

Pakiet ROOT

selekcja przypadków

Maciej Trzebiński

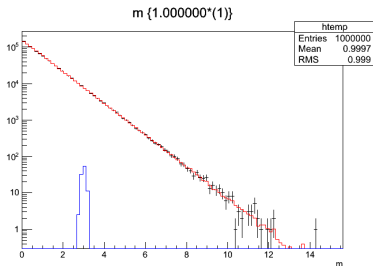
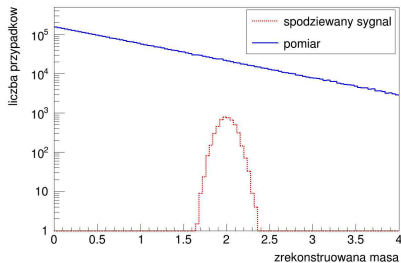
Instytut Fizyki Jądrowej
Polskiej Akademii Nauki



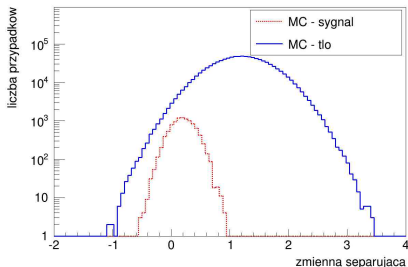
Praktyki studenckie na LHC
IFJ PAN, 6 lipca 2017

Konieczność selekcji

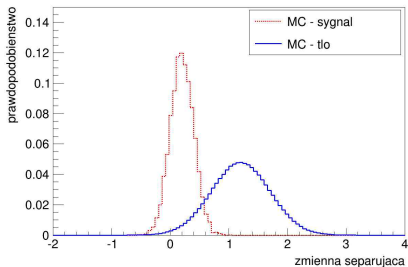
- Cząstki, które produkujemy w wyniku zderzeń są zwykle krótkożyciowe. Np. czas życia bozonu Higgsa to ok. 10^{22} s.
- W detektorach nie obserwujemy samych cząstek, tylko produkty ich rozpadu. Np. jeden ze sposobów rozpadu bozonu Higgsa, to produkcja dwóch fotonów.
- Gdy jednak weźmiemy wszystkie przypadki z dwoma fotonami zarejestrowane w detektorze, to okaże się, że jest ich zdecydowanie więcej, niż spodziewanego sygnału.
- W szczególności może powstać sytuacja jak na poniższym rysunku:
 - spodziany sygnał jest znacznie mniejszy niż tło,
 - nie widać go w danych – fluktuacje statystyczne ($\sqrt{\text{liczba przypadków}}$) są znacznie większe.



- Oprócz informacji o zrekonstruowanej masie, mamy przeważnie wiele innych informacji o właściwościach przypadku.
- Używając modeli Monte Carlo możemy wygenerować próbki o odpowiednich właściwościach. Następnie można bezpośrednio porównać sygnał i tło.
- Ponieważ ilość danych w próbkach może się znacznie różnić (zwykle generujemy znacznie więcej przypadków tła niż sygnału), to dobrze porównywać rozkłady znormalizowane do prawdopodobieństwa.

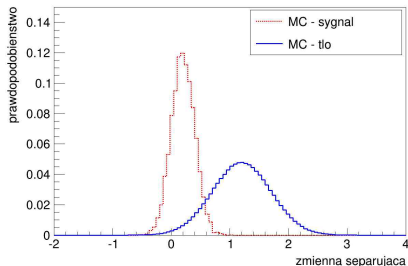


Rozkład sygnału i tła znormalizowany do ilości wygenerowanych przypadków.

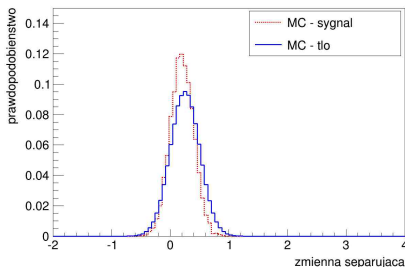


Rozkład sygnału i tła znormalizowany do prawdopodobieństwa (jedynki).

- Istnieją różne metody optymalizacji cięć – jest to temat szeroki i wykraczający poza ćwiczenie.
- My posłużymy się najprostszą metodą:
 - znormalizować oba rozkłady do prawdopodobieństwa,
 - zobaczyć, w który miejscu się przecinają,
 - ustawić wartość cięcia ma miejscu przecięcia, pod warunkiem, że:
 - cięcie usunie znacznie więcej tła niż sygnału ORAZ
 - po cięciu pozostanie co najmniej 50% sygnału.

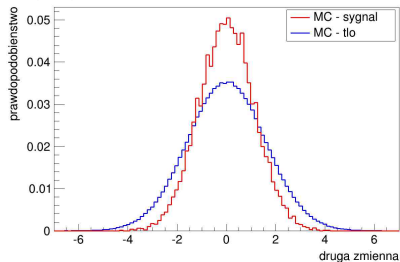
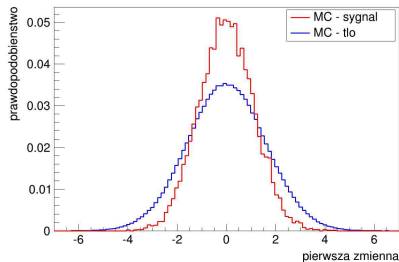


Przykład zmiennej, która może zostać użyta do separacji. Wymaganie wartości mniejszych od 0.6 odrzuci sporo tła i zostawi większość sygnału.

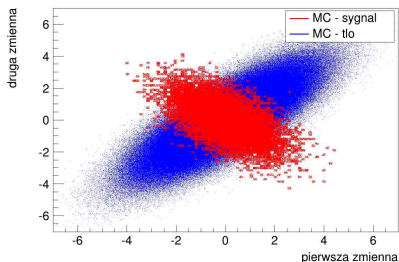


Przykład zmiennej, która nie nadaje się do separacji.

- Czasem pojedyncze zmienne nie nadają się do separacji:



- ale ich korelacja może prowadzić do efektywnego cięcia:



Po prawidłowym wykonaniu selekcji w ćwiczeniu, sygnał rezonansu powinien być wyraźnie widoczny w danych:

