Faculty of Physics, Niels Bohr Institute
Advanced Methods In Applied Statistics 2017

sFit: a method for background subtraction in maximum likelihood fit

Summary presentation

Mads Juul Damgaard and Tue Holm-Jensen

March 9, 2017

Content

We have met 2 concepts during this course:

1. Likelihood fits
2. sPlots

Combined by

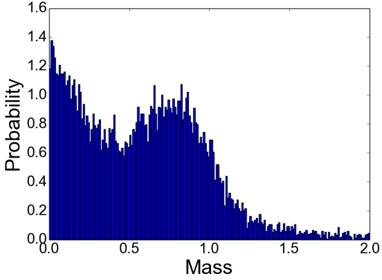
$$L(x;\theta) = \prod_{i}^{N} P_s(x_i;\theta)^{W_s(y_i)},$$

where $W_s(y_i)$ is the sWeight encountered earlier

Introduction
Motivating example

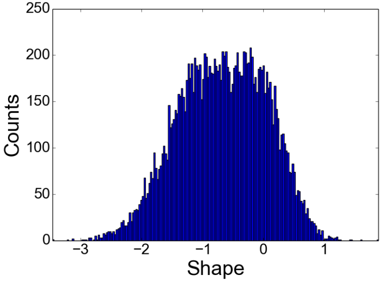
Mass:

1. Signal: Gaussian
2. Background: Exponential



Shape:

1. Signal: Gaussian
2. Background: ???

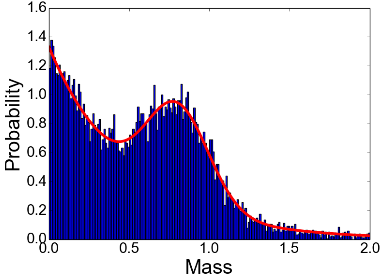


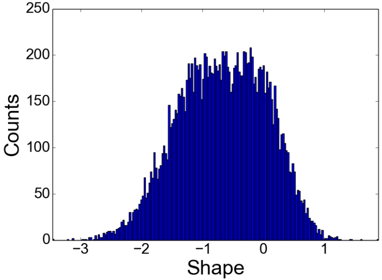
Mass:

1. Signal: Gaussian
2. Background: Exponential

Shape:

1. Signal: Gaussian
2. Background: ???



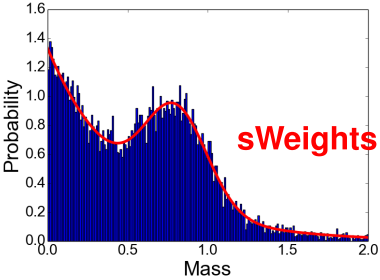


Mass:

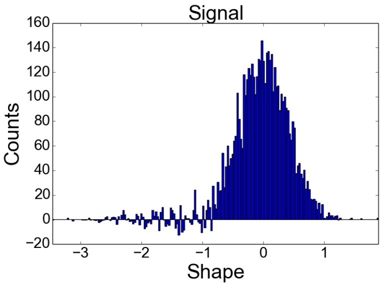
1. Signal: Gaussian
2. Background: Exponential

Shape:

1. Signal: Gaussian
2. Background: ???



sWeights ⇒



Introduction

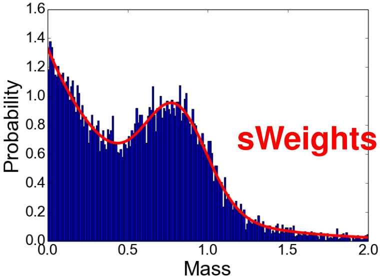
Motivating example

Mass:

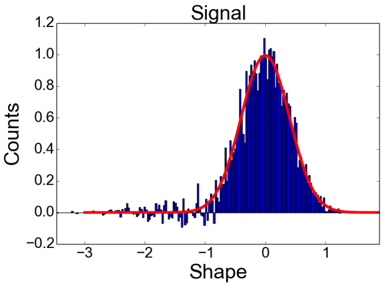
1. Signal: Gaussian
2. Background: Exponential

Shape:

1. Signal: Gaussian
2. Background: ???



sWeights ⇒

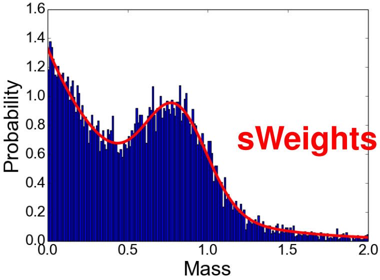


Mass:

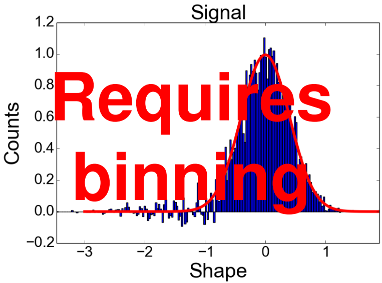
1. Signal: Gaussian
2. Background: Exponential

