

Vježbe 7

Toni Šćulac

Zadatak 1

U prošlim vježbama razmatrali smo samo signalne događaje, tj. događaje u kojima se proizveo Higgsov bozon i raspao u četiri leptonu. Međutim, većina događaja koje rekonstruiramo s četiri leptonu dolaze od pozadinskih događaja gdje se nije proizveo Higgsov bozon. U novoj mapi Vježbe_7 unaprijedite kod iz prijašnjih vježbi na način da crtate četiri distribucije leptonu iz prošlih vježbi (*LepP*, *LepEta*, *LepPhi* i *LepBDT*) za signalne događaje (ggH125) i pozadinske događaje (qqZZ) na dva odvojena platna.

Zadatak 2

Nacrtajte zbroj signalnih i pozadinskih događaja za distribuciju rekonstruirane mase četiri leptonu u rasponu od 70 do 170 GeV prizivajući događaje u binovima širine 2 GeV. Koristite klasu *THStack* koja će vam omogućiti da u ukupnom histogramu naznačite kontribucije signala i pozadine. Što primjećujete na novoj distribuciji? Usporedite vaš rezultat sa zadnjim rezultatima CMS kolaboracije. Uočavate li ikakve razlike? Ako da, što mislite odakle one dolaze?

Zadatak 3

Za još bolje razlikovanje signalnih od pozadinskih događaja razvijena je varijabla koju nazivamo kinematički diskriminator i koja je definirana kao:

$$\mathcal{D}_{\text{bkg}}^{\text{kin}} \equiv \left[1 + \frac{c \cdot \mathcal{P}_{\text{bkg}}^{\text{qq}}(\vec{\Omega}^{\text{H} \rightarrow 4\ell} | m_{4\ell})}{\mathcal{P}_{\text{sig}}^{\text{gg}}(\vec{\Omega}^{\text{H} \rightarrow 4\ell} | m_{4\ell})} \right]^{-1},$$

gdje je signalna vjerojatnost spremljena pod nazivom *p_GG_SIG_ghg2_1_ghz1_1_JHUGen*, pozadinska pod nazivom *p_QQB_BKG_MCFM*, a $c = 70$ je konstanta skaliranja. Nacrtajte tu distribuciju normaliziranu na 1 za signalne i pozadinske događaje na posebnom platnu.

Zadatak 4

Receiver operating characteristic (ROC) krivulja je način prezentacije moći separacije neke varijable. Računa se na način da se na osi x prikazuje postotak pozadinskih događaja, a na osi y signalnih događaja koji prolaze odabranu vrijednost diskriminante. Npr. u našem slučaju ROC krivulja s 10 točaka bi računala koliki postotak signalnih (pozadinskih) događaja ima vrijednost manju od 0.1, 0.2, ..., 1.0 te ucrtala sve te točke na jedan graf. Unaprijedite svoj kod tako da uz distribuciju kinematičke diskriminante na podjeljenom platnu crta s desne strane i ROC krivulju. ROC krivulju prikažite kao točke u objektu klase *TGraph*. Koliko točaka je potrebno da na grafu bude vidljiva struktura krivulje?

Zadatak 5

Podjelite platno na 4 prozora i u donja dva nacrtajte dvodimenzionalne histograme gdje će jedna os predstavljati rekonstruiranu masu četiri leptonu, a druga kinematičku diskriminantu. U donji lijevi prozor nacrtajte

2D histogram za pozadinu a u donji desni za signal. Kako biste iskoristili ova četiri rezultata za statističku analizu podataka? Koje metode s predavanja bi primjenili za analiziranje stvarnih podataka i donošenje odluke da li je Higgs bozon pronađen u njima i s kojom vjerojatnosti?