

Exercícios de estatística para análise de dados em HEP

Professores: Eliza Melo, Dilson Damião e Mauricio Thiel *Nome:* Jorge Júlio Barreiros Venuto de Siqueira

Leitura dos arquivos

A primeira tarefa que foi pedida foi usar os dados do CMS opendata distribuídas anteriormente, utilizei o Python com auxílio do ROOT para ler os arquivos e selecionar as informações relativísticas das partículas.

Lendo o arquivo

```
diretorios = [
    "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOB/UL2016_MiniAOBv2_
    NanoAOBv9-v1/100000/*.root",
    "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOB/UL2016_MiniAOBv2_
    NanoAOBv9-v1/1010000/*.root",
    "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOB/UL2016_MiniAOBv2_
    NanoAOBv9-v1/250000/*.root"
]

# Expandir caminhos com glob e carregar arquivos
arquivos = glob.glob(diretorios[0]) + glob.glob(diretorios[1]) + glob.glob(diretorios[2])
#print(arquivos)

file_1 = uproot.open(arquivos[0])
file_2 = uproot.open(arquivos[1])
file_3 = uproot.open(arquivos[2])

nElectron_1 = tree_1["nElectron"].array()
eletron_pt_1 = tree_1["Electron_pt"].array()
eletron_eta_1 = tree_1["Electron_eta"].array()
eletron_phi_1 = tree_1["Electron_phi"].array()
eletron_mass_1 = tree_1["Electron_mass"].array()
eletron_charge_1 = tree_1["Electron_charge"].array()

nElectron_2 = tree_2["nElectron"].array()
eletron_pt_2 = tree_2["Electron_pt"].array()
eletron_eta_2 = tree_2["Electron_eta"].array()
eletron_phi_2 = tree_2["Electron_phi"].array()
eletron_mass_2 = tree_2["Electron_mass"].array()
eletron_charge_2 = tree_2["Electron_charge"].array()

nElectron_3 = tree_3["nElectron"].array()
eletron_pt_3 = tree_3["Electron_pt"].array()
eletron_eta_3 = tree_3["Electron_eta"].array()
eletron_phi_3 = tree_3["Electron_phi"].array()
eletron_mass_3 = tree_3["Electron_mass"].array()
eletron_charge_3 = tree_3["Electron_charge"].array()

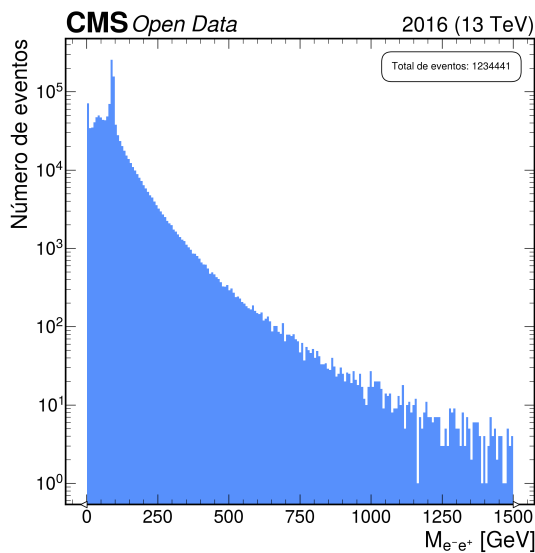
# Combinar os arrays
Neletron = np.concatenate([nElectron_1, nElectron_2, nElectron_3])
eletron_pt = np.concatenate([eletron_pt_1, eletron_pt_2, eletron_pt_3])
eletron_eta = np.concatenate([eletron_eta_1, eletron_eta_2, eletron_eta_3])
eletron_phi = np.concatenate([eletron_phi_1, eletron_phi_2, eletron_phi_3])
eletron_mass = np.concatenate([eletron_mass_1, eletron_mass_2, eletron_mass_3])
eletron_charge = np.concatenate([eletron_charge_1, eletron_charge_2, eletron_charge_3])
```

Para a análise dos dados, foram adotadas as seguintes estratégias de seleção e cortes nos eventos:

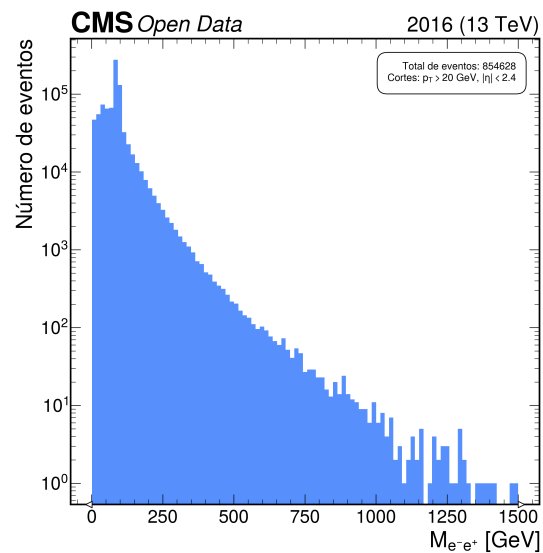
- **Seleção de eventos com dois elétrons de cargas opostas:** Apenas eventos contendo dois elétrons de cargas opostas foram considerados, garantindo que os pares de partículas relevantes fossem analisados.
- **Corte no momento transversal (P_T):** Foi aplicado um corte no momento transversal (P_T) de 20 GeV. Esse corte visa excluir partículas com baixo P_T , que frequentemente não contribuem de maneira significativa para os processos físicos de interesse ou são difíceis de detectar. Dessa forma, é possível reduzir o fundo de eventos irrelevantes, concentrando a análise em partículas de maior energia, mais associadas a eventos físicos relevantes.
- **Corte na pseudo-rapidez (η):** Foi aplicado um corte na pseudo-rapidez dos elétrons, restringindo o intervalo de η entre -2.4 e 2.4. Esse intervalo corresponde à região de aceitação do detector, onde a eficiência de detecção é máxima. Partículas fora dessa faixa têm menor probabilidade de serem detectadas corretamente, o que pode comprometer a precisão da análise. A exclusão dessas partículas melhora a eficiência da análise, garantindo que apenas eventos com partículas detectáveis sejam considerados.

Esses cortes são fundamentais para refinar a seleção dos dados, eliminando eventos irrelevantes e garantindo que a análise se concentre em regiões do espaço de fase com maior relevância física e precisão para realizar o cálculo da massa invariante dos di-elétrons.

Gráfico da massa invariante dos elétrons



(a) Distribuição da Massa invariante dos di-elétrons sem os cortes



(b) Distribuição da Massa invariante dos di-elétrons com os cortes de P_T e η

Fica evidente a queda no número de eventos e uma melhora aparente na região de sinal em torno de 90 GeV após aplicar os cortes de P_T e η .

A fim de estudar essa ressonância, será dada uma atenção especial a essa região, focando uma análise mais detalhada.

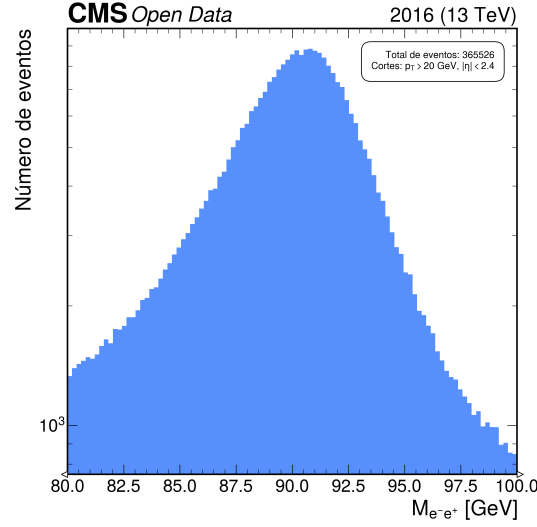
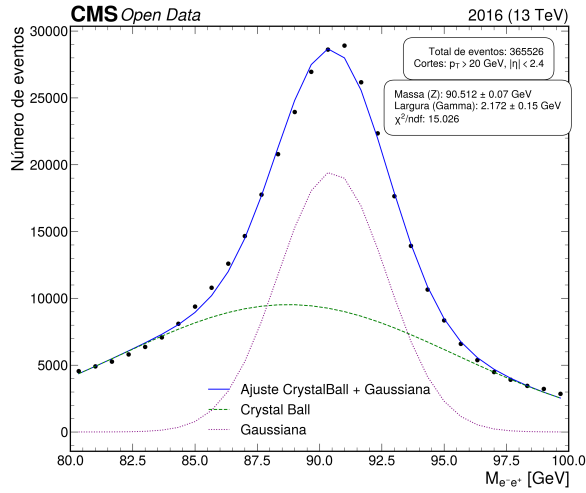
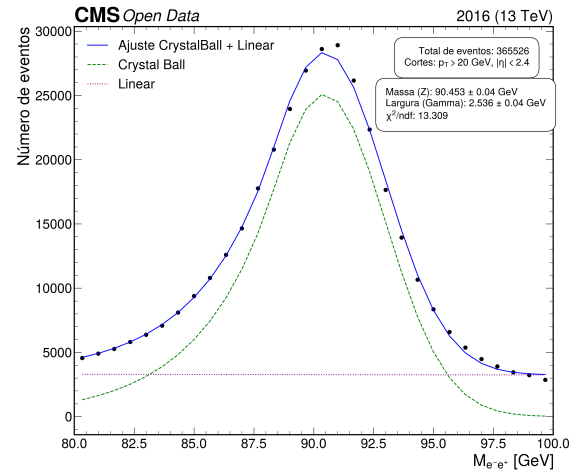


Figura 2: Massa invariante dos di-elétrons em torno de 90GeV

Com o objetivo de estudar essa ressonância, foram realizados diferentes ajustes de funções sobre os dados, a fim de determinar seus parâmetros físicos.



(a) CrystalBall + Gaussiana



(b) CrystalBall + Polinômio de grau 1

Com os ajustes realizados, observou-se que as funções ajustadas não conseguem representar de maneira completa a distribuição da região analisada. Os valores obtidos para χ^2/ndf para os ajustes (a) e (b) são 15,026 e 13,309, respectivamente, ambos significativamente superiores ao valor esperado de 1. Contudo, os valores determinados para o centro da ressonância e sua largura estão em boa concordância com o valor da massa invariante do bóson Z , conforme ilustrado na figura. Segundo o Particle Data Group (PDG), a massa do bóson Z é dada por $M_Z = 91,1880 \pm 0,0020 \text{ GeV}$ e sua largura de decaimento é $\Gamma_Z = 2,4955 \pm 0,0023 \text{ GeV}$. A fim de avaliar a compatibilidade entre os parâmetros determinados e os valores de referência fornecidos pelo PDG, será realizada uma análise de compatibilidade entre os parâmetros ajustados e os valores de referência.

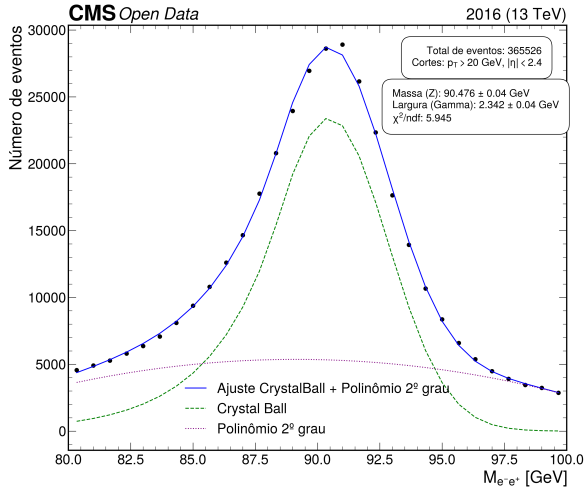
$$C_a^M = \frac{|M_Z^{\text{PDG}} - M_Z^a|}{\sqrt{\sigma_{\text{PDG}}^2 + \sigma_a^2}} = 9,65\sigma \quad C_b^M = \frac{|M_Z^{\text{PDG}} - M_Z^b|}{\sqrt{\sigma_{\text{PDG}}^2 + \sigma_b^2}} = 18,35\sigma$$

$$C_a^\Gamma = \frac{|\Gamma_Z^{\text{PDG}} - \Gamma_Z^a|}{\sqrt{\sigma_{\text{PDG}}^2 + \sigma_a^2}} = 2,15\sigma \quad C_b^\Gamma = \frac{|\Gamma_Z^{\text{PDG}} - \Gamma_Z^b|}{\sqrt{\sigma_{\text{PDG}}^2 + \sigma_b^2}} = 1,01\sigma$$

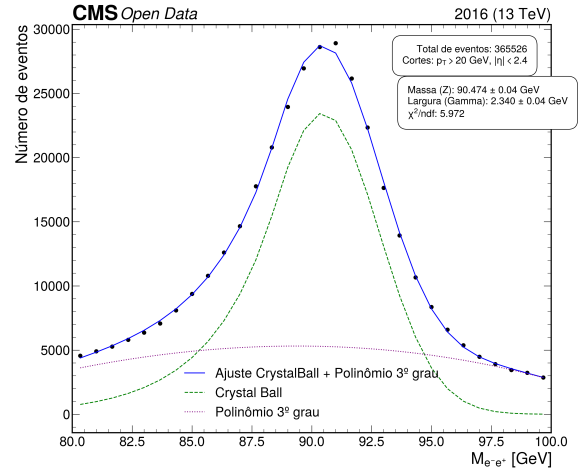
Compatibilidade	Valor
C_a^M	9,65 σ
C_a^Γ	2,15 σ
C_b^M	18,35 σ
C_b^Γ	1,01 σ

