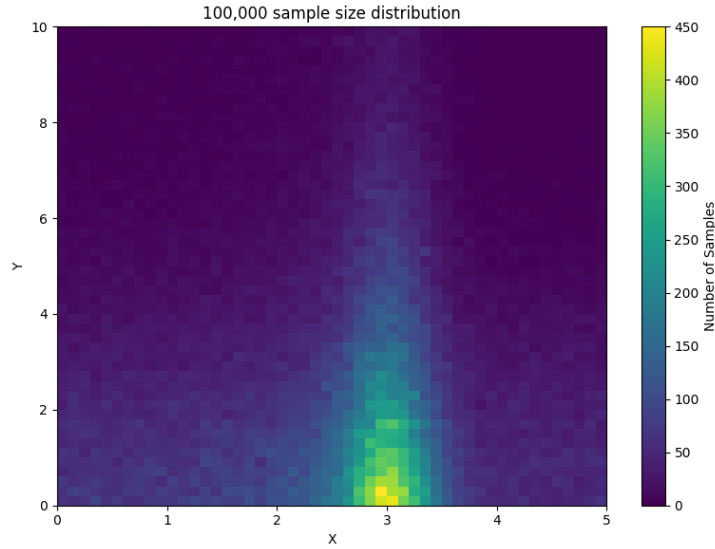


Figure 3: The distribution of the high-statistic sample collected from the joint pdf using the accept-reject generator

The EML fit was then executed using the `iminuit` package. A limit of  $\beta > 0$  and  $m > 1$  was enforced according to the mathematical constraints in  $f(X, Y)$ . This ensured that the algorithm did not converge into an invalid region of the likelihood space.

After convergence, a log likelihood minimum is found at  $-1.46 \times 10^6$ , and the estimates of the parameters are displayed in table2. Next to the parameter estimates are the Hesse errors, which are the symmetric uncertainty intervals of the parameter estimates, and can be used as a measure of parameter uncertainty [1].

Table 2: Parameter estimates and Hesse Errors

Name	Value	Hesse Error
$n$	100.00e3	0.32e3
$f$	0.598	0.004
$\mu$	2.9960	0.0026
$\sigma$	0.3038	0.0024
$\beta$	1.032	0.023
$m$	1.35	0.06
$\lambda$	0.2996	0.0021
$\mu_b$	0.03	0.08
$\sigma_b$	2.49	0.04

The covariance matrix of the parameters shown in figure.4 shows strong negative correlations between  $\sigma_b$  and  $\mu_b$ . There is also a strong positive correlation between  $m$  and  $\beta$ . These correlations are as expected, as these parameters originate from the same signal  $g_s(X)$  and background  $h_b(Y)$  distributions respectively.

	n	f	mu	sigma	beta	m	lambda_val	mu_b	sigma_b
n	1.01e+05	-2.496e-3 (-0.003)	1.331e-3 (0.002)	-1.650e-3 (-0.003)	-18.6e-3 (-0.003)	0.049 (0.002)	-825e-6 (-0.002)	-0.052 (-0.002)	0.0238 (0.002)
f	-2.496e-3 (-0.003)	1.3e-05 (0.000)	-0e-6 (0.000)	3e-6 (0.319)	0.010e-3 (0.121)	-0.088e-3 (-0.385)	2e-6 (0.263)	0.058e-3 (0.197)	-0.033e-3 (-0.241)
mu	1.331e-3 (0.002)	-0e-6 (-0.003)	6.61e-06 (0.000)	-3e-6 (-0.534)	-30e-6 (-0.506)	55e-6 (0.337)	-0e-6 (-0.005)	-3e-6 (-0.014)	2e-6 (0.019)
sigma	-1.650e-3 (-0.003)	3e-6 (0.319)	-3e-6 (-0.534)	6.01e-06 (0.451)	26e-6 (0.451)	-54e-6 (-0.348)	0e-6 (0.085)	19e-6 (0.095)	-11e-6 (-0.122)
beta	-18.6e-3 (-0.003)	0.010e-3 (0.121)	-30e-6 (-0.506)	26e-6 (0.451)	0.000546 (0.000)	-1.3e-3 (-0.876)	1e-6 (0.018)	0.1e-3 (0.053)	-0.1e-3 (-0.071)
m	0.049 (0.002)	-0.088e-3 (-0.385)	55e-6 (0.337)	-54e-6 (-0.348)	-1.3e-3 (-0.876)	0.00399 (0.000)	-13e-6 (-0.096)	-0.001 (-0.123)	0.0004 (0.160)
lambda_val	-825e-6 (-0.002)	2e-6 (0.263)	-0e-6 (-0.005)	0e-6 (0.085)	1e-6 (0.018)	-13e-6 (-0.096)	4.27e-06 (0.062)	10e-6 (0.062)	4e-6 (0.047)
mu_b	-0.052 (-0.002)	0.058e-3 (0.197)	-3e-6 (-0.014)	19e-6 (0.095)	0.1e-3 (0.053)	-0.001 (-0.123)	10e-6 (0.062)	0.00658 (0.000)	-0.0029 (-0.931)
sigma_b	0.0238 (0.002)	-0.033e-3 (-0.241)	2e-6 (0.019)	-11e-6 (-0.122)	-0.1e-3 (-0.071)	0.0004 (0.160)	4e-6 (0.047)	-0.0029 (-0.931)	0.00145 (0.000)

Figure 4: Covariance Matrix of the parameters estimated from the 100,000 high statistics sample

Both running EML and generating the high statistics sample using an accept generator are computationally expensive tasks. Compared to generating a random sample size of 100,000 using `np.random.normal()`, generating 100,000 samples using the accept-reject method is 162 times slower, while fitting 100,000 samples using EML is approximately 202 times slower.

