Final Presentation

CERN-Korean Summer Student Program

Juhoon Son[1]

[1]Department of Astronomy, Yonsei University

2023-07-07

Table of contents

Look Elsewhere Effect

Example

Suppose you throw a coin 10 times, and you've got 10 heads.

- It's very unusual.
- Can you quantify how unusual this result is?

In particular, can you say the probability for this kind of peculiarity happening is $(1/2)^{10}$?

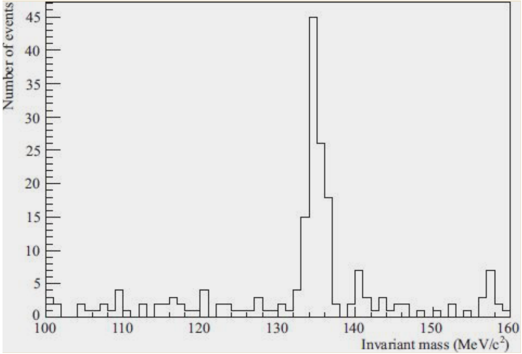
If not, what must then be the correct answer?

Example

Larkoski *"Elementary Particle Physics"*, Exercise 5.6

- Often in experimental particle physics it is known where to look for a signal. For example, if you want to study the properties of the Z boson, then you can tune the collision energy of your e^+e^- collider to the mass of the Z boson.
- However, when searching for possible new physics, where the mass scale is unknown, experimentalists look for excesses over many bins. Purely as a result of finite statistics, one will find excesses when enough bins are considered.
- Thus, the significance of an excess in any one bin is reduced, simply because that excess could have been anywhere. This is called the look-elsewhere effect.

$\pi^0 \to e^+e^-\gamma$



- 100 signal & 100 bkg events are generated over [100 MeV, 160 Mev].
- Histogram in 60 bins.

(a) local vs. global significance

- Assume you are looking for excesses in a collection of N_{bins} bins of data. Let's also assume that you only think that excesses are interesting if they deviate by more than $X\sigma$ from the expected number of events in a bin.
- If the probability for any one bin to have an excess of at least $X\sigma$ is p_X, **determine the probability that at least one of the N_{bins} bins has such an excess.** In this problem, only assume that the different bins are independent; don't assume anything in particular about the probability p_X.

(a) local vs. global significance

Solution

$$
\begin{aligned}
P(\text{at least 1 bin with excess}) &= 1 - P(\text{no bins with excess}) \\
&= 1 - (1 - p_X)^{N_{\text{bins}}}.
\end{aligned} \tag{1}
$$

(b) in the limit $p_X \to 0$

- Now, expand your result from part (a) to lowest order in the limit where $p_X \to 0$. How much larger is the probability for at least one bin to have an excess than p_X?
- Find, in this limit, the probability p_X^{global} of an excess at least as large as $X\sigma$ anywhere in your data.

Solution

$$P_1 = p_X^{\text{global}} = \lim_{p_X \to 0} \left[1 - (1 - p_X)^{N_{\text{bins}}}\right] = N_{\text{bins}} p_X. \tag{2}$$

(c) case for evidence?

- Excesses are considered "interesting" or "evidence" if they are at least a 3σ deviation from the expected value from the null hypothesis.
- If the local significance (the significance in one bin) of an excess is 3σ, what is the global significance, which includes the *look-elsewhere effect* (LEE)? Assume a number of bins to be 100.
- Including the LEE, do you think such an excess is still interesting?

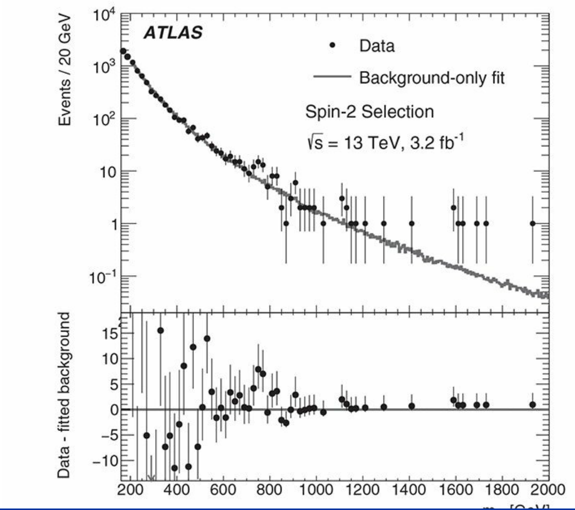
Solution

For 3σ excess, $P_3 = \frac{1}{2}(1 - \mathrm{erf}(\frac{3}{\sqrt{2}})) = 0.13\%$.

