

# Ćwiczenie 5

## ROOT – selekcja przypadków

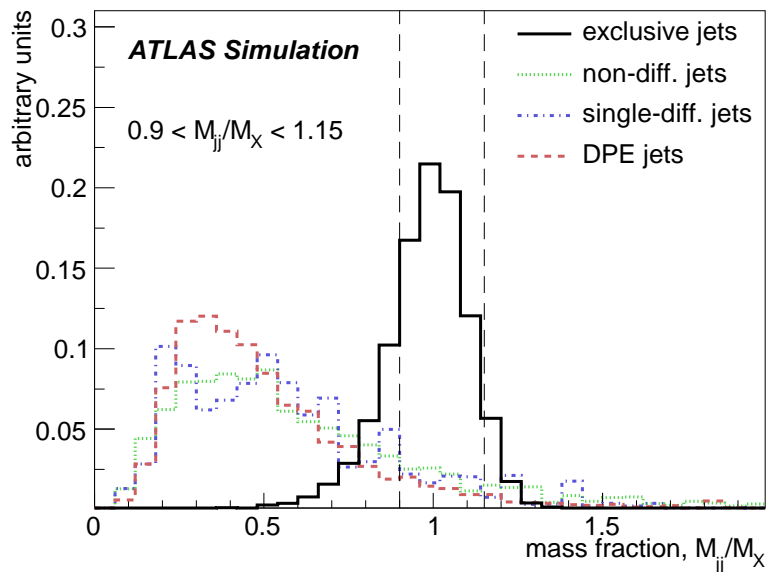
### 1. Wstęp teoretyczny

W eksperymentach w fizyce wysokich energii bardzo często jesteśmy zainteresowani badaniem rzadkich procesów, np. bozon Higgsa jest produkowany na LHC mniej więcej raz na 100 000 000 zderzeń proton-proton. W związku z tym istnieje bardzo duże prawdopodobieństwo, że oprócz interesującego nas sygnału zarejestrujemy przypadki tła, które go „udają”. Jednym z problemów podczas analizy danych jest więc oddzielenie sygnału od tła. Separacji tej dokonujemy na podstawie znanych (lub spodziewanych) różnic właściwości sygnału i tła. Różnice te są wnioskowane na podstawie analizy przewidywań teoretycznych i modelowane za pomocą generatorów Monte Carlo.

Dla przykładu rozważmy sytuację zobrazowaną na Rys. 1:

- wykres został narysowany w funkcji pewnej zmiennej ( $M_{jj}/M_X$ ), w której spodziewaliśmy się różnego zachowania sygnału oraz tła,
- sygnał zaznaczony jest ciągłą, czarną linią,
- tła narysowane są przerywanymi, kolorowymi liniami.

Gołym okiem widać, że rozkład sygnału jest inny niż rozkład tła. Ponadto, na podstawie analizy statystycznej można było określić optymalne cięcia ( $0.9 < M_{jj}/M_X < 1.15$ ).



Rysunek 1. Rozkład sygnału (ciągła, czarna linia) oraz tła (przerywane, kolorowe linie).

#### 1.1. Cele ćwiczenia

Celem ćwiczenia jest zapoznanie się z podstawowymi metodami selekcji przypadków. Przygotowane zostały następujące pliki:

- `data.root` – przypadki danych,
- `mc_signal.root` – przypadki Monte Carlo wygenerowane w oparciu o model sygnału (patrz pkt. 3),
- `mc_background.root` – przypadki Monte Carlo wygenerowane w oparciu o model tła (patrz pkt. 4).

Na podstawie analizy właściwości przypadków Monte Carlo należy zdefiniować odpowiednie cięcia, a następnie zastosować je do separacji sygnału i tła w danych.