## 6 Part E: Parametric bootstrapping for determining bias and uncertainty of $\lambda$

This section aims to assess the sensitivity of EML for estimating parameter values with increasing sample sizes. In particular, the parameter estimation of the decay constant  $\lambda$  from  $h_s(y)$  with increasing sample sizes will be investigated.

The methodology first requires 250 toys to be made for each sample size, with a poisson variation on each toy's sample size. This is done by bootstrapping the parameter estimates from part D, to create 250 toys with a poisson varied sample size of 500, 1000, 2500, 5000, 10000. EML is then fitted, and the mean  $\lambda$  estimates for each sample size and their uncertainties are analysed.

Even though this section is only concerned with the bias and uncertainty of the decay constant  $\lambda$  from  $h_s(y)$  with increasing sample sizes, it is important to examine the other parameter distributions too. This can ensure that changes in bias and uncertainty of  $\lambda$  is not attributed to dependency on other parameters, but from the variable studied (the sample size, in the context of this study).

All parameter estimates, uncertainty and pulls for all sample sizes are calculated and plotted as histograms in figures 12 - 16 located in the appendix. Generally, parameter estimates were normally distributed, except for  $\mu_b$  and  $m$  estimates. While  $\mu_b$  approached a normal distribution with larger sample sizes in accordance to the central limit theorem,  $\mu$  remained right skewed. This is due to the imposed limit at  $m > 1$ , and the true  $m$  parameter defined as  $m = 1.4$ . Generally, all histograms approach a normal distribution with larger sample sizes, and there is no clear dependency between  $\lambda$  and other parameters by visual inspection. Thus, any uncertainty or bias inferred for  $\lambda$  in later sections is unlikely a result of multicollinearity.

