Profs. Javier Duarte and R. Sekhar Chivukula
University of California San Diego
Department of Physics
PHYS 239: Modern Collider Physics
Spring 2023

Problem Set 6

Before you begin this problem set, keep in mind the following:

- **Due date is Friday, May 26, 5pm for the draft.**
- **Due date is Tuesday, May 30, 5pm for the corrections.**
- You should try each problem to the best of your ability. You can work with your class peers and consult internet resources in discussing a problem, but when writing/coding up your solution, you should not be consulting any other source specific to the problem.
- Leave space for your corrections. Do not try to cram as many solutions into as small a space as possible.

1. **DELPHES: Generator-level vs. reconstruction-level** [20 points].

   In this problem, we will investigate the effect of the parameterized detector effects in Delphes. First, generate 10,000 Drell–Yan $pp \to \mu\mu$ events at $\sqrt{s} = 13\,\text{TeV}$ assuming only standard model interactions with Pythia+MadGraph+Delphes.

   (a) Plot the momentum resolution formula for muons used in Delhpes (see `https://github.com/delphes/delphes/blob/master/cards/delphes_card_CMS.tcl#L186C1-L191`). Note that there are three separate $|\eta|$ bins. Plot all three on the same canvas from $p_T = 1$ to $100\,\text{GeV}$. What is the approximate range of momentum resolution values?

   (b) Plot the invariant mass of the two leading-p_T generator-level muons and compare it to the invariant mass of the two leading-p_T reconstructed muons. Which distribution is wider? Note, to a good approximation, we can ignore the masses of the muons, and the squared invariant mass is

   $$m^2 = p_1 p_2 (1 - \cos\theta_{12}) \, , \tag{1}$$

   where p_1 and p_2 are the momenta of the two muons and θ_{12} is the opening angle between them.

   (c) The generator-level distribution approximately follows a Breit-Wigner (or Cauchy) distribution. Fit the generator-level data to this distribution and extract the pole mass and decay width of the Z boson.

(d) The reconstruction-level distribution should approximately follow the convolution of a Breit-Wigner and a Gaussian, known as a Voigtian. Fit the reconstructor-level data to this distribution and extract the pole mass, decay width of the Z boson, and the overall effective mass resolution. Does it make sense given the formulas? Note: if the muons have momentum resolution σ_{p_i}/p_i, then we can expect the resolution on the mass to be

$$\frac{\sigma_m}{m} = \frac{1}{2}\sqrt{\left(\frac{\sigma_{p_1}}{p_1}\right)^2 + \left(\frac{\sigma_{p_2}}{p_2}\right)^2} \approx \frac{1}{\sqrt{2}}\frac{\sigma_p}{p} \tag{2}$$

in terms of an overall effective momentum resolution σ_p/p.

2. **Single-bin counting experiment and $s/\sqrt{b}$** [20 points].

In this problem, you will derive the figure of merit "$s/\sqrt{b}$" as a measure of expected discovery significance. For a single-bin counting experiment, with expected (known) signal s and background b, the likelihood is given by

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)}\,, \tag{3}$$

where μ is the signal strength (and "parameter of interest") and there are no nuisance parameters (because s and b are taken to be known). As shown in class, the corresponding maximum likelihood estimate for μ is $\hat{\mu} = (n - b)/s$.

For the usual case where the signal models correspond to positive μ, one may test the $\mu = 0$ (background-only) hypothesis with the test statistic q_0 where[1]

$$q_\mu = \begin{cases} -2\ln\dfrac{L(\mu)}{L(\hat{\mu})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{4}$$

(a) In the large-sample limit [1], the discovery significance is given by $Z_0 = \sqrt{q_0}$. Assuming this, show that

$$Z_0 = \sqrt{2\left(n\ln\frac{n}{b} + b - n\right)} \tag{5}$$

(b) To approximate the median expected discovery significance $\mathbb{E}(Z_0|\mu = 1)$, we can simply replace n with the expected value $\mathbb{E}(n|\mu = 1) = s + b$ (the so-called "Asimov data set"). Show that if you make this replacement and take the limit $s \ll b$, you arrive at

$$\mathbb{E}(Z_0|\mu = 1) \approx \frac{s}{\sqrt{b}} \tag{6}$$

[1] This modified definition for $\hat{\mu} < 0$ is to prevent the rejection of the $\mu = 0$ hypothesis because of a downward fluctuation of the data. If the data fluctuate such that one finds fewer events than even predicted by background processes alone, then $\hat{\mu} < 0$ and one has $q_0 = 0$. As the event yield increases above the expected background, i.e., for increasing $\hat{\mu}$, one finds increasingly large values of q_0, corresponding to an increasing level of incompatibility between the data and the $\mu = 0$ hypothesis.

3. **Final Project Proposal** [30 points]

   Propose your final project in one or two paragraphs. The basic idea is to generate and simulate a new physics model and approximately estimate the sensitivity with a simple data analysis at the LHC.

   - The physics motivation (what research question is the new physics model trying to solve?);
   - Description of the new physics model (what new particles or interactions?) and resulting experimental signature;
   - Which software tools you will use and why; and
   - How the event selection/optimization/sensitivity estimate will be carried out (e.g., what backgrounds will be considered?)

References

[1] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: Eur.Phys.J.C 73, 2501 (2013)].