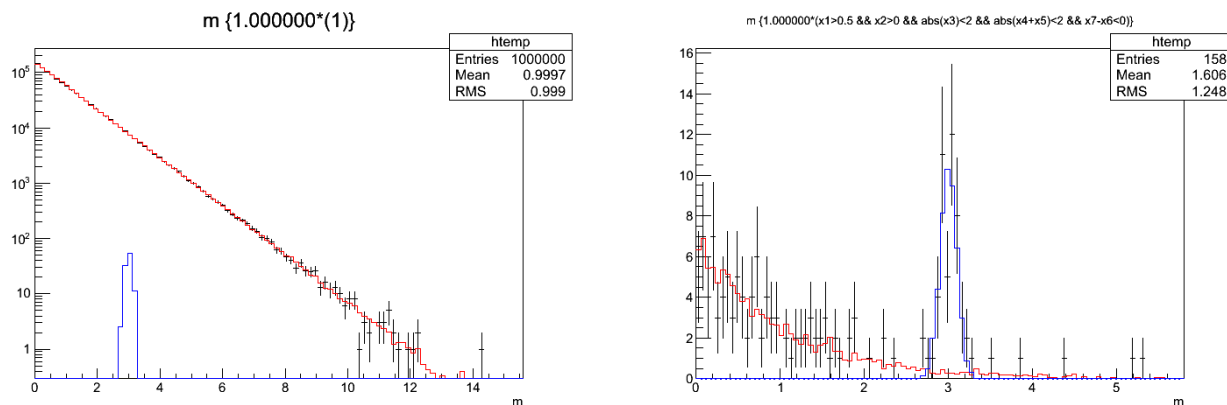
## 1.2. Struktura plików

Każdy przypadek zawiera informację o ośmiu zmiennych:

- $m$  – masa niezmiennicza (w jej rozkładzie szukamy rezonansu będącego naszym sygnałem),
- $\mathbf{x} = (x_1, x_2, \dots, x_7)$  – inne zmienne kinematyczne na których trzeba dokonać cięć. Ich interpretacja fizyczna jest nieistotna dla celów ćwiczenia.

## 2. Wykonanie ćwiczenia

1. Utworzyć program wykonujący pętlę po przypadkach wczytywanych z pliku podawanego jako pierwszy argument linii poleceń (`argc`, `argv`). Potrzebne pliki znajdują się w katalogu `/home/praktyki/cwiczenie5/` (nie trzeba ich kopiować, wystarczy odpowiednio podlinkować).
2. Program ma zapisywać tworzone histogramy do pliku typu `.root` o nazwie podawanej jako drugi argument linii poleceń.
3. Wykorzystując powyższy program utworzyć rozkłady zmiennej  $m$  dla wszystkich trzech próbek.
4. Napisać makro wczytujące histogramy z trzech plików i rysujące je na jednym wykresie. Rozkłady MC narysować liniami, rozkład dla danych punktami z słupkami błędów; znormalizować je zgodnie z poniższymi informacjami:
  - przekrój czynny dla próbki sygnału wynosi 1 nb,
  - przekrój czynny dla próbki tła wynosi 10000 nb,
  - scałkowana świetlność odpowiadająca zebranej próbce danych wynosi  $100 \text{ nb}^{-1}$ .
5. W oparciu o próbki MC opracować selekcję przypadków tak, by zredukować tło.
6. Dopasować funkcję Gaussa do rezonansu obserwowanego w danych po selekcji.



Rysunek 2. Przykładowy wynik wykonania ćwiczenia. Rysunek po lewej pokazuje początkowy rozkład masy niezmienniczej. Rysunek po prawej pokazuje ten sam rozkład dla przypadków przechodzących selekcję. Na obu rysunkach czarne punkty reprezentują dane, czerwona linia – przewidywania MC dla tła, niebieska linia – przewidywania MC dla sygnału.