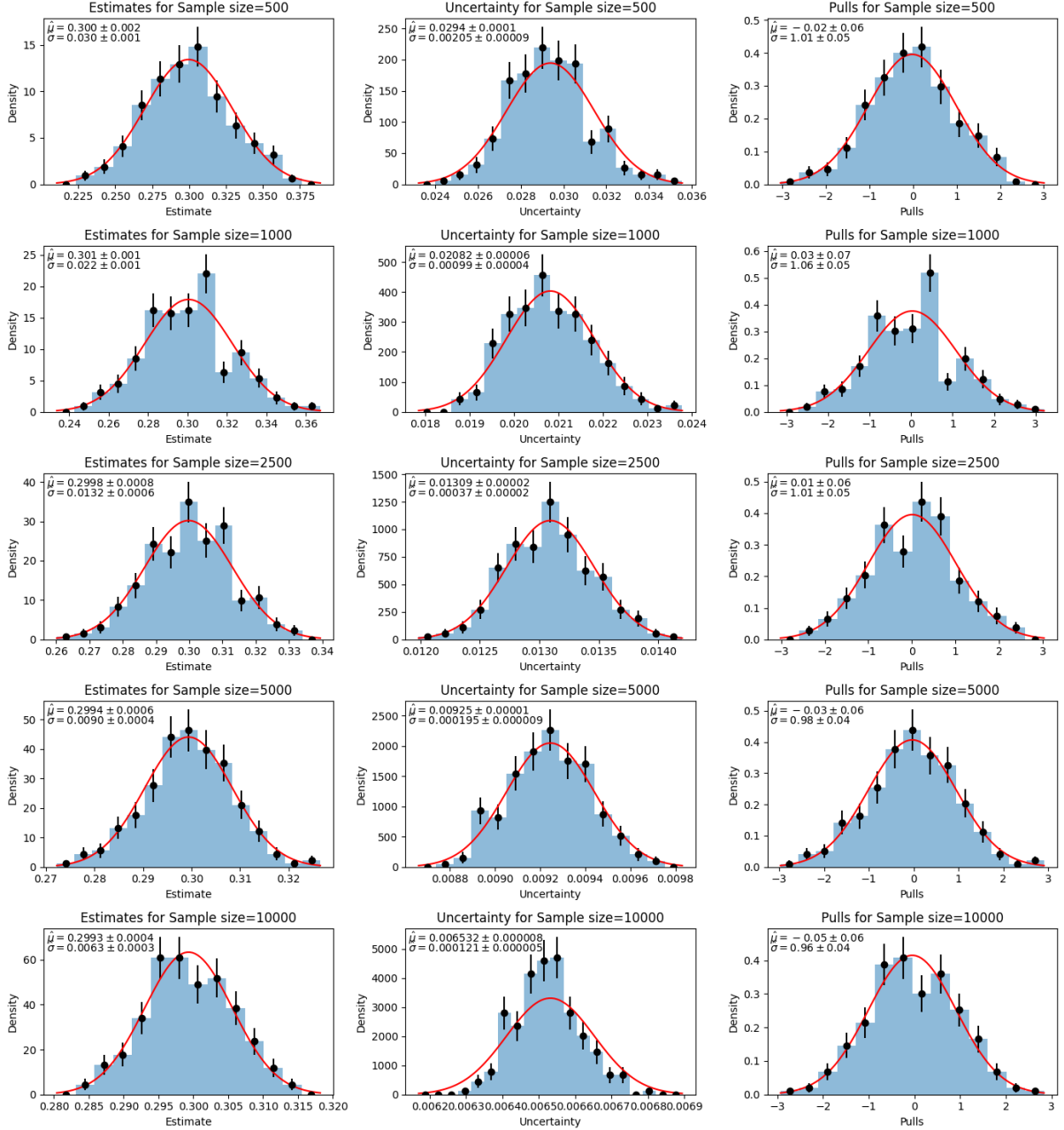


Figure 5: Estimate, Pull, Bias and uncertainty distributions of Lambda

### 6.1 Lambda bias and uncertainty with increasing sample sizes

Figure 5 shows that all  $\lambda$  estimates, uncertainty and pulls are normally distributed. The estimate of  $\lambda$  have a normal distribution centered around a mean of 0.3, and a standard deviation of 0.03 for the 500 sample size. As the sample size gets larger, the  $\lambda$  lambda estimator decreases to 0.2993, and the standard deviation of the estimator decreases by one whole magnitude. This is similarly displayed in the second column of uncertainty plots, showing a whole magnitude decrease in uncertainty for a sample size of 10,000. Finally, the pull distribution, defined by  $\frac{\lambda_{estimate} - \lambda_{true}}{error}$  generally is centered around 0, with a standard deviation of 1. These normally distributed histograms of the pull implies a good quality of fit for  $\lambda$ .

After visual inspection of the fit of  $\lambda$  through histograms, this section will compute the bias of the  $\lambda$  estimator for each sample size. Recall that the bias of the estimator is defined as the difference between the expectation of  $\lambda$  and the true value for a given sample size  $N$ , such that:



$$b_N(\hat{\theta}) = E[\hat{\theta}] - \theta_0$$

$E[\hat{\theta}]$  is calculated by computing the mean of the  $\lambda$  estimates for each sample size, and  $\theta_0$  is 0.2996, as estimated from part D. The estimator, and the bias of each  $\lambda$  estimator is plotted in figure 6. The bias transitions from positive bias in small sample sizes, to negative bias in very large sample sizes such as 10,000. This implies a systematic bias in  $\lambda$  estimation potentially due to the methodology of using EML.

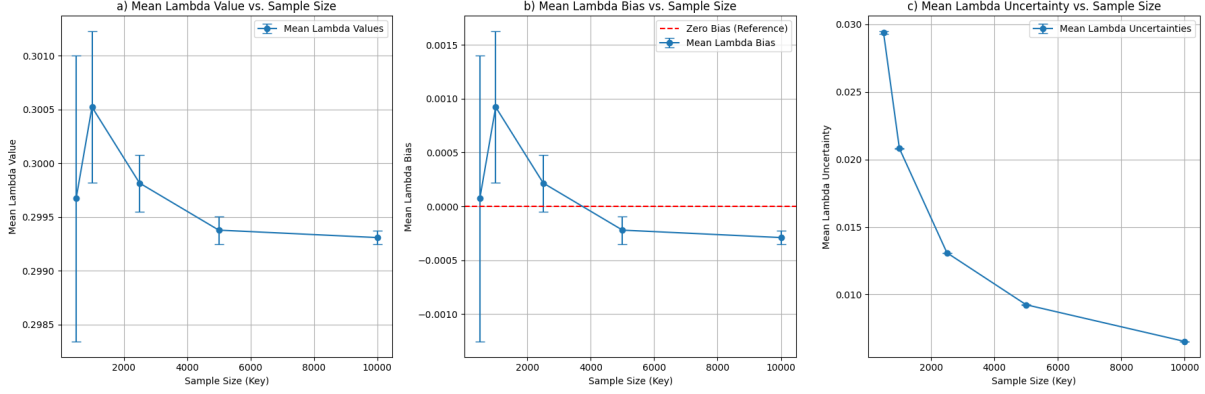


Figure 6: Mean Lambda a) estimator, b) estimator bias and c) uncertainty with increasing sample sizes

Figure 6 also shows that uncertainty of the estimator decreases with increasing sample size. This aligns well with the foundations of a maximum likelihood estimator, which is a consistent estimator with an asymptotic normality property, such that larger sample sizes will lead to a convergence to a normal distribution [3].

## 7 Part F: Using sWeights for $\lambda$ projection

Given the signal and background models are independent in X and Y for  $f(X, Y)$ , sWeights could be used as an alternative for parameter estimation. sWeights is used to differentiate the masked signal from background distribution in Y, given known signal and background contributions in X. This section will aim to use the bootstrapped X variable from part E to project out a signal density in Y using sWeights.

Firstly, an EML fit was performed on the 500, 1000, 2500, 5000, and 10000 bootstrapped values for the X variable. figure 7 is an example of the EML fit, showing the parameters estimated for minimizing the negative log likelihood, and fitting the first toy of the 5000 sample size onto the expected X model curve. Note that two extra parameters, Ns and Nb, are introduced as parameters to be fitted. Ns and Nb represent the number of points belonging to the signal and background fraction respectively.

Name	Value
Ns	$2.98 \times 10^3$
Nb	$1.87 \times 10^3$
$\beta$	0.93
$m$	1.49
$\mu$	3.010
$\sigma$	0.305

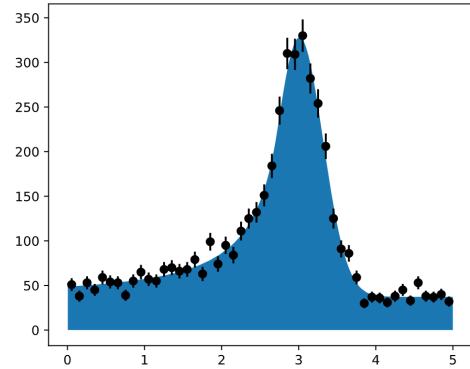


Figure 7: Parameter estimate results for the first toy of the 5000 sample size (left) and the fitted points onto the X model (right)

After obtaining the parameter estimates for all toys in each sample size, SWeight was executed to extract the signal and background weights of the X variable. These weights can be visualised in figure 8. This is representative of figure 1, such that in the region of  $x < 2$ , the background distribution  $g_b(X)$  contributes more to the overall X model than the signal distribution  $g_s(X)$ .  $g_s(X)$  contributes more weight to the X model in the maximum of the Gaussian core. Both signal and background at any point of  $x$  adds up to 1.