Observa-se que os valores dos parâmetros obtidos pelos ajustes realizados em (a) não são compatíveis com os valores de referência do PDG, uma vez que a análise de compatibilidade revelou um desvio superior a 3σ para a massa invariante, e um desvio próximo de 2σ para a largura de decaimento Γ , o que sugere inconclusividade. Por outro lado, os parâmetros do ajuste (b) apresentaram resultados mais distintos: enquanto a massa invariante se mostrou incompatível, com um erro associado superior a 3σ , a largura parcial de decaimento exibiu uma compatibilidade aceitável, com o desvio resultante sendo inferior a 2σ . Com base nesses resultados, conclui-se que os ajustes realizados não oferecem resultados satisfatórios.

Afim de buscar mais resultados, foi realizado ajuste de novas funções sobre os dados, como nas figuras abaixo.



(c) CrystalBall + Polinômio de grau 2



(d) CrystalBall + Polinômio de grau 3

Novamente, observa-se que o valor de χ^2/ndf em ambos os ajustes não é próximo de 1, porém, apresentam resultados melhores em comparação aos casos anteriores de (a) e (b). Além disso, é possível concluir que os resultados de (c) e (d) estão muito próximos entre si. A fim de determinar o melhor ajuste entre eles, será realizado novamente o cálculo da compatibilidade entre os parâmetros determinados e os valores de referência do PDG.

Compatibilidade	Valor
C_c^M	$17,78\sigma$
C_c^Γ	$3,83\sigma$
C_d^M	$17,83\sigma$
C_d^Γ	$3,88\sigma$

Os resultados das análises de compatibilidade ainda indicam uma incompatibilidade entre os parâmetros do PDG e os valores determinados. No entanto, observa-se também que, ao alterar o grau do polinômio, os valores obtidos não apresentaram uma grande variação, o que sugere a necessidade de uma investigação mais aprofundada sobre o conjunto de dados ou os ajustes realizados.

A figura a seguir apresenta os parâmetros determinados, juntamente com suas incertezas associadas, para os ajustes das figuras (a), (b), (c) e (d), respectivamente.

```
*****
Minimizer is Minuit2 / Migrad
Chi2      =      330.563
NDF       =         22
Edm       = 7.33343e-07
NCalls    =      1217
p0        = 9530.54 +/- 144.023
p1        = 88.7357 +/- 0.0395081
p2        = 6.71738 +/- 0.0723128
p3        = 408054 +/- 0.25
p4        = 1.05895 +/- 2.23607
p5        = 19472.1 +/- 142.257
p6        = 90.5123 +/- 0.0103987
p7        = 2.17156 +/- 0.0158591
```

(a) CrystalBall + Gaussiana

```
*****
Minimizer is Minuit2 / Migrad
Chi2      =      306.106
NDF       =         23
Edm       = 1.89156e-07
NCalls    =      1266
p0        = 25094.7 +/- 82.1739
p1        = 90.4526 +/- 0.0101598
p2        = 2.53634 +/- 0.0112235
p3        = 0.824661 +/- 0.0124043
p4        = 4.49378e+06 +/- 259053
p5        = 3507.5 +/- 455.169
p6        = -2.69059 +/- 4.69932
```

(b) CrystalBall + Polinômio de grau 1

```
*****
Minimizer is Minuit2 / Migrad
Chi2      =      130.791
NDF       =         22
Edm       = 3.73691e-06
NCalls    =      1217
p0        = 23422.7 +/- 82.754
p1        = 90.4761 +/- 0.0106632
p2        = 2.34208 +/- 0.0113115
p3        = 0.891706 +/- 0.0120504
p4        = 1.51941e+06 +/- 224350
p5        = -172325 +/- 163.182
p6        = 3987.83 +/- 2.12353
p7        = -22.3742 +/- 0.0187332
```

(c) CrystalBall + Polinômio de grau 2

```
-----
Minimizer is Minuit2 / Migrad
Chi2      =      125.415
NDF       =         21
Edm       = 1.02603e-06
NCalls    =      852
p0        = 23469.3 +/- 136.428
p1        = 90.4744 +/- 0.0109544
p2        = 2.33956 +/- 0.0170722
p3        = 0.879594 +/- 0.0154861
p4        = 1068.11 +/- 72.9181
p5        = -83121.1 +/- 5970.07
p6        = 1010.93 +/- 69.6444
p7        = 10.5942 +/- 0.727171
p8        = -0.121216 +/- 0.00830888
```

(d) CrystalBall + Polinômio de grau 3