Introdução à Análise de dados em FAE **(18/11/2024)**

Exercicios de estatística para análise de dados em HEP

Professores: Eliza Melo, Dilson Damião e Mauricio Thiel *Name:* Thiago de Andrade Rangel Monteiro

EXERCÍCIO 1

```
1
2  import ROOT
3  import glob
4  import uproot
5  import awkward as ak
6  import time
7  import matplotlib.pyplot as plt
8  import vector
9  import hist
10 import numpy as np
11 import mplhep as hep
12
13
14 diretorios = [
15     "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOD/UL2016_MiniAODv2_NanoAODv9-
           v1/100000/*.root:Events",
16     "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOD/UL2016_MiniAODv2_NanoAODv9-
           v1/1010000/*.root:Events",
17     "/opendata/eos/opendata/cms/Run2016G/DoubleEG/NANOAOD/UL2016_MiniAODv2_NanoAODv9-
           v1/250000/*.root:Events"
18 ]
19
20 variaveis = [
21     "nElectron",
22     "Electron_pt",
23     "Electron_eta",
24     "Electron_phi",
25     "Electron_mass",
26     "Electron_charge"
27 ]
28
29 arrays = uproot.concatenate(diretorios, filter_name=variaveis)
30
31 nElectron = arrays["nElectron"]
32 eletron_pt = arrays["Electron_pt"]
33 eletron_eta = arrays["Electron_eta"]
34 eletron_phi = arrays["Electron_phi"]
35 eletron_mass = arrays["Electron_mass"]
36 eletron_charge = arrays["Electron_charge"]
37
38 mask_di_eletron = nElectron == 2
39
40 eletron_p4 = vector.zip({'pt': eletron_pt,
41                          'eta': eletron_eta,
42                          'phi': eletron_phi,
43                          'mass': eletron_mass})
44 two_eletron_p4 = eletron_p4[mask_di_eletron]
```

```
1
2
3  two_eletron_charges = eletron_charge[mask_di_eletron]
4  opposite_sign_eletron_mask = two_eletron_charges[:, 0] != two_eletron_charges[:, 1]
5  two_eletron_p4 = two_eletron_p4[opposite_sign_eletron_mask]
6
7  first_eletron_p4 = two_eletron_p4[:, 0]
8  second_eletron_p4 = two_eletron_p4[:, 1]
9  di_eletron_p4 = first_eletron_p4 + second_eletron_p4
10
11 pt = di_eletron_p4.pt
12 eta = di_eletron_p4.eta
13 mass = di_eletron_p4.mass
14
15 bins = np.linspace(80, 100, 100)
16
17 di_eletron_mass_hist = hist.Hist(hist.axis.Variable(bins, label=r'$M_{e^{-}e^{+}}$ [
       GeV]'))
18 di_eletron_mass_hist.fill(mass)
19
20 pt_hist = hist.Hist(hist.axis.Regular(50, 0, 500, label=r'$p_T$ [GeV]'))
21 pt_hist.fill(pt)
22
23 eta_hist = hist.Hist(hist.axis.Regular(50, -3, 3, label=r'$\eta$'))
24 eta_hist.fill(eta)
25
26 hep.style.use('CMS')
27 fig, axs = plt.subplots(1, 3, figsize=(18, 5))
28
29 hep.histplot(di_eletron_mass_hist, histtype='fill', ax=axs[0])
30 axs[0].set_title(r"Histogram of $M_{e^{-}e^{+}}$")
31 axs[0].set_xlabel(r"$M_{e^{-}e^{+}}$ [GeV]")
32 axs[0].set_ylabel("Number of Events")
33 axs[0].set_xlim(80, 100)
34 axs[0].set_yscale('log')
35
36 hep.histplot(pt_hist, histtype='fill', ax=axs[1])
37 axs[1].set_title(r"Histogram of $p_T$")
38 axs[1].set_xlabel(r"$p_T$ [GeV]")
39 axs[1].set_ylabel("Number of Events")
40 #axs[1].set_yscale('log')
41
42 hep.histplot(eta_hist, histtype='fill', ax=axs[2])
43 axs[2].set_title(r"Histogram of $\eta$")
44 axs[2].set_xlabel(r"$\eta$")
45 axs[2].set_ylabel("Number of Events")
46 #axs[2].set_yscale('log')
47
48 total_events = int(di_eletron_mass_hist.sum())
49 print(f"Total events: {total_events}")
50
51 plt.tight_layout()
52 plt.show()
53 plt.savefig('histogram_1.png')
```