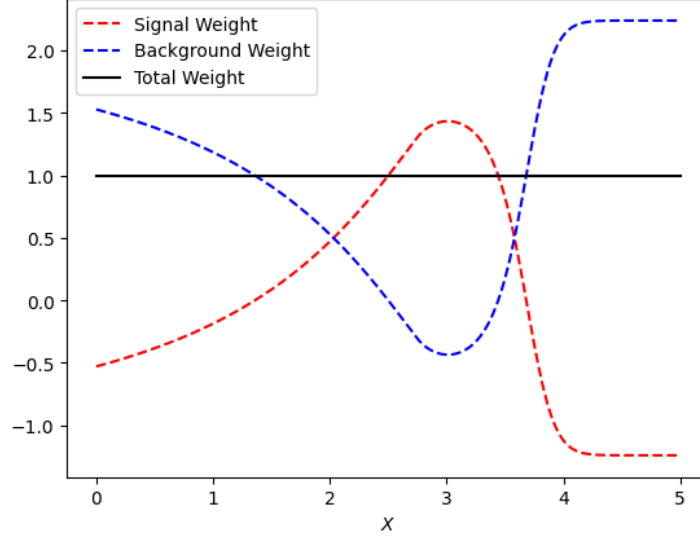


Figure 8: Signal, background of total weights for the first toy of 5000 sample size

These weights are applied to the Y variable data to project the Y signal and background distribution as shown in Figure 9. The projected signal and background curves resemble the predefined  $h_s(Y)$  and  $h_b(Y)$ , suggesting sWeights has projected out the Y signal successfully.

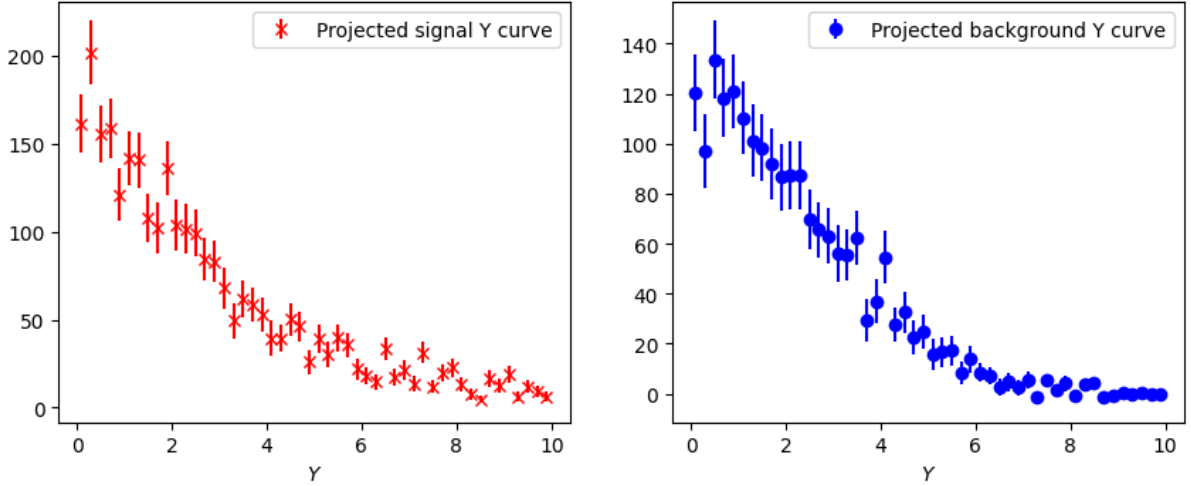


Figure 9: Projected Y signal and background distribution

After running sWeights through 250 toys for each sample size, EML was used to fit the projected Y signal to estimate the decay parameter  $\lambda$ . The bias of  $\lambda$  is calculated through subtracting the true  $\lambda$  parameter of 0.2996 from the estimated  $\lambda$ , and the mean  $\lambda$  is plotted for each sample size in Figure 10. The error bars are calculated by  $\sigma/\sqrt{N}$  to give the standard error on the mean. With increasing sample sizes, a consistent negative bias emerges, and the error on the bias decreases drastically. Generally, the bias on  $\lambda$  is small, and sWeights accurately estimated Y's decay constant  $\lambda$  value using projection of weights from the known X distribution.