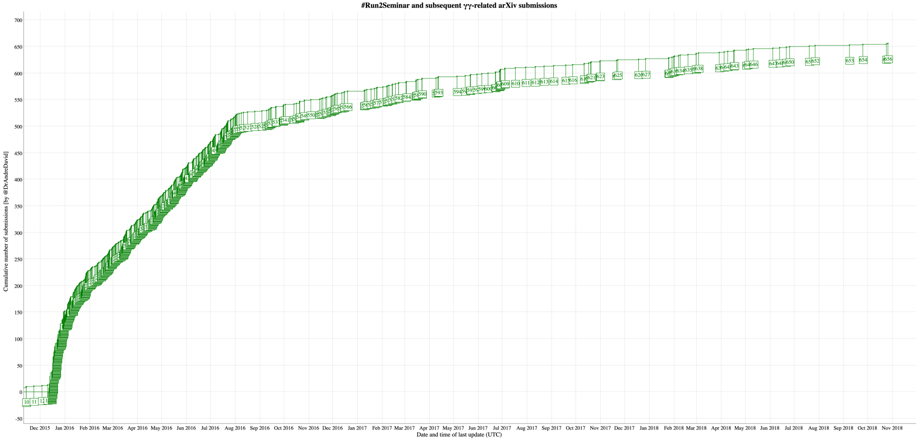
With $N_{\mathrm{bins}} = 100$, $P_3^{\mathrm{global}} \approx 100 \times 0.13 = 13\%$ (Not interesting)

cf. $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$

Diphoton (or digamma) event at $M_{\gamma\gamma} \approx 750$ GeV on December 15, 2015

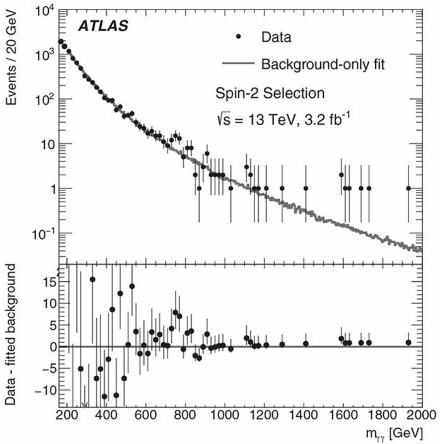


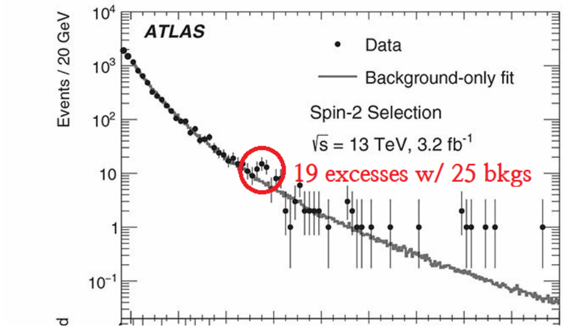
Diphoton (or digamma) event at $M_{\gamma\gamma} \approx 750$ GeV on December 15, 2015



(d) the $X(750)$?

- Estimate the local significance σ_{local} of the $\gamma\gamma$ excess.
- To do this, use the three bins that range over 720 – 780 GeV to calculate the significance and ignore the look-elsewhere effect.
- Would you be interested in such an excess?





Solution

In the 3 bins (720-780 GeV), $\exists \sim 19$ excess events. From the main plot, $N_{\rm bkg} \simeq 25$ events.

$\Rightarrow \delta N_b \simeq \sqrt{25} = 5.$

$\Rightarrow 19/5 = 3.8\sigma$ (very interesting!)

(d) the $X(750)$?

- Estimate the global significance σ_{global} of the excess. Use the same bins as in previous part and now include the LEE.
- To include the LEE, you will need to determine how many sets of three neighboring bins there are in these data. You can safely use the approximation you derived in part (b). Would you be interested in such an excess?

(d) the $X(750)$?

Solution

For global significance, we need to determine how many possibilities of '3-bin sets' in the measurement. Between 200 and 1200 GeV, $\exists$ 50 bins $\Rightarrow$ 48 cases of 3-bin sets.

$\therefore$ Including LEE,

$$P_{3.8}^{\text{global}} \approx 48 \times 7.2 \times 10^{-5} = 0.35\% > 0.13\% \quad \text{(not interesting)}$$

(d) the $X(750)$?

- Assuming that this excess is just a statistical fluctuation of the null hypothesis, how many more events need to be added to these three bins to reduce the local significance of the excess to 1σ?

Solution

To make 19 event excess be 1σ excess we need to have $N_{\text{bkg}} \simeq 19^2$events $= 361$ events.
Currently, in the plot, we already have $19 + 25 = 44$ events.
$\therefore$ We need $361 - 44 = 317$ more events.

맺음말