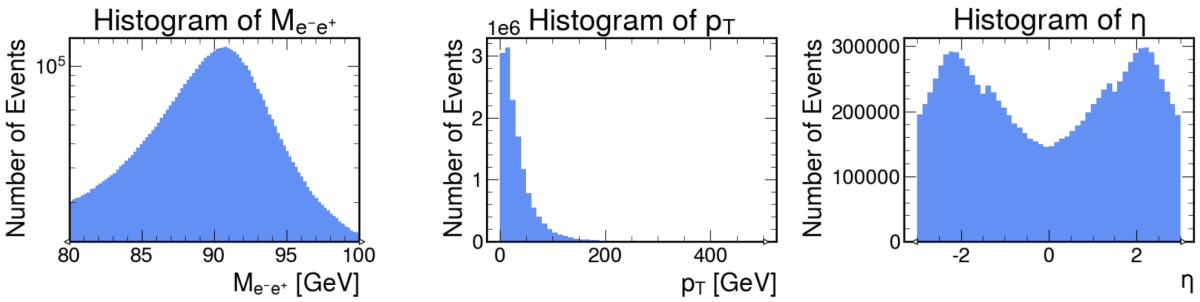

Figura 1: Plots de massa invariante, momento transverso e pseudorapidez antes do corte

Continuando agora fazendo os corte de momento transverso e pseudorapidez, nesse caso temos um corte de $p_T > 20$ Gev e $|\eta| < 2.4$ para ambos os elétrons. Esses cortes foram feitos com os mesmos critérios que no exercício anterior.

```
cut_pt_1 = first_eletron_p4.pt > 20
cut_pt_2 = second_eletron_p4.pt > 20

cut_eta_1 = first_eletron_p4.eta < np.abs(2.4)
cut_eta_2 = second_eletron_p4.eta < np.abs(2.4)

final_cut = cut_pt_1 & cut_pt_2 & cut_eta_1 & cut_eta_2

two_eletron_p4_cut = two_eletron_p4[final_cut]

first_eletron_p4 = two_eletron_p4_cut[:,0]
second_eletron_p4 = two_eletron_p4_cut[:,1]
di_eletron_p4_cut = first_eletron_p4 + second_eletron_p4

pt = di_eletron_p4_cut.pt
eta = di_eletron_p4_cut.eta
mass = di_eletron_p4_cut.mass

bins = np.linspace(80, 100, 100)

di_eletron_mass_hist = hist.Hist(hist.axis.Variable(bins, label=r'$M_{e^{-}e^{+}}$ [
    GeV]'))
di_eletron_mass_hist.fill(mass)

pt_hist = hist.Hist(hist.axis.Regular(50, 0, 500, label=r'$p_T$ [GeV]'))
pt_hist.fill(pt)

eta_hist = hist.Hist(hist.axis.Regular(50, -3, 3, label=r'$\eta$'))
eta_hist.fill(eta)

hep.style.use('CMS')
fig, axs = plt.subplots(1, 3, figsize=(18, 5))

hep.histplot(di_eletron_mass_hist, histtype='fill', ax=axs[0])
axs[0].set_title(r"Histogram of $M_{e^{-}e^{+}}$")
axs[0].set_xlabel(r"$M_{e^{-}e^{+}}$ [GeV]")
axs[0].set_ylabel("Number of Events")
axs[0].set_xlim(80, 100)
axs[0].set_yscale('log')

hep.histplot(pt_hist, histtype='fill', ax=axs[1])
axs[1].set_title(r"Histogram of $p_T$")
axs[1].set_xlabel(r"$p_T$ [GeV]")
axs[1].set_ylabel("Number of Events")
#axs[1].set_yscale('log')
```

```

hep.histplot(eta_hist, histtype='fill', ax=axs[2])
axs[2].set_title(r"Histogram of $\eta$")
axs[2].set_xlabel(r"$\eta$")
axs[2].set_ylabel("Number of Events")
#axs[2].set_yscale('log')

total_events = int(di_eletron_mass_hist.sum())
print(f"Total events: {total_events}")

plt.tight_layout()
plt.show()
plt.savefig('histogram_2.png')
```

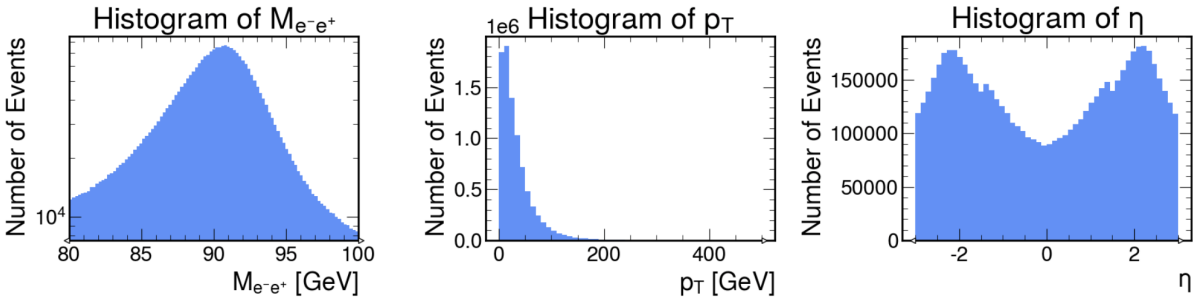

Figura 2: Plots de massa invariante, momento transverso e pseudorapidez após o corte