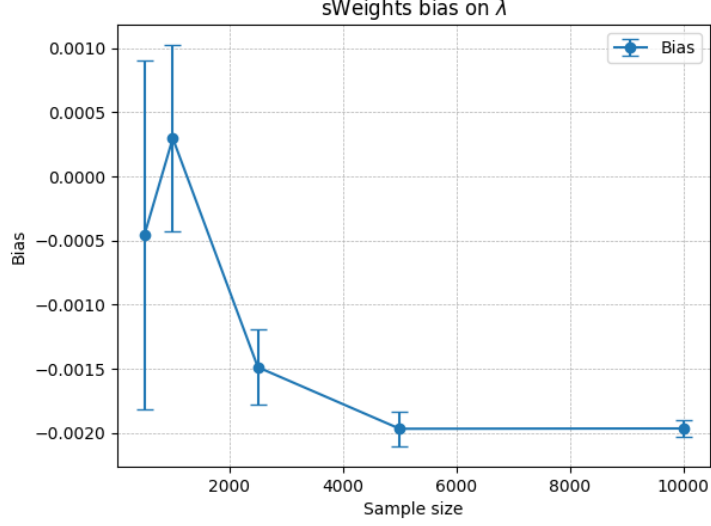


Figure 10: Mean lambda bias using sWeight for parameter estimation

Comparing to part E, where the bias and uncertainty of  $\lambda$  was obtained through EML fitting of samples collected from the joint  $X$  and  $Y$  distribution  $f(X, Y)$ , sWeights obtained very similar results as shown in Figure 11. In particular, a negative bias with a smaller standard error is observed tending to larger sample sizes. Both methods also show that uncertainty of the  $\lambda$  estimate decreases with sample sizes. The sWeights method obtained a lower uncertainty overall for all the sample sizes, implying more consistent estimation of the  $\lambda$  parameter. However, sWeights consistently underestimated  $\lambda$  to a greater extent than the method in part E.

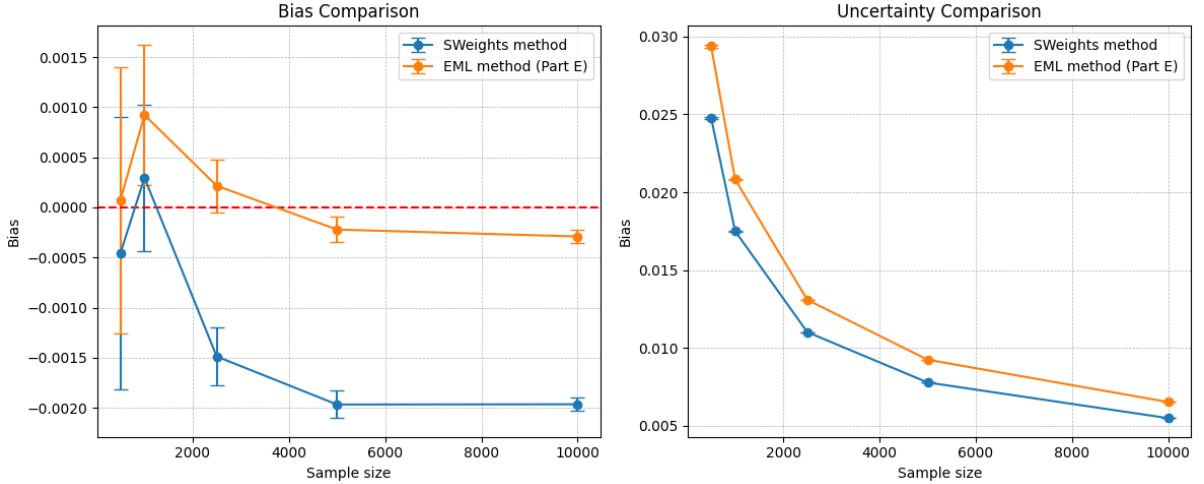


Figure 11: Bias and Uncertainty comparison between EML and sWeights

## 8 Part G: Evaluation of methods

Both methods consistently underestimated the  $\lambda$  parameter. The systematic bias observed is likely not due to insufficient sample size, as the bias is consistently negative with larger sample size. Studies have suggested this may be due to the nature of the truncated distributions defined in part B [2], such that the bias of the estimators was found to increase with higher degrees of truncation. Recall that the  $Y$  signal  $h_s(Y)$  is defined as a truncated exponential between  $Y \in (0, 10)$ . This degree of truncation is reflected in the likelihood function, which may have led to consistently smaller  $\lambda$  estimated, ultimately causing a systematic negative bias.

The main drawback of using sWeights is a consistently greater systematic bias that is not resolved with larger sample sizes. In other words, sWeights consistently underestimates the  $\lambda$  parameter more than the part E method. Though very few literature have investigated this issue, an educated guess would be that sWeights accounts for an additional bias in weights associated with the X distribution. For example, given that the X signal  $g_s(X)$  is a crystal ball function truncated between  $X \in (0,5)$ , the estimated parameters  $N_s$  and  $N_b$  may contain errors and biases, which are then carried forward into the weights, the projection of the Y signal  $h_s(Y)$ , and finally the  $\lambda$  estimation.

However, the method in part E, which entails using EML for parameters based on the known signal and background models of a joint distribution, comes with drawbacks too. In many applied sciences, data can be noisy, implying that the Y signal cannot be easily recovered from the mixture of signal and background information that makes up the Y data. For example, this is a common problem in particle physics, where signal properties like lifetime, spin and amplitudes are masked by the background contribution. In these situations where the Y background distribution is not clearly defined and a simple fitting of the data using EML cannot be performed, sWeights should be used. However, it should be noted that both signal and background should factorise for the valid application of sWeights. For this specific task, since the distributions are clearly defined as in part A, a simple EML should be used as it can better estimate  $\lambda$  values to a smaller degree of error.

## 9 Conclusion

In conclusion, both EML and sWeights are valid methods for parameter estimation. However, it is recommended that EML is used when both the X and Y variable signal and background fractions are known. sWeights should be used as an alternative when the Y signal cannot be easily recovered from a unknown mixture of background and signal.

## 10 Appendix

This report used generation tools such as ChatGPT and Claude.ai to facilitate LaTeX formatting, such as inserting figures, making tables and inserting equations. ChatGPT was used to help debugging code in the associated Jupyter notebook.