### 3. Model Sygnału

Przypadki MC sygnału wygenerowane zostały z rozkładu:

$$f^S(m, \mathbf{x}) = f_0^S(m) \cdot f_1^S(x_1) \cdot f_2^S(x_2) \cdot f_3^S(x_3) \cdot f_{45}^S(x_4, x_5) \cdot f_{67}^S(x_6, x_7)$$

gdzie funkcje  $f_i$  zdefiniowane są wzorami:

$$\begin{aligned} f_0^S(m) &= \frac{1}{\sqrt{2\pi c_0}} \exp\left(-\frac{(m-d_0)^2}{2c_0^2}\right), \\ f_1^S(x_1) &= \frac{1}{\sqrt{2\pi c_1}} \exp\left(-\frac{(x_1-d_1)^2}{2c_1^2}\right), \\ f_2^S(x_2) &= \frac{1}{\sqrt{2\pi c_2}} \exp\left(-\frac{(x_2-d_2)^2}{2c_2^2}\right), \\ f_3^S(x_3) &= \frac{1}{\sqrt{2\pi c_3}} \exp\left(-\frac{x_3^2}{2c_3^2}\right), \\ f_{45}^S(x_4, x_5) &= \frac{1}{\sqrt{2\pi c_4}} \exp\left(-\frac{(x_4+x_5)^2}{2c_4^2}\right) \cdot \frac{1}{\sqrt{2\pi c_5}} \exp\left(-\frac{(x_4-x_5)^2}{2c_5^2}\right), \\ f_{67}^S(x_6, x_7) &= \frac{1}{\sqrt{2\pi c_6}} \exp\left(-\frac{(x_6+x_7)^2}{2c_6^2}\right) \cdot \frac{1}{\sqrt{2\pi c_7}} \exp\left(-\frac{(x_6-x_7-d_7)^2}{2c_7^2}\right). \end{aligned}$$

a parametry przyjmują wartości:

$$d_0 = 3.0, d_1 = 1.0, d_2 = 1.0, d_7 = 0.5,$$

$$c_0 = 0.1, c_1 = 0.7, c_2 = 2.0, c_3 = 1.0, c_4 = 1.0, c_5 = 2.0, c_6 = 1.0, c_7 = 0.3.$$

### 4. Model Tła

Przypadki MC tła wygenerowane zostały z rozkładu:

$$f^B(m, \mathbf{x}) = f_0^B(m) \cdot f_1^B(x_1) \cdot f_2^B(x_2) \cdot f_3^B(x_3) \cdot f_{45}^B(x_4, x_5) \cdot f_{67}^B(x_6, x_7)$$

gdzie funkcje  $f_i$  zdefiniowane są wzorami:

$$\begin{aligned} f_0^B(m) &= \lambda e^{-\lambda m}, \\ f_1^B(x_1) &= \frac{1}{\sqrt{2\pi a_1}} \exp\left(-\frac{(x_1-b_1)^2}{a_1^2}\right), \\ f_2^B(x_2) &= \frac{1}{\sqrt{2\pi a_2}} \exp\left(-\frac{(x_2-b_2)^2}{a_2^2}\right), \\ f_3^B(x_3) &= \frac{1}{\sqrt{2\pi a_3}} \exp\left(-\frac{(x_3)^2}{a_3^2}\right), \\ f_{45}^B(x_4, x_5) &= \frac{1}{\sqrt{2\pi a_4}} \exp\left(-\frac{(x_4+x_5)^2}{2b_4^2}\right) \cdot \frac{1}{\sqrt{2\pi a_5}} \exp\left(-\frac{(x_4-x_5)^2}{2a_5^2}\right), \\ f_{67}^B(x_6, x_7) &= \frac{1}{\sqrt{2\pi a_6}} \exp\left(-\frac{(x_6+x_7-b_6)^2}{2a_6^2}\right) \cdot \frac{1}{\sqrt{2\pi a_7}} \exp\left(-\frac{(x_6-x_7-b_7)^2}{2a_7^2}\right), \end{aligned}$$

a parametry  $a_i$  i  $b_i$  przyjmują wartości:

$$\lambda = 1.0, b_1 = -1.0, b_2 = -1.0, b_3 = 0.0, b_4 = 0.0, b_5 = 0.0, b_6 = 0.0, b_7 = -0.5,$$

$$a_1 = 0.7, a_2 = 1.0, a_3 = 5.0, a_4 = 3.0, a_5 = 1.0, a_6 = 1.0, a_7 = 1.0.$$