Agora, vamos colocar os dois plots de antes em um único canvas para poder comparar visualmente os resultados.

```
    import numpy as np
import hist
import mplhep as hep
import matplotlib.pyplot as plt

pt_no_cut = di_eletron_p4.pt
eta_no_cut = di_eletron_p4.eta
mass_no_cut = di_eletron_p4.mass

pt_cut = di_eletron_p4_cut.pt
eta_cut = di_eletron_p4_cut.eta
mass_cut = di_eletron_p4_cut.mass

bins = np.linspace(80, 100, 100)

di_eletron_mass_hist_no_cut = hist.Hist(hist.axis.Variable(bins, label=r'$M_{e^{-}e^{+}}$ [GeV]'))
di_eletron_mass_hist_no_cut.fill(mass_no_cut)

di_eletron_mass_hist_cut = hist.Hist(hist.axis.Variable(bins, label=r'$M_{e^{-}e^{+}}$ [GeV]'))
di_eletron_mass_hist_cut.fill(mass_cut)

pt_hist_no_cut = hist.Hist(hist.axis.Regular(50, 0, 500, label=r'$p_T$ [GeV]'))
pt_hist_no_cut.fill(pt_no_cut)

pt_hist_cut = hist.Hist(hist.axis.Regular(50, 0, 500, label=r'$p_T$ [GeV]'))
pt_hist_cut.fill(pt_cut)

eta_hist_no_cut = hist.Hist(hist.axis.Regular(50, -3, 3, label=r'$\eta$'))
eta_hist_no_cut.fill(eta_no_cut)

```

```
eta_hist_cut = hist.Hist(hist.axis.Regular(50, -3, 3, label=r'$\eta$'))
eta_hist_cut.fill(eta_cut)

hep.style.use('CMS')
fig, axs = plt.subplots(1, 3, figsize=(18, 5))

hep.histplot([di_eletron_mass_hist_no_cut, di_eletron_mass_hist_cut], histtype='fill', ax=axs[0], label=['No Cut', 'With Cut'], alpha=0.6)
axs[0].set_title(r"Histogram of $M_{e^{-}e^{+}}$")
axs[0].set_xlabel(r"$M_{e^{-}e^{+}}$ [GeV]")
axs[0].set_ylabel("Number of Events")
axs[0].set_xlim(80, 100)
axs[0].set_yscale('log')
axs[0].legend()

hep.histplot([pt_hist_no_cut, pt_hist_cut], histtype='fill', ax=axs[1], label=['No Cut', 'With Cut'], alpha=0.6)
axs[1].set_title(r"Histogram of $p_T$")
axs[1].set_xlabel(r"$p_T$ [GeV]")
axs[1].set_ylabel("Number of Events")
axs[1].set_yscale('log')
axs[1].legend()

hep.histplot([eta_hist_no_cut, eta_hist_cut], histtype='fill', ax=axs[2], label=['No Cut', 'With Cut'], alpha=0.6)
axs[2].set_title(r"Histogram of $\eta$")
axs[2].set_xlabel(r"$\eta$")
axs[2].set_ylabel("Number of Events")
axs[2].set_yscale('log')
axs[2].legend()

plt.tight_layout()
plt.show()
plt.savefig('histogram_comparacao.png')

total_events_no_cut = int(di_eletron_mass_hist_no_cut.sum())
total_events_cut = int(di_eletron_mass_hist_cut.sum())

print(f"Total events (no cut): {total_events_no_cut}")
print(f"Total events (with cut): {total_events_cut}")
```

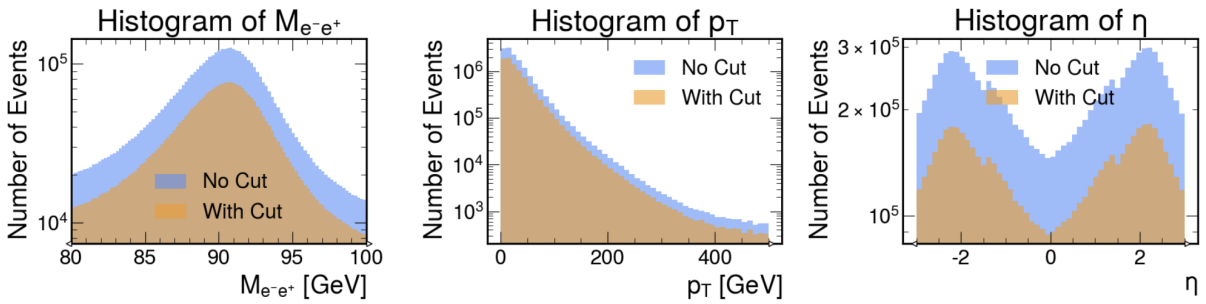

Figura 3: Plots de massa invariante, momento transverso e pseudorapidez antes e após os cortes

Temos um total de events sem corte de: 5248323 e um total events com corte de: 4630099, com isso temos aproximadamente uma redução de 13% da quantidade de eventos totais.