## References

- [1] Hans Dembinski. Introduction to iminuit, 2019. Accessed: 2024-06-17.
- [2] James T. Hattaway. Parameter estimation and hypothesis testing for the truncated normal distribution with applications to introductory statistics grades. Master's thesis, Brigham Young University, 2010. Accessed: 2024-12-17.
- [3] ScienceDirect Topics. Asymptotic normality, 2024. Accessed: 2024-12-18.

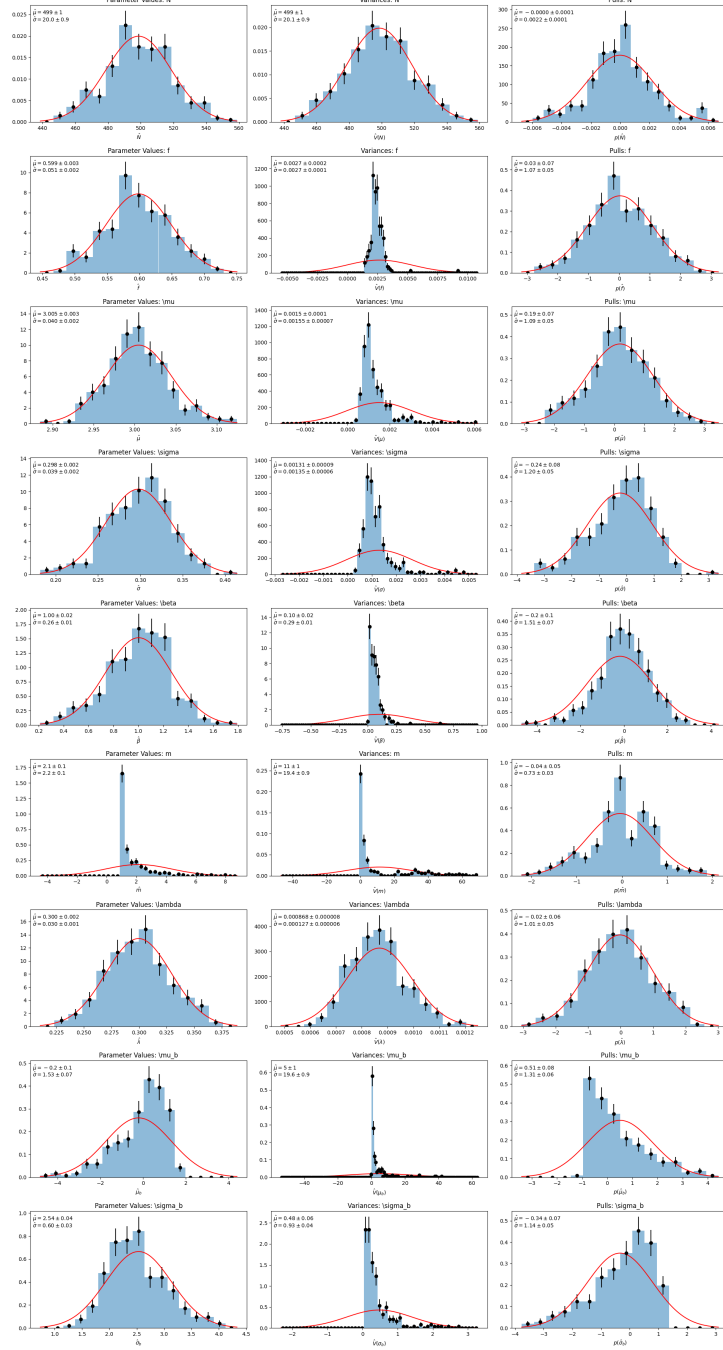


Figure 12: Histograms of 250 toys for a sample size of 500 showing the parameter value estimates, uncertainties and pull distribution.

