

Ćwiczenie 3

ROOT – selekcja przypadków

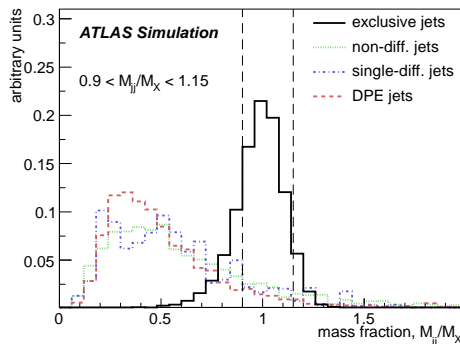
1. Wstęp teoretyczny

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 (proces ekskluzywnej produkcji jetów) zaznaczony jest ciągłą, czarną linią,
- tła narysowane są przerywanymi, kolorowymi liniami.

Gołym okiem widać, że rozkład sygnału jest inny niż rozkład tła. Można więc wybrać taki zakres zmiennej M_{jj}/M_X , w której jest znacznie więcej sygnału niż tła. Ponadto, na podstawie analizy statystycznej można było określić optymalne cięcia (tu $0.9 < M_{jj}/M_X < 1.15$).



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. Zakładamy, że poszukujemy pewnej cząstki (takiej jak np. bozon Z czy Higgs). Wiemy, że masa takiej cząstki jest dobrze określona – na wykresie masy niezmienniczej (m) cząstka taka będzie widoczna jako pik (rezonans). O tle wiemy, że posiada bardziej płaski w zmiennej m . Dodatkowo, sygnał oraz tło różnią się rozkładami pewnych charakterystycznych zmiennych np. krotności (ilości) produkowanych cząstek, korelacji pomiędzy kątami pod jakimi są produkowane itp. W tym ćwiczeniu sygnałem jest hipotetyczny rezonans R , który rozpada się na parę mionów o przeciwnych ładunkach $R \rightarrow \mu^+ \mu^-$. Tłem są przypadki, w pojawia się losowa para mionów. Przygotowane zostały następujące pliki (można je znaleźć w na stronie praktyk):

- przypadki danych: `data.root`,
- przypadki Monte Carlo wygenerowane w oparciu o model sygnału: `MC_resonance_search_signal.root`,
- przypadki Monte Carlo wygenerowane w oparciu o model tła:
 - `MC_resonance_search_background_small.root` – mała próbka do wstępnej analizy,
 - `MC_resonance_search_background_X.root`, gdzie $X = 0, 1, 2, 3$ lub 4 – wszystkie wygenerowane przypadki tła.

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

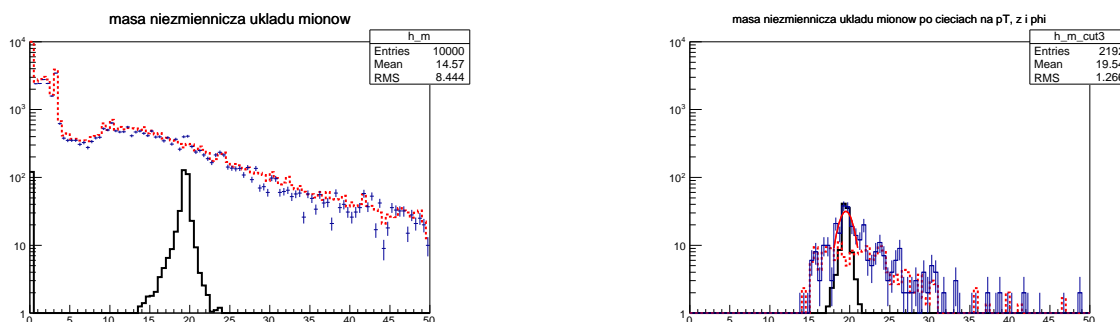
1.2. Struktura plików

Każdy przypadek zawiera informację o dziesięciu zmiennych opisujących każdą ze zmierzonych cząstek:

- `vector<double> *x`; – miejsce produkcji cząstki w x zrekonstruowane przez detektor,
- `vector<double> *y`; – jw, ale w y ,
- `vector<double> *z`; – jw, ale w z ,
- `vector<double> *px`; – pęd cząstki w kierunku x ,
- `vector<double> *py`; – jw, ale w y ,
- `vector<double> *pz`; – jw, ale w z ,
- `vector<double> *e`; – energia cząstki,
- `vector<int> *id`; – numer identyfikacyjny cząstki, dla mionów $id_{\mu^-} = 13$, $id_{\mu^+} = -13$; o konwencji nadawania numerów można przeczytać więcej np. tutaj pdg.lbl.gov/2007/reviews/montecarlo/rpp.pdf,
- `vector<double> *charge`; – ładunek cząstki,
- `vector<double> *mu_like`; – zmienna określająca prawdopodobieństwo, że dana cząstka jest mionem.

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 na stronie praktyk. Za odpowiedź może służyć program wykonany w ramach ćw. 2. Warto również użyć polecenia `MakeClass()`.
2. Program ma zapisywać tworzone histogramy do pliku typu `.root` o nazwie podawanej jako drugi argument linii poleceń.
3. Ustalić selekcję mionów na podstawie zmiennej `mu_like`. W tym celu należy, używając próbek MC, porównać rozkłady zmiennej `mu_like` dla cząstek z `id` mionu oraz wszystkich innych `id`. Wykresy zapisać do osobnego pliku z histogramami. Napisać program do odczytywania wykresów i porównać rozkłady. Wykresy powinny zostać znormalizowane do 1.
4. Wykorzystując powyższy program utworzyć rozkłady masy niezmienniczej dla wszystkich trzech próbek.
5. 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 5 nb,
 - przekrój czynny dla próbki tła wynosi 500 nb,
 - scałkowana świetlność odpowiadająca zebranej próbce danych wynosi 100 nb^{-1} .
6. Czy sygnał jest dobrze widoczny w danych?
7. W oparciu o próbki MC opracować selekcję przypadków tak, by zredukować tło. W tym celu należy porównać ze sobą rozkłady odpowiednich zmiennych (znormalizowane do 1) i wprowadzić cięcia.
8. Z doświadczenia (i fizyki!) wiadomo, że pewne zmienne powinny separować sygnał od tła lepiej niż inne. Selekcja takiego rezonansu będzie prawdopodobnie bardzo podobna do ćwiczenia z poszukiwania bozonu Z^0 : odległość pomiędzy parą mionów w z , pęd poprzeczny mionów (zwłaszcza korelacje), kąt pomiędzy mionami. Można również sprawdzić czy istnieją korelacje między zmiennymi.
9. 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 niebieskie punkty reprezentują dane, czerwona linia – przewidywania MC dla tła, czarna linia – przewidywania MC dla sygnału.