Shape:

1. Signal: Gaussian
2. Background: ???



sWeights ⇒



Likelihood function

$$L(x;\theta;f_s) = \prod_{i}^{N} \left[f_s\, P_s(x_i;\theta) + (1 - f_s)\, P_b(x_i;\theta) \right],$$

P_s signal PDF
P_b background PDF
f_s relative strength of signal

Likelihood function

$$L(x;\theta;f_s) = \prod_i^N \left[f_s\, P_s(x_i;\theta) + (1 - f_s)\, P_b(x_i;\theta) \right],$$

P_s signal PDF
P_b background PDF, *unknown*
f_s relative strength of signal

Likelihood function

$$L(x;\theta;f_s) = \prod_i^N \left[f_s\, P_s(x_i;\theta) + (1-f_s)\, P_b(x_i;\theta) \right],$$

P_s signal PDF
P_b background PDF, *unknown*
f_s relative strength of signal

Weighted likelihood

$$L(x;\theta) = \prod_i^N P_s(x_i;\theta)^{W_s(y_i)}$$

$w_s(y_i)$ sWeight encountered earlier

Monte Carlo parent distributions

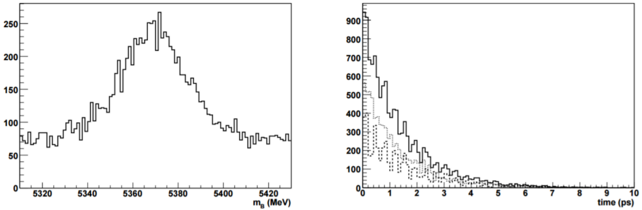


Figure 1: Distributions from a data set with $S_m/\sigma_m = 6$, $N_s = 5000$ and $N_b/N_s = 1.5$. Left: the B mass distribution; right: the total time distribution (solid), as well as the signal time distribution (dashed) and background time distribution (dot-dashed) reconstructed

Parameter bootstrapping

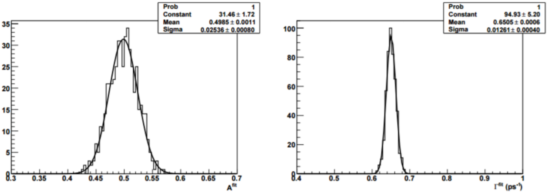


Figure 3: Distributions of the estimated values of A (left) and Γ (right) obtained with the sFit method, with superimposed gaussian fits, for the scenario $S_m/\sigma_m = 6$, $N_s = 5000$ and $N_b/N_s = 1.5$. The input values are $A = 0.5$ and $\Gamma = 0.65\,\mathrm{ps}^{-1}$.

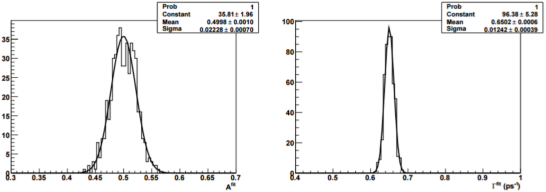


Figure 4: Distributions of the estimated values of A (left) and Γ (right) obtained with the reference method, with superimposed gaussian fits, for the scenario $S_m/\sigma_m = 6$, $N_s = 5000$ and $N_b/N_s = 1.5$. The input values are $A = 0.5$ and $\Gamma = 0.65\,\mathrm{ps}^{-1}$.

Parameter bootstrapping

S_m/σ_m, N_s, N_b/N_s	$\sigma(A)$	mean of A	$\sigma(\Gamma)$ (ps^{-1})	mean of Γ (ps^{-1})
4, 5000, 1	0.0304	0.502	0.0134	0.6504
6, 5000, 1.5	0.0254	0.498	0.0126	0.6504
4, 5000, 0.5	0.0243	0.501	0.0115	0.6511
6, 5000, 0.75	0.0223	0.501	0.0107	0.6496

Table 1: Statistical errors and mean values of A and Γ from 500 fits using the sFit method for different scenarios. Errors of the numbers are on the last digits. The input values are $A = 0.5$ and $\Gamma = 0.65\,\text{ps}^{-1}$.

S_m/σ_m, N_s, N_b/N_s	$\sigma(A)$	mean of A	$\sigma(\Gamma)$ (ps^{-1})	mean of Γ (ps^{-1})
4, 5000, 1	0.0251	0.502	0.0129	0.6506
6, 5000, 1.5	0.0223	0.500	0.0124	0.6502
4, 5000, 0.5	0.0215	0.500	0.0113	0.6511
6, 5000, 0.75	0.0211	0.501	0.0105	0.6496

Table 2: Statistical errors and mean values of A and Γ from 500 fits using the conventional maximum likelihood method for different scenarios. Errors of the numbers are on the last digits. The input values are $A = 0.5$ and $\Gamma = 0.65\,\text{ps}^{-1}$.

Thank you for your attention