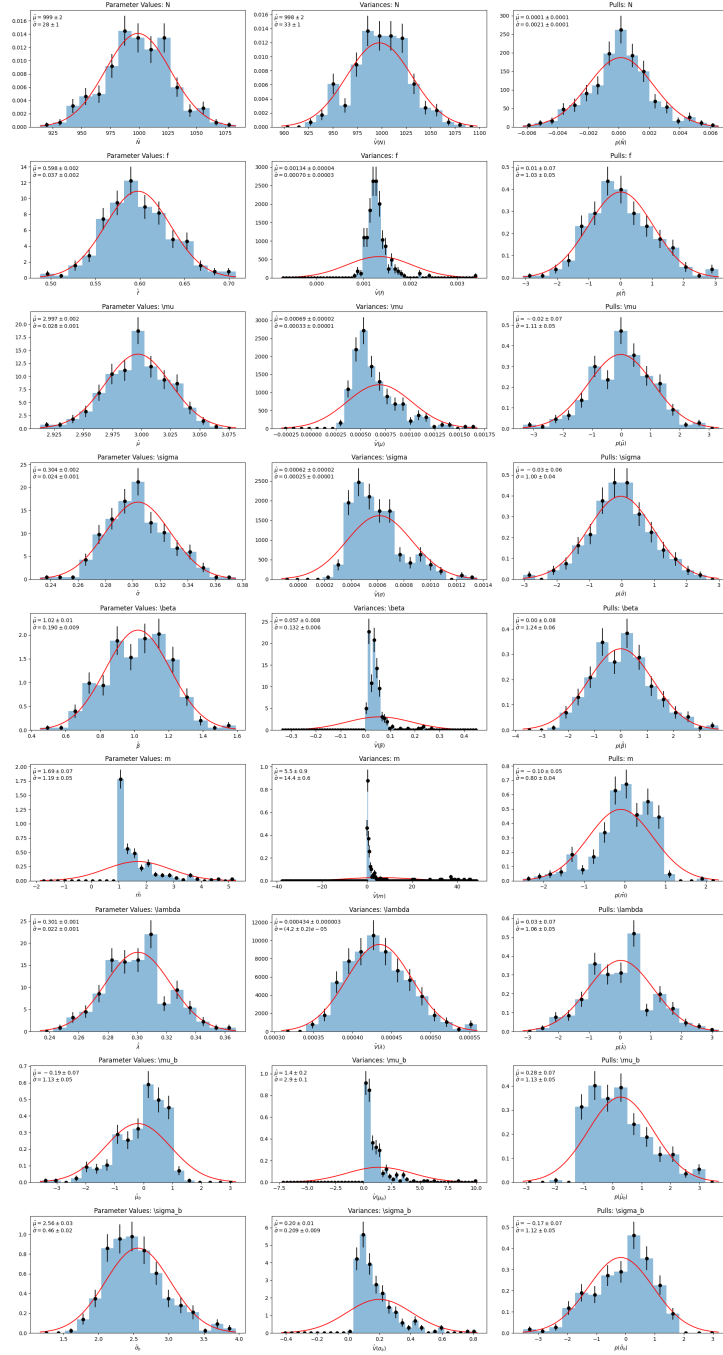


Figure 13: Histograms of 250 toys for a sample size of 1000 showing the parameter value estimates, uncertainties and pull distribution.

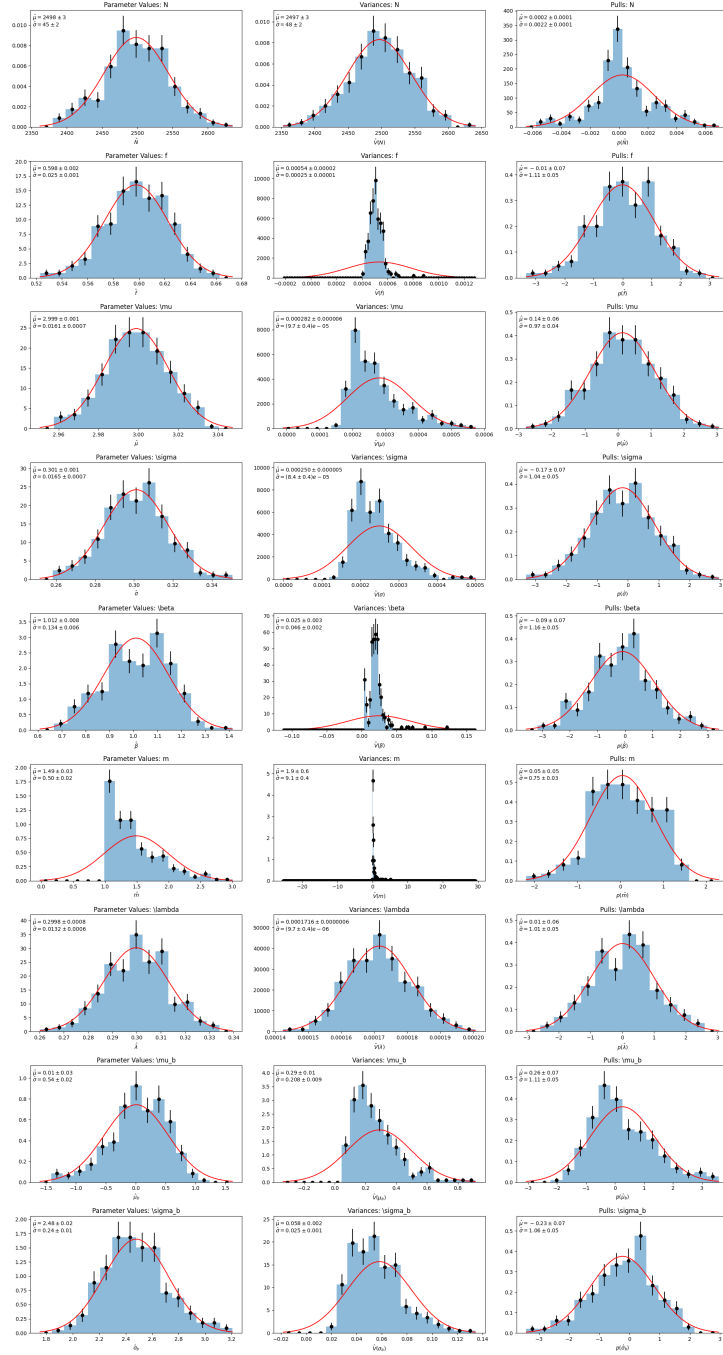


Figure 14: Histograms of 250 toys for a sample size of 2500 showing the parameter value estimates, uncertainties and pull distribution.

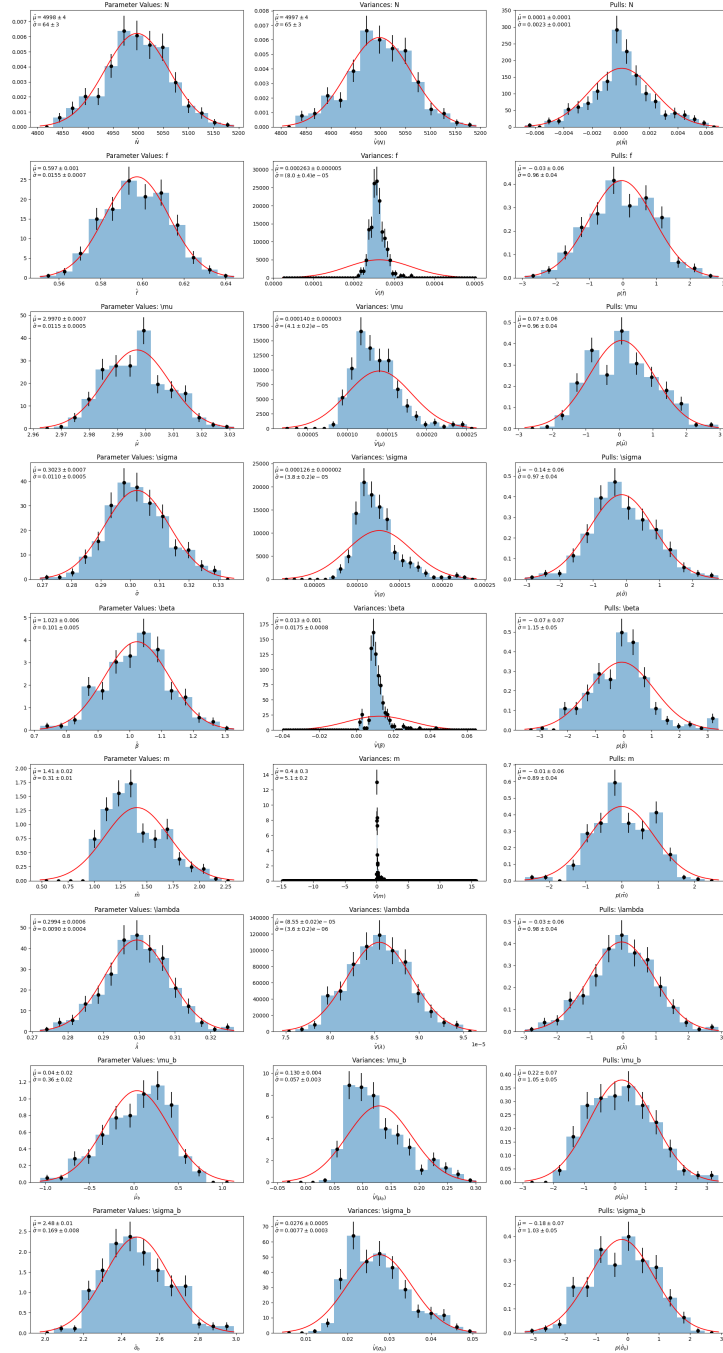


Figure 15: Histograms of 250 toys for a sample size of 5000 showing the parameter value estimates, uncertainties and pull distribution.



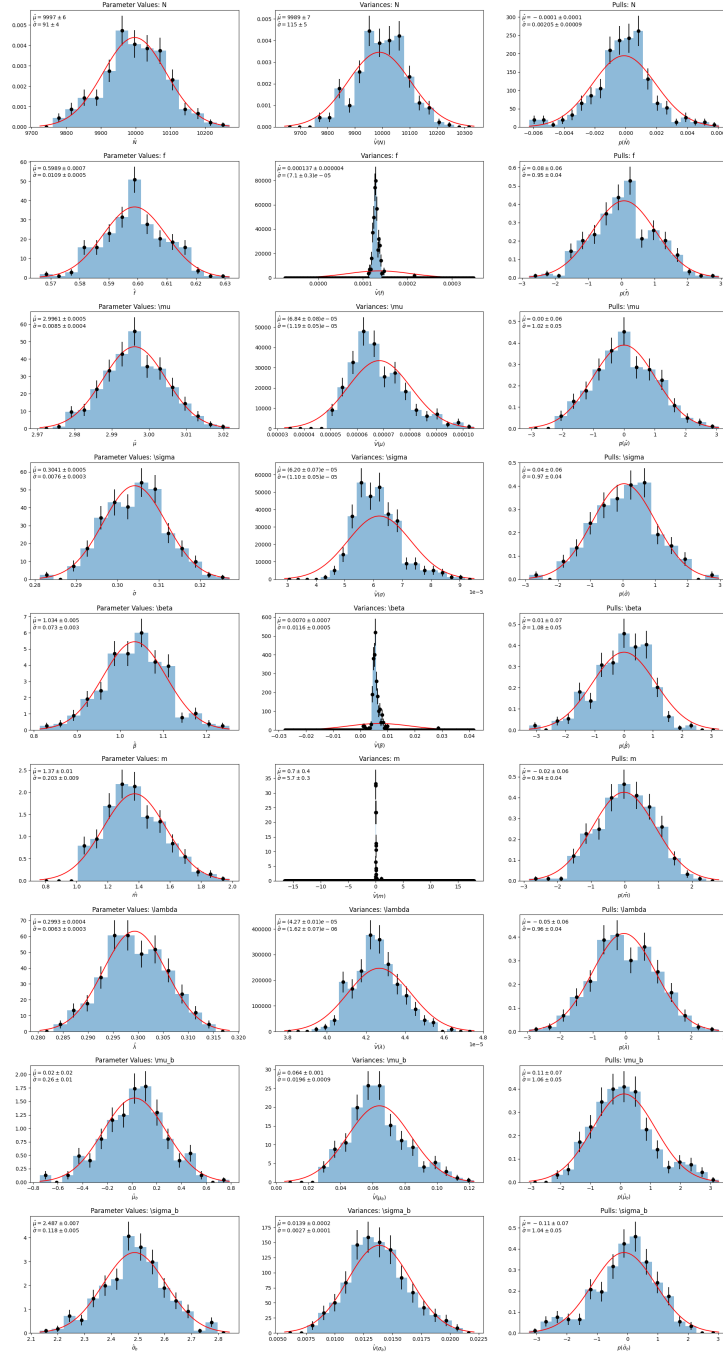


Figure 16: Histograms of 250 toys for a sample size of 10000 showing the parameter value estimates, uncertainties and